PRIME NUMBER THEOREM FOR REGULAR TOEPLITZ SUBSHIFTS

KRZYSZTOF FRĄCZEK, ADAM KANIGOWSKI, AND MARIUSZ LEMAŃCZYK

Abstract. We prove that neither a prime nor a semiprime number theorem holds in the class of regular Toeplitz subshifts but when the regularity (with respect to the periodic structure) is strengthened to the regularity with respect to the values of the Euler totient function of the periods then they both do hold.

1. Introduction

Given a topological dynamical system $(X,T)$, where $T$ is a homeomorphism of a compact metric space $X$, one says that a prime number theorem (PNT) holds if the limit

$$
\lim_{N \to \infty} \frac{1}{\pi(N)} \sum_{p \leq N} f(T^p x)
$$

(\(p\) stands always for a prime number) exists for each $x \in X$, an arbitrary $f \in C(X)$ and $\pi(N)$ denotes the number of primes up to $N$. In fact, then, via Riesz theorem, for all $f \in C(X)$, we have

$$
\lim_{N \to \infty} \frac{1}{\pi(N)} \sum_{p \leq N} f(T^p x) = \int_X f \, d\nu_x,
$$

for a probability (Borel) measure $\nu_x$ on $X$, where $\nu_x$ depends only on $x \in X$.

Let us first consider the cyclic case: $X = \mathbb{Z}/k\mathbb{Z}$ and $Tx = x + 1$. Fix $x \in X$ and notice that (1) indeed holds by the classical prime number theorem in arithmetic progressions, where $\nu_x$ is the uniform probability measure on the “coset” $\{a < k : (a,k) = 1\} + x$. Hence, a PNT holds in cyclic (hence, in finite) systems.

If instead of cyclic systems we consider the procyclic case, that is, an odometer system $(H,T)$:

$$
H = \lim_{t \to \infty} \mathbb{Z}/n_t \mathbb{Z}, \quad Tx = x + (1,1,\ldots)
$$

(here $n_t|n_{t+1}$ for $t \geq 0$) then a PNT still holds, but a reason for it is that, compared to the finite case, we did not add too many continuous functions. Indeed, in the odometer, there is a natural partition $D^t = (D^t_0, \ldots, D^t_{n_t - 1})$, $TD^t_i = D^t_{i+1} \mod n_t$, consisting of closed (hence clopen) sets, having the same diameter (as $T$ is an isometry) and the diameter is going to 0 when $t \to \infty$ (reflecting the fact that the levels of the towers $D^t$ form a basis of topology on $H$). Hence, each $f \in C(H)$ can be approximated uniformly by functions which are constant on the levels of towers $D^t$ (it is not hard to see that each character of the group $H$ is constant on the levels of the towers $D^t$ for $t$ sufficiently large) and a PNT holds because it does in the finite case.

2000 Mathematics Subject Classification. 37B10, 37A45, 11N05, 11N13.

Key words and phrases. prime number theorem, Toeplitz systems, semi-prime numbers, non-conventional ergodic theorems.

Research supported by Narodowe Centrum Nauki grant 2019/33/B/ST1/00364.
We will deal with extensions of odometers but before we describe our results toward PNT, let us first discuss basic topological properties of odometers. These are zero entropy topological systems which are minimal (all $T$-orbits are dense) and uniquely ergodic (there is only one $T$-invariant measure - Haar measure in this case). But not so much can be said about a PNT for a general uniquely ergodic system: while (1) holds a.e. with respect to the unique invariant measure $\mu$, one can easily construct a counterexample to the validity of (1) for all $x \in X$. Indeed, denote by $\mathbb{P}$ the set of prime numbers and consider the left shift $S$ on $\{0,1\}^\mathbb{Z}$ and the subshift $(X_{1 \mathbb{P} \to \mathbb{P}} , S)$ obtained by the orbit closure of the characteristic function $1_{\mathbb{P} \to \mathbb{P}}$ of the “symmetrized” primes. It has a unique invariant measure (which is the Dirac measure at the fixed point $\ldots 0.00 \ldots$) and a PNT fails in it (see e.g. [8] for details). Now, this particular uniquely ergodic model of the one-point system implies paradoxically that each ergodic dynamical system has a uniquely ergodic model $(X,T)$ in which a PNT does not hold. To see this, take any uniquely ergodic model $(Y,\nu, R)$ of the given measure-theoretic dynamical system. Since the one-point system is (Furstenberg) disjoint with any other system, the product system $(X_{1 \mathbb{P} \to \mathbb{P}} \times Y, S \times R)$ is still uniquely ergodic, with the unique invariant measure $\delta_{\ldots 0.00 \ldots} \otimes \nu$. It is not hard to see that the product measure-theoretic system is still measure-theoretically isomorphic to the original system. Since the new system has $(X_{1 \mathbb{P} \to \mathbb{P}} , S)$ as its topological factor, a PNT does not hold in $(X_{1 \mathbb{P} \to \mathbb{P}} \times Y, S \times R)$. Hence, if we think about a necessary condition for a PNT to hold, it looks reasonable to add the minimality assumption to avoid a problem of “exotic” orbits on which PNT does not hold (we also recall that a uniquely ergodic system has a unique subsystem which is strictly ergodic). In this class however still one can produce counterexample to a PNT, see [18] for the first symbolic (counter)examples (though the entropy is not determined in examples in [18]), or more recent non-symbolic (counter)examples from [18]. On the other hand, we have quite a few classes in which a PNT holds, including systems of algebraic origin [11], [21], symbolic systems [4], [10], [15], [16] or recently [13] in the category of smooth systems, where a PNT has been proved in the class of analytic Anzai skew products. Finding a sufficient dynamical condition for a PNT to hold, postulated a few years ago by P. Sarnak [20] seems to be an important and difficult task in dynamics, however we rather expect the following:

**Working Conjecture:** Each ergodic and aperiodic measure-theoretic dynamical system has a strictly ergodic model in which a PNT fails

which, if true, makes Sarnak’s postulate even harder to realize. The present paper should be viewed as introductory steps in trying to understand the conjecture.

A PNT can be reformulated as the existence of a limit of

$$\frac{1}{N} \sum_{n<N} f(T^n x) \Lambda(n),$$

where $\Lambda$ stands for the von Mangoldt function: $\Lambda(p^\ell) = \log p$ for $\ell \geq 1$ and 0 otherwise. In the last decade, the PNT problem in dynamics goes in parallel with a
related to it Sarnak’s conjecture [19] on Möbius orthogonality:

\[ \lim_{N \to \infty} \frac{1}{N} \sum_{n < N} f(T^n x) \mu(n) = 0 \]

for each \( x \in X \), \( f \in C(X) \) in each \textbf{zero entropy} dynamical system \((X, T)\) \( \mu \) stands for the Möbius function: \( \mu(1) = 1 \), \( \mu(p_1 \ldots p_k) = (-1)^k \) for different primes \( p_1, \ldots, p_k \), and \( \mu(n) = 0 \) for the remaining \( n \in \mathbb{N} \). Here, the class of objects for which we expect the positive answer is precisely defined. In fact, in quite a few cases: cf. [4], [5], [9], [10], [11], [15], [16], one can observe the following principle: once we can prove Sarnak’s conjecture for \( p \), cf. [4], [5], [9], [10], [11] [15], [16], one can observe the following principle: once we can prove Sarnak’s conjecture for \((X, T)\) with a “sufficient” speed of convergence to zero in (2) then a PNT holds in \((X, T)\).

With all the above in mind we come back to extensions of odometers we intend to study. We stay in the \textbf{zero entropy} category of systems and we assume \textbf{minimality}. Further, we assume that the systems are \textbf{almost 1-1 extensions} of odometers.\(^5\) We also assume that our systems are \textbf{symbolic} \(^3\). All these natural assumptions determine however a very precise class of topological systems, namely Toeplitz subshifts \((X_x, S)\), where \( x \) is a Toeplitz sequence over a finite alphabet \( A \), see Downarowicz’s survey [7], Section 7. That is, \( x \in A^\mathbb{Z} \) and for \( a \in A \) there is \( \ell \in \mathbb{N} \) such that \( x(a) = x(a + k\ell) \) for each \( k \in \mathbb{Z} \) and \( X_x \) stands for the set of all \( y \in A^\mathbb{Z} \) whose each subblock appears also in \( x \). One shows then that there is a sequence \( n_t|n_{t+1} \) such that if \( \text{Per}_{n_t}(x) := \{ a \in \mathbb{Z} : x(a) = x(a + kn_t) \} \) for each \( k \in \mathbb{Z} \) then

\[ \bigcup_{t \geq 0} \text{Per}_{n_t}(x) = \mathbb{Z}. \]

Moreover, there is a natural continuous factor map \( \pi : X_x \to H \), where \( H \) stands for the odometer determined by \( (n_t) \). In fact, we will still restrict the class of Toeplitz subshifts to so called \textit{regular} Toeplitz subshifts, whose formal definition is that the density of \( \bigcup_{t=0}^M \text{Per}_{n_t}(x) \) is going to 1. Regular Toeplitz subshifts are \textbf{zero entropy} strictly ergodic systems, measure-theoretically isomorphic to rotations given by the maximal equicontinuous factors. Although in [7] there are four other equivalent conditions for regularity (see Theorem 13.1 in [7]), we will choose a different path. Namely, first notice that since \( \pi : X_x \to H \) is a \textbf{continuous and equivariant surjection},

\[ E_t^i := \pi^{-1}(D_t^i) = (E_0^i, \ldots, E_{n_t-1}^i) \text{ with } E_j^i = \pi^{-1}(D_j^i) \]

is an \( S \)-tower of height \( n_t \) whose level are closed (hence clopen). By the minimality of \((X_x, S)\) there is a unique tower with clopen levels and of fixed height. Let us consider on \( A^\mathbb{Z} \) a metric inducing the product topology being given by

\[ d(x, y) = 2^{-\inf(|n| : x(n) \neq y(n))}. \]

The diameters of the levels of towers \( E_t^i \) do \textbf{not} converge to zero, unless \( x \) is periodic. Moreover, the diameters of different levels are in general different as the shift \( S \) is not an isometry. Let us consider the diameter of the tower \( E_t^i \) given by:

\[ \delta(E_t^i) := \sum_{0 \leq j < n_t} \text{diam}(E_j^i). \]

\(^3\)If \((H, T)\) is a factor of \((X, S)\) via \( \pi : X \to H \), then \((X, S)\) is called an \textit{almost 1-1 extension} of \((H, T)\) if there is a point \( h \in H \) such that \( |\pi^{-1}(h)| = 1 \); in fact, in this case the set of points with singleton fibers is \( G_\delta \) and dense.

\(^4\)We recall that each zero entropy system has an extension which is symbolic [2], and clearly if a PNT holds for a system, it does for a factor.
It is not hard to see (see Appendix A) that the regularity of a Toeplitz sequence is equivalent to

\[
\lim_{t \to \infty} \frac{\delta(E^t)}{n_t} = 0.
\]

It is also not hard to see that this property does not depend on the choice of \((n_t)\) satisfying (3). We recall that the Möbius orthogonality of subshifts given by regular Toeplitz sequences has been proved in [1]. Here are two first results of the paper proved in Section 2 and Section 4, respectively:

**Theorem A.** A PNT does not hold in the class of minimal almost 1-1 symbolic extensions satisfying (1) of odometers. That is, a PNT need not hold in a strictly ergodic subshift determined by a regular Toeplitz sequence.

**Theorem B.** A PNT holds in the class of minimal almost 1-1 symbolic extensions of odometers in which (1) holds with a speed:

\[
\lim_{t \to \infty} \frac{\delta(E^t)}{\varphi(n_t)} = 0,
\]

where \(\varphi\) denotes the Euler totient function.

As for all Toeplitz dynamical systems constructed in the proof of Theorem A we have

\[
0 < \liminf_{t \to \infty} \frac{\delta(E^t)}{\varphi(n_t)} \leq \limsup_{t \to \infty} \frac{\delta(E^t)}{\varphi(n_t)} < +\infty,
\]

the condition (5) in Theorem B is optimal to have a PNT.

We then turn our attention to a semi-prime number theorem (SPNT) which is much less explored than the PNT case and which, for the first time in dynamics, is studied in [1] (for some smooth Anzai skew products). In Section 3 and Section 5 we provide sketches of proofs of the exact analogues of Theorems A and B for an SPNT for regular Toeplitz subshifts.

In Section 6.1 we prove a new non-conventional ergodic theorem:

\[
\lim_{N \to \infty} \frac{1}{N} \sum_{n \leq N} f(S^{P(n)}x) \text{ exists}
\]

for monic polynomials \(P\) with positive integer coefficients for all symbolic minimal almost 1-1 extensions of odometers with a modified condition (4). In Section 6.2 we provide a regular Toeplitz subshift which does not satisfy the non-conventional ergodic theorem for squares but it satisfies a PNT.

While Theorem A confirms Working Conjecture for a subclass of odometers, we have been unable to confirm it for the whole class of odometers. Confirming Working Conjecture for the class of automorphisms with discrete spectrum seems to be the first step toward a possible general statement. In Appendix B, we provide a simple argument showing that in the classical symbolic models of irrational rotations given by Sturmian sequences, a PNT holds. The Sturmian systems are strictly ergodic and are almost 1-1 extensions of irrational rotations.
2. Regular Toeplitz subshifts which do not satisfy PNT (proof of Theorem A)

For all $K, n \in \mathbb{N}$ and $a \in \mathbb{Z}$ let
\[ \pi(K; n, a) = \{1 \leq p \leq K : p \in \mathbb{P}, p = a \mod n\}. \]

**Theorem 2.1** (Dirichlet). For any natural $n$ and any integer $a$ with $(a, n) = 1$ we have
\[ \lim_{K \to \infty} \frac{\pi(K; n, a)}{\pi(K)} = 1. \]

We construct a Toeplitz sequence $x \in \{0, 1\}^\mathbb{Z}$ with the period structure $(n_t)$:
\[ n_{t+1} = k_{t+1}n_t, \quad (k_{t+1}, n_t) = 1 \]
for each $t \geq 1$. We will show that for this $x$:
\[ \lim_{t \to \infty} \frac{1}{\pi(n_t)} \sum_{p < n_t} F(S^p x) \text{ does not exist,} \]
where $F(y) = (-1)^{y(0)}$. At stage $t$, $x$ is approximated by the infinite concatenation of $x_t[0, n_t - 1] \in \{0, 1, ?\}^n$ (that is, we see a periodic sequence of $0, 1, ?$ with period $n_t$). Successive "?" will be filled in in the next steps of construction of $x$. We require that:
\[ \frac{\varphi(n_t)}{n_t} \leq \frac{1}{2^{t+1}}, \]
\[ \{0 \leq i < n_t : x_t(i) = ?\} \subset \{0 \leq j < n_t : (j, n_t) = 1\}, \]
\[ \#\{0 \leq i < n_t : x_t(i) = ?\} \geq \left(1 - \sum_{l=1}^{t} \frac{1}{100^l}\right) \varphi(n_t), \]
\[ \#\{p < n_t : x_t(p) = ?\} \geq \frac{1}{2} \varphi(n_t). \]

We choose $k_{t+1}$ satisfying (11) and:
\[ \frac{\varphi(k_{t+1})}{k_{t+1}} \leq \frac{1}{2}, \]
\[ \varphi(k_{t+1}) \geq 100t+1, \]
\[ 8 \log n_{t+1} \leq \pi(n_{t+1}), \quad 8\pi(n_t) \leq \pi(n_{t+1}) \]
and for each $0 < a < n_t$, $(a, n_t) = 1$, we have
\[ \#\{a + jn_t : j = 0, \ldots, k_{t+1}\} \cap \mathbb{P} = \pi(n_{t+1}; n_t, a) \leq 2 \frac{\pi(n_{t+1})}{\varphi(n_t)}. \]

The latter we obtain from the Dirichlet theorem (remembering that $n_t$ is fixed, so the number of $a$ is known, we can obtain the accuracy as good as we want by taking $k_{t+1}$ sufficiently large).

We need two simple observations:
\[ \text{Note that if } p_i \text{ stand for the } i \text{-th prime then } \sum_{j=1}^{s} 1/p_j = +\infty, \text{ whence remembering that } \varphi(p_1p_2 \cdots p_{s+1}) = p_1p_2 \cdots p_{s+1} \prod_{j=0}^{s} (1 - 1/p_{i+j}), \text{ we have } \prod_{j=0}^{s} (1 - 1/p_{i+j}) \to 0, \text{ and therefore } \prod_{j=0}^{s} (1 - 1/p_{i+j}) < 1/2 \text{ for } s \text{ large enough.} \]
that the known number which we had from stage
\[ \{0 \leq k < n_{t+1} : (k, n_{t+1}) = 1\} \subset \bigcup_{0 \leq a < n_t} \{a + jn_t : j = 0, \ldots, k_{t+1} - 1\}. \]

Lemma 2.2. For every \(0 \leq a < n_t\) with \((a, n_t) = 1\), we have
\[ \#\{0 \leq j < k_{t+1} : (a + jn_t, n_{t+1}) = 1\} = \varphi(k_{t+1}). \]

Proof. First note that \((a + jn_t, n_{t+1}) = 1\) if \((a + jn_t, k_{t+1}) = 1\). Indeed, assume that
\((a + jn_t, k_{t+1}) = 1\). If for some prime \(p\) we have \(p|(a + jn_t)\) and \(p|n_{t+1} = nt_{k_t+1}\), then \(p|k_{t+1}\). Otherwise, we have \(p|n_t\), so \(p|a\). As \((a, n_t) = 1\), this gives a contradiction.

Thus \((a + jn_t, k_{t+1}) = 1\) implies \((a + jn_t, n_{t+1}) = 1\). The opposite implication is obvious. Thus
\[ \{0 \leq j < k_{t+1} : (a + jn_t, n_{t+1}) = 1\} = \{0 \leq j < k_{t+1} : (a + jn_t, k_{t+1}) = 1\}. \]

Let us consider the affine map
\[ \mathbb{Z}/k_{t+1}\mathbb{Z} \ni j \mapsto a + jn_t \in \mathbb{Z}/k_{t+1}\mathbb{Z}. \]
If \(J := \{0 \leq \ell < k_{t+1} : (\ell, k_{t+1}) = 1\}\) then
\[ \{0 \leq j < k_{t+1} : (a + jn_t, k_{t+1}) = 1\} = A^{-1}(J). \]
Since \((n_t, k_{t+1}) = 1\), the map \(A\) is a bijection. It follows that
\[ \#\{0 \leq j < k_{t+1} : (a + jn_t, k_{t+1}) = 1\} = \#\{0 \leq \ell < k_{t+1} : (\ell, k_{t+1}) = 1\} = \varphi(k_{t+1}), \]
which completes the proof. \(\Box\)

We need to describe now which and how we fill "?" in \(x_{t+1}[0, n_{t+1} - 1]\). This block is divided into \(k_{t+1}\) subblocks
\[ x_{[0, n_t - 1]}x_{[0, n_t - 1]} \cdots x_{[0, n_t - 1]} \]
We fill in all "?" in the first block \(x_{[0, n_t - 1]}\) in such a way to “destroy” a PNT for the time \(n_t\), namely
\[ \frac{1}{\pi(n_t)} \sum_{p < n_t} F(S^p x) = \frac{1}{\pi(n_t)} \sum_{p < n_t} (-1)^{x(p)} + \]
\[ \frac{1}{\pi(n_t)} \left( \sum_{p < n_t} \frac{1}{\pi(n_t)} \left( 1 - \sum_{(p, n_t) = 1} \frac{1}{x_i(p) = 0} \sum_{(p, n_t) = 1} \frac{1}{x_i(p) = 1} \right) \right). \]

As the number of the primes dividing \(n_t\) is bounded by \(\log n_t\), it is negligible compared to \(\pi(n_t) = n_t/\log n_t\). It follows that
\[ \left| \frac{1}{\pi(n_t)} \sum_{p < n_t} (-1)^{x(p)} \right| \leq \frac{\log n_t}{\pi(n_t)} = o(1), \]
so the first summand does not affect the asymptotic of the averages in (11). Since the number of \(p\) in the last summand is at least \(1/\pi(n_t)\) in view of (111), we can fill in \(x_{t+1}\) at places \(\{p < n_t : (p, n_t) = 1, x_i(p) = ?\}\) to obtain the sum completely different that the known number which we had from stage \(t - 1\).
We fill in (in an arbitrary way) all remaining places between 0 and \( n_t - 1 \) and all places \( a + jn_t \) for \( 0 \leq j < k_{t+1} \) such that this number is not coprime with \( n_{t+1} \), so that (5) will be satisfied at stage \( t + 1 \). We must remember that for certain \( 0 < a < n_t \) coprime to \( n_t \), \( x_t(a) \) was already defined at previous stages, so along the corresponding arithmetic progressions \( a + jn_t \), \( 0 \leq j < k_{t+1} \), these places are also filled in previously. On the other side, if \( x_{t+1}(a + jn_t) \neq \) (that is, \( x_{t+1}(a + jn_t) = 0 \) or \( x_{t+1}(a + jn_t) = 1 \)) and \( (a + jn_t, n_{t+1}) = 1 \) for some \( 0 < j < k_{t+1} \) then \( x_t(a) \neq \).

By (8), Lemma 2.2, (9) and (12), it follows that

\[
\begin{align*}
\{0 \leq i < n_{t+1} : (i, n_{t+1}) = 1, x_{t+1}(i) \neq \} \\
\subset \{0 < a < n_t : (a, n_t) = 1, x_t(a) \neq \} \\
\cup \{a + jn_t : 0 < j < k_{t+1}, (a + jn_t, n_{t+1}) = 1\}.
\end{align*}
\]

By (8), Lemma 2.2, (9) and (12), it follows that

\[
\begin{align*}
\#\{0 \leq i < n_{t+1} : (i, n_{t+1}) = 1, x_{t+1}(i) \neq \} \\
\leq \varphi(n_t) + \#\{0 \leq a < n_t : (a, n_t) = 1, x_t(a) \neq \} \varphi(k_{t+1}) \\
\leq \varphi(n_t) + \left( \sum_{k=1}^{t} \frac{1}{100^k} \right) \varphi(n_t) \varphi(k_{t+1}) = \left( \frac{1}{\varphi(k_{t+1})} \right) + \sum_{k=1}^{t} \frac{1}{100^k} \varphi(n_t) \\
\leq \sum_{k=1}^{t+1} \frac{1}{100^k} \varphi(n_{t+1}) \leq \frac{1}{99} \varphi(n_{t+1})
\end{align*}
\]

In particular, at stage \( t + 1 \), also (9) is satisfied.

Similar arguments show that

\[
\begin{align*}
\{p < n_{t+1} : x_{t+1}(p) \neq \} \subset \{p < n_{t+1} : p \mid n_{t+1} \} \cup \{p < n_t : x_{t+1}(p) \neq \} \\
\cup \{a + jn_t : 0 < j < k_{t+1}, a + jn_t \in \mathbb{P}\}.
\end{align*}
\]

In view of (14), (9) and (13), it follows that

\[
\begin{align*}
\#\{p < n_{t+1} : x_{t+1}(p) \neq \} \\
\leq \log n_{t+1} + \pi(n_t) + 2 \#\{0 \leq a < n_t : (a, n_t) = 1, x_t(a) \neq \} \frac{\pi(n_{t+1})}{\varphi(n_t)} \\
\leq \log n_{t+1} + \pi(n_t) + \frac{2}{99} \varphi(n_t) \frac{\pi(n_{t+1})}{\varphi(n_t)} \leq \frac{1}{2} \pi(n_{t+1}).
\end{align*}
\]

Therefore, at stage \( t + 1 \), also (10) is satisfied.

Finally, note that

\[
\frac{\varphi(n_{t+1})}{n_{t+1}} = \frac{\varphi(n_t) \varphi(k_{t+1})}{n_t k_{t+1}} \leq \frac{\varphi(n_t)}{n_t} \frac{1}{2},
\]

so (7) holds and the resulting Toeplitz sequence is regular.
3. Toeplitz subshifts for which an SPNT does not hold

We now intend to give an example of a (regular) Toeplitz sequence \( x \) such that an SPNT does not hold for the corresponding subshift, in fact:

\[
\lim_{t \to \infty} \frac{1}{\pi_2(n_t)} \sum_{p_1 p_2 < n_t} F(S^{p_1 p_2} x) \text{ does not exist,}
\]

where \( \pi_2(N) \) stands for the number of semiprimes \( < N \). We recall that \( \pi_2(N) \sim \frac{N \log \log N}{\log N} \). We also recall briefly that if we look for semiprime numbers in arithmetic progressions \( a + m \mathbb{Z} \) then (for our purposes), we may assume that \( (a, m) = 1 \). Indeed, if \( (a, m) \in \mathbb{S} \mathbb{P} \) (or has more than two prime divisors) then in \( a + m \mathbb{N} \) we can have at most one semiprime number. If \( (a, m) \in \mathbb{P}, \) say \( (a, m) = p \), then the number of semiprimes \( \leq N \) in \( a + m \mathbb{N} \) is the same as the number of primes \( \leq N/p \) in \( a + (m/p) \mathbb{Z} \). Thus (as \( m \) is fixed)

\[
\sum_{(a, m) > 1} \sum_{p_1 p_2 \equiv a \pmod{m} \atop N \log \log N} \frac{N/p}{\varphi(m)} = o\left( \pi_2(N) \right)
\]

when \( N \to \infty \) (with “\( o \)” which depends on \( m \)).

Now, suppose that \( (a, m) = 1 \). Assume that \( p_1, p_2 \) are prime numbers such that \( p_1 p_2 \leq N, p_1 p_2 = a \pmod{m} \) and \( p_1 \leq p_2 \). Then \( p_1 \leq \sqrt{N} \). Since \( (p_1, m) = 1 \), if \( a(p_1) < m \) there exists a unique \( 0 \leq a(p_1) < m \) such that \( p_1 \cdot a(p_1) = a \pmod{m} \) and \( (a(p_1), m) = 1 \). Then

\[
\sum_{p_1 p_2 \equiv a \pmod{m} \atop p_1 p_2 \leq N} 1 = \sum_{p_1 \leq \sqrt{N}} \sum_{p_2 \equiv a(p_1) \pmod{m} \atop p_2 \leq N} 1 = \sum_{p_1 \leq \sqrt{N}} \pi([N/p_1]; m, a(p_1)).
\]

As \( p_1 \leq \sqrt{N} \) implies \( N/p_1 \geq \sqrt{N} \), in view of Theorem 2.1 for every \( \varepsilon > 0 \) there exists \( N_\varepsilon \) such that for all \( N \geq N_\varepsilon \) and \( p_1 \leq \sqrt{N} \) with \( (p_1, m) = 1 \), we have

\[
(1 - \varepsilon) \frac{\pi([N/p_1])}{\varphi(m)} < \pi([N/p_1]; m, a(p_1)) \leq (1 + \varepsilon) \frac{\pi([N/p_1])}{\varphi(m)}.
\]

Since \( \pi_2(N) = \sum_{p_1 \leq \sqrt{N}} \pi([N/p_1]) \), it follows that

\[
(1 - \varepsilon) \frac{\pi_2(N)}{\varphi(m)} < \sum_{p_1 p_2 \equiv a \pmod{m} \atop p_1 p_2 \leq N} 1 < (1 + \varepsilon) \frac{\pi_2(N)}{\varphi(m)}
\]

for every \( N \geq N_\varepsilon \). Therefore

\[
\sum_{p_1 p_2 \equiv a \pmod{m} \atop p_1 p_2 \leq N} 1 \sim \frac{\pi_2(N)}{\varphi(m)}.
\]

Now, we repeat the scheme of the previous construction almost word for word, although we have to take care how to choose \( k_{t+1} \).

First of all, we require that \( k_{t+1} \) is large enough so that (16) and (17) hold (with \( m = n_t \)). So, we replace (10) by

\[
\# \{ p_1 p_2 < n_t: x_t(p_1 p_2) = ? \} \geq \frac{1}{2} \pi_2(n_t)
\]
and requiring (instead of (13)) that for \((a, n_t) = 1\), we have

\[ \# \{a + jn_t : 0 \leq j < k_{t+1} \} \cap \mathbb{P} \leq 2 \frac{\pi_2(n_{t+1})}{\varphi(n_t)}. \]

Furthermore, we replace (13) by the requirement that

\[ \sum_{a,n_t \geq 1 \atop p_1 p_2 \equiv a \pmod{n_t}} 1 \leq \frac{1}{8} \pi_2(n_{t+1}) \]

which is possible as explained at the beginning of this section.

To carry over the previous proof, it remains to show that

\[ \frac{1}{\pi_2(n_t)} \sum_{p_1 p_2 < n_t, (p_1 p_2, n_t) > 1} (-1)^{x(p_1 p_2)} = o(1). \]

To show this, for simplicity, we assume additionally that \(k_t\) are chosen to be perfect squares. If so then also all numbers \(n_t\) are perfect squares and

(18) \quad \text{if } p|n_t \text{ then } \frac{\varphi}{p} \geq \sqrt{n_t}.

Now, using that \(\sum_{p|k} \frac{1}{p} \ll \log \log \log k\), see e.g. [12] and (18), it follows that

\[
\sum_{p_1 p_2 < n_t, (p_1 p_2, n_t) > 1} 1 \ll \log^2 n_t + \sum_{p|n_t} \pi(n_t/p) = \\
\log^2 n_t + \sum_{p|n_t} O\left(\frac{n_t/p}{\log(n_t/p)}\right) = \log^2 n_t + n_t \sum_{p|n_t} \frac{1}{p} O\left(\frac{1}{\log(n_t/p)}\right) = \\
\log^2 n_t + \frac{2n_t}{\log n_t} O\left(\sum_{p|n_t} \frac{1}{p}\right) \ll \log^2 n_t + \frac{2n_t \log \log \log n_t}{\log n_t} = o(\pi_2(n_t))
\]

as needed.

4. Regular Toeplitz subshifts which satisfy a PNT (proof of Theorem B)

Let \(x \in \mathcal{A}^\mathbb{Z}\) be a regular Toeplitz sequence. Then, for every \(k \in \mathbb{N}\), there is an \(n_k\)-periodic sequence \(x_k \in (\mathcal{A} \cup \{?\})^\mathbb{Z}\) so that

\[ x_k(j) \neq ? \text{ implies } x(j) = x_k(j) = x_l(j) \text{ for all } l \geq k \]

and

\[ ?_k = ?_k(x) := \# \{0 \leq j < n_k : x_k(j) = ? \} = o(n_k). \]

For every Toeplitz sequence \(x \in \mathcal{A}^\mathbb{Z}\) and natural \(m\) let us consider a new Toeplitz sequence \(x^{(m)} \in (\mathcal{A}^{2m+1})^\mathbb{Z}\) given by

\[ x^{(m)}(j) = (x(j-m), \ldots, x(j+m)) \text{ for every } j \in \mathbb{Z}. \]

If \((n_t)_{t \geq 1}\) is a periodic structure of \(x\), then it is also a periodic structure of \(x^{(m)}\). Moreover,

(19) \quad ?_k(x^{(m)}) \leq (2m + 1)?_k(x) \text{ for every } k \geq 1.

Hence, the regularity of \(x\) implies the regularity of \(x^{(m)}\).

Theorem B follows directly from Lemma A.1 and the following result.
Theorem 4.1. Suppose that \((X_x, S)\) is a Toeplitz system such that
\[ ?_k = o(\varphi(n_k)). \]
Then \((X_x, S)\) satisfies a PNT.

Proof. To show a PNT for \((X_x, S)\), we need to show that for every continuous
\(F : X_x \rightarrow \mathbb{C}\) and every \(\varepsilon > 0\) there exists \(N_\varepsilon\) so that for every \(N, M \geq N_\varepsilon\) and every
\(r \in \mathbb{Z}\), we have
\[ \left| \frac{1}{\pi(N)} \sum_{p \leq N} F(S^{p+r}x) - \frac{1}{\pi(M)} \sum_{p \leq M} F(S^{p+r}x) \right| < \varepsilon. \]
We first assume that \(F : X_x \rightarrow \mathbb{R}\) depends only on the zero coordinate, i.e. \(F(y) = f(y(0))\) for some \(f : \mathcal{A} \rightarrow \mathbb{R}\).
Fix \(\varepsilon > 0\). Fix also \(k \geq 1\) so that
\[ ?_k < \frac{\varepsilon}{8}\varphi(n_k). \]
Next choose \(N_\varepsilon\) such that for every \(N \geq N_\varepsilon\), we have
\[ \left| \pi(N; n_k, a) - \frac{\pi(N)}{\varphi(n_k)} \right| < \varepsilon \frac{\pi(N)}{8\varphi(n_k)} \quad \text{for all } a \in \mathbb{Z} \text{ with } (a, n_k) = 1, \]
\[ \#\{p \leq N : p|n_k\} \leq \log n_k < \frac{\varepsilon}{8}\pi(N). \]
We will show that for all \(N \geq N_\varepsilon\) and \(r \in \mathbb{Z}\) we have
\[ \left| \frac{1}{\pi(N)} \sum_{p \leq N} F(S^{p+r}x) - \frac{1}{\varphi(n_k)} \sum_{0 \leq a < n_k} F(S^ax) \right| \leq \varepsilon \left\| F \right\|_{sup}, \]
which implies \((20)\).
Recall that \(x_k \in (\mathcal{A} \cup {?})^\mathbb{Z}\) is an \(n_k\)-periodic sequence (used to construct \(x\) at stage \(k\)). If for some \(a \in \mathbb{Z}\) we have
\[ x_k(a) \neq ?, \]
then
\[ x(a + j \cdot n_k) = x_k(a) \text{ for every } j \in \mathbb{Z}. \]
This implies that if \(p \leq N\) and \(x_k(p + r \mod n_k) \neq ?, \) then
\[ F(S^{p+r}x) = F(S^{p+r \mod n_k}x). \]
Note that
\[ \#\{p \leq N : x_k(p + r \mod n_k) = ?\} \]
\[ \leq \sum_{0 \leq a < n_k \atop (a-r,n_k) = 1} \#\{p \leq N : p = a - r \mod n_k\} \]
\[ + \sum_{0 \leq a < n_k \atop (a-r,n_k) > 1} \#\{p \leq N : p = a - r \mod n_k\}. \]
Assume that \(N \geq N_\varepsilon\). By \((22)\) and \((23)\), for every integer \(v\) with \((v, n_k) = 1\) we have
\[ \#\{p \leq N : p = v \mod n_k\} = \pi(N; n_k, v) \leq (1 + \varepsilon/8)\frac{\pi(N)}{\varphi(n_k)} \]
and
\[(26)\quad \sum_{0 \leq a < n_k \atop (a-r,n_k) > 1} \#\{p \leq N : p = a - r \mod n_k\} \leq \#\{p \leq N : p|n_k\} < \frac{\varepsilon}{8}\pi(N),\]
where left inequality follows from the fact that if \((a - r, n_k) > 1\) and \(p_a = a - r\) mod \(n_k\) for a prime \(p_a\), then \((a - r, n_k) = p_a\) and
\[\{p \leq N : p = a - r \mod n_k\} = \{p_a\}.
It follows that (use also \((21)\))
\[\#\{p \leq N : x_k(p + r \mod n_k) = ?\}
\leq \#\{0 \leq a < n_k : (a - r, n_k) = 1, x_k(a) = ?\}(1 + \varepsilon/8)\frac{\pi(N)}{\varphi(n_k)} + \varepsilon\frac{1}{8}\pi(N)
\leq?_k(1 + \varepsilon/8)\frac{\pi(N)}{\varphi(n_k)} + \varepsilon\frac{1}{8}\pi(N) \leq \varepsilon\frac{1}{2}\pi(N).

Let
\[P_N := \{p \leq N : x_k(p + r \mod n_k) \neq ?\}.
Then by the above, for every \(N \geq N_\varepsilon\),
\[(27)\quad \left| \frac{1}{\pi(N)} \sum_{p \leq N} F(S^{p+r}x) - \frac{1}{\pi(N)} \sum_{p \leq P_N} F(S^{p+r}x) \right| \leq \frac{\varepsilon}{2} \|F\|_{\text{sup}}.
But by \((25)\),
\[\sum_{p \leq P_N} F(S^{p+r}x) = \sum_{0 \leq a < n_k \atop x_k(a) \neq ?} \sum_{p \leq N \atop p \equiv a - r \mod n_k} F(S^a x)
= \sum_{0 \leq a < n_k \atop x_k(a) \neq ?} F(S^a x)\#\{p \leq N, p = a - r \mod n_k\}.
If \((a - r, n_k) = 1\), then again by \((22)\), we have
\[\left| \#\{p \leq N, p = a - r \mod n_k\} - \pi(N)\frac{\pi(N)}{\varphi(n_k)} \right| = \left| \pi(N; n_k, a - r) - \pi(N)\frac{\pi(N)}{\varphi(n_k)} \right| < \frac{\varepsilon}{8}\frac{\pi(N)}{\varphi(n_k)}.
In view of \((26)\), it follows that
\[\left| \frac{1}{\pi(N)} \sum_{p \leq P_N} F(S^{p+r}x) - \frac{1}{\pi(n_k)} \sum_{0 \leq a < n_k \atop (a-r,n_k)=1} F(S^a x) \right|
= \left| \sum_{0 \leq a < n_k \atop x_k(a) \neq ?} F(S^a x)\pi(N; n_k, a - r) - \frac{1}{\pi(n_k)} \sum_{0 \leq a < n_k \atop (a-r,n_k)=1} F(S^a x) \right|
\leq \frac{1}{\pi(N)} \sum_{0 \leq a < n_k \atop (a-r,n_k)=1} \left| F(S^a x) \pi(N; n_k, a - r) - \frac{\pi(N)}{\varphi(n_k)} \right| + \frac{\varepsilon}{8} \|F\|_{\text{sup}}
\leq \|F\|_{\text{sup}} \left( \frac{\varepsilon}{8} \#\{0 \leq a < n_k : x_k(a) \neq ?, (a - r, n_k) = 1\} \varphi(n_k) + \frac{\varepsilon}{8} \right) \leq \|F\|_{\text{sup}} \frac{\varepsilon}{2}.
Together with (27), this gives (24), which completes the proof in the case of $F$ depending only on the zero coordinate.

Now suppose that $F : X_x \to \mathbb{C}$ depends only on finitely many coordinates. Then there exists natural $m$ and $f : A^2^{m+1} \to \mathbb{C}$ such that $F(y) = f(y(-m), \ldots, y(m))$ for every $y = (y(k))_{k \in \mathbb{Z}} \in X_x$. Denote by $X_{x(m)} \subset (A^2^{m+1})^2$ the orbit closure of $x^{(m)} \in (A^2^{m+1})^2$. Then every $y^{(m)} \in X_{x(m)}$ is of the form $y^{(m)}(k) = (y(k-m), \ldots, y(k+m))$ for some $y = (y(k))_{k \in \mathbb{Z}} \in X_x$.

In view of (19), $(X_{x(m)}, S)$ is a regular Toeplitz shift with $\varphi(k(x^{(m)})) = o(\varphi(n_k))$. Let us consider $\tilde{F} : X_{x(m)} \to \mathbb{C}$ given by $\tilde{F}(y^{(m)}) = f(y^{(m)}(0)) = f(y(-m), \ldots, y(m))$ for $y^{(m)} \in X_{x(m)}$. Since $\tilde{F}$ depends only on the zero coordinate, by (20) applied to $x^{(m)}$ and the map $\tilde{F}$, for every $\varepsilon > 0$ there exists $N_\varepsilon$ such that for $N, M \geq N_\varepsilon$, we have

$$\left| \frac{1}{\pi(N)} \sum_{p \leq N} F(S^{p+r}x) - \frac{1}{\pi(M)} \sum_{p \leq M} F(S^{p+r}x) \right| < \varepsilon.$$ 

Thus (20) holds for every $F : X_x \to \mathbb{C}$ depending only on finitely many coordinates. As the set of such functions is dense in $C(X_x)$, (20) also holds for every $F \in C(X_x)$, which completes the proof.

As $\varphi(n) \to \infty$ when $n \to \infty$, we obtain the following result.

**Corollary 4.2.** If $x$ is Toeplitz for which the sequence $(\varphi(k))$ is bounded then $(X_x, S)$ satisfies a PNT.

5. **TOEPLITZ SUBSHIFTS FOR WHICH AN SPNT HOLDS**

**Theorem 5.1.** Suppose that $(X_x, S)$ is a Toeplitz system such that $\varphi(k) = o(\varphi(n_k))$.

Then, for every $F \in C(X_x)$ and $y \in X_x$, the limit

$$\lim_{N \to \infty} \frac{1}{\pi_2(N)} \sum_{p_1p_2 < N} F(S^{p_1p_2}y)$$

exists.

**Proof.** The proof proceeds along the same lines as the proof of Theorem 4.1. It relies on the following counterpart of (24): for every $\varepsilon > 0$ there exists a natural $N_\varepsilon$ such that for all $N \geq N_\varepsilon$ and $r \in \mathbb{Z}$, we have

$$\left| \frac{1}{\pi_2(N)} \sum_{p_1p_2 < N} F(S^{p_1p_2+r}x) - \frac{1}{\varphi(n_k)} \sum_{0 \leq a < n_k \atop (a-r,n_k) = 1 \atop x_k(a) \neq ?} F(S^ax) \right| \leq \varepsilon \|F\|_{sup}.$$ 

In turn, the proof of (24) is based on only two elements: (22) and (26). Their semi-prime counterparts follows directly from (17) and (16), respectively. Now, we repeat the arguments of the proof of (24) almost word for word, replacing (22) and (26) by their semi-prime counterparts.

**Remark 5.2.** In view of (24) and (28), under the assumption $\varphi(k) = o(\varphi(n_k))$, we have

$$\lim_{N \to \infty} \frac{1}{\pi_2(N)} \sum_{p_1p_2 < N} F(S^{p_1p_2}y) = \lim_{N \to \infty} \frac{1}{\pi(N)} \sum_{p < N} F(S^p y).$$
for every $F \in C(X_x)$ and $y \in X_x$, so a PNT and an SPNT fully coincide for this class of regular Toeplitz systems.

6. Ergodic averages along polynomial times

Let $P$ be a monic polynomial\footnote{The leading coefficient of $P$ equals 1. This assumption is only for simplicity. In fact, Theorem \ref{thm:main} below is true whenever the set of (non-zero) coefficients of $P - P(0)$ is coprime, see the proof of Corollary \ref{cor:main} and the assumptions of Albid theorem in \cite{Albi}.} of degree $d > 1$ with non-negative integer coefficients. Note that, under these assumptions, $P(\cdot)$ is a strictly increasing function on $\mathbb{N}$. For every $n \in \mathbb{N}$, let

$$R_n^P := \{0 \leq a < n : a = P(m) \mod n \text{ for some } m \in \mathbb{N}\}$$

and \psi(n) := \#R_n^P.

For all $N, n \in \mathbb{N}$ and $a \in R_n^P$, let

$$\rho^P(N; n, a) = \#\{1 \leq m \leq N : P(m) = a \mod n\}.$$ 

and

$$\rho^P(n, a) := \rho^P(n; n, a), \quad \rho^P(n) := \max_{a \in R_n^P} \rho^P(n; n, a).$$

**Lemma 6.1.** The function $\psi^P$ is multiplicative, i.e. $\psi^P(n_1 n_2) = \psi^P(n_1) \psi^P(n_2)$ if $(n_1, n_2) = 1$. If $a \in \mathbb{Z}/n\mathbb{Z}$, $n_1, \ldots, n_k$ are pairwise coprime and $n = n_1 \cdots n_k$ then $a \in R_n^P$ iff $a_i \in R_{n_i}^P$ for $i = 1, \ldots, k$, where $0 \leq a_i < n_i$ is the remainder of $a$ when divided by $n_i$ (that is, $0 \leq a_i < n_i$ and $a_i = a \mod n_i$). Moreover,

$$\rho^P(n, a) = \prod_{i=1}^k \rho^P(n_i, a_i). \tag{29}$$

**Proof.** Note that the multiplicativity of $\psi^P$ follows from the second part of the lemma.

Moreover, note that $a \in R_n^P$ iff $a = P(m) \mod n$ for some $0 \leq m < n$. Indeed, if $a = P(m) \mod n$ for some $m \in \mathbb{N}$, then $a = P(m') \mod n$, where $0 \leq m' < n$ is the remainder of $m$ when divided by $n$.

If $a \in R_n^P$, i.e. $a = P(m) \mod n_1 \cdots n_k$ for some $0 \leq m < n$, then $a_i = a = P(m) = P(m_i) \mod n_i$ for every $i = 1, \ldots, k$, where $0 \leq m_i < n_i$ is the remainder of $m$ when divided by $n_i$.

Now, suppose $a \in \mathbb{Z}/n\mathbb{Z}$, $a_i = a \mod n_i$ and $a_i \in R_{n_i}^P$ for $i = 1, \ldots, k$. Then, for every $i = 1, \ldots, k$, there exists $0 \leq m_i < n_i$ such that $a_i = P(m_i) \mod n_i$. By the Chinese Remainder Theorem, there exists a unique $0 \leq m < n$ such that $m = m_i \mod n_i$ for $i = 1, \ldots, k$. It follows that

$$P(m) = P(m_i) = a_i = a \mod n_i \text{ for all } i = 1, \ldots, k.$$ 

This yields $a = P(m) \mod n_1 \cdots n_k$ and $a \in R_n^P$.

The argument above also shows \cite{209}. \hfill \square

**Remark 6.2.** Note that in the argument above we used the fact that the $a_i$’s determine $a$ as by the ChRT there exists only one $0 \leq a < n$ such that $a = a_i \mod n_i$ for each $i = 1, \ldots, k$.

For any natural $n$ denote by $\omega(n)$ the number of its prime divisors (counted without multiplicities) and by $p(n)$ the product of its prime divisors.

**Corollary 6.3.** The arithmetic function $\rho_P$ is multiplicative and $\rho^P(n) \leq \frac{d^c(n)}{p(n)} n$. 
Proof. The multiplicativity of $\rho^P$ follows directly from \[29\]. By Albis theorem (see Corollary 3 of Theorem 1.23 in [17]), for any prime number we have $\rho^P(p^n) \leq d p^{n-1}$. This result combined with the multiplicativity of $\rho^P$ gives the required bound of $\rho^P(n)$. \qed

Lemma 6.4. For all $n \in \mathbb{N}$, $a \in R^P_n$ and $N \geq P(n)$, we have
\[
\rho^P(n, a) \left( \frac{P^{-1}(N)}{n} - 1 \right) \leq \# \{ m \in \mathbb{N} : 1 \leq P(m) \leq N, P(m) = a \mod n \}
\]
\[
\leq \rho^P(n, a) \left( \frac{P^{-1}(N)}{n} + 1 \right).
\]

Proof. Let $s := \rho^P(n, a)$ and let $1 \leq m_1 < \ldots < m_s \leq n$ be all numbers such that $P(m_i) = a \mod n$. Note that a natural number $m$ satisfies $P(m) \leq N$ and $P(m) = a \mod n$ iff $m = jm + r$ with $0 \leq j \leq (P^{-1}(N) - r)/n$ and $0 < r \leq n$ satisfies $P(r) = a \mod n$. Thus, $r = m_i$ for some $i = 1, \ldots, s$. It follows that
\[
\rho := \# \{ m \in \mathbb{N} : 1 \leq P(m) \leq N, P(m) = a \mod n \}
\]
\[
= \sum_{i=1}^{s} \left( \left\lfloor \frac{P^{-1}(N) - m_i}{n} \right\rfloor + 1 \right).
\]

Since
\[
\frac{P^{-1}(N)}{n} - 1 \leq \frac{P^{-1}(N) - m_i}{n} < \left\lfloor \frac{P^{-1}(N) - m_i}{n} \right\rfloor + 1
\]
\[
\leq \frac{P^{-1}(N) - m_i}{n} + 1 < \frac{P^{-1}(N)}{n} + 1,
\]
by summing up, this gives
\[
s \left( \frac{P^{-1}(N)}{n} - 1 \right) \leq \rho \leq s \left( \frac{P^{-1}(N)}{n} + 1 \right).
\]
\[\Box\]

Remark 6.5. As $P$ is an increasing function, we can apply the above inequalities to $P(N)$ instead of $N$ (as $P(N) \geq N$). Then $P(m) \leq P(N)$ iff $m \leq N$, and the result of the lemma implies
\[
\rho^P(n, a) \left( \frac{N}{n} - 1 \right) \leq \rho^P(N; n, a) \leq \rho^P(n, a) \left( \frac{N}{n} + 1 \right).
\]

We now focus on the simplest case when $P(n) = n^2$. We continue to write $R$ for $R^P$, $\psi$ for $\psi^P$ and $\rho$ for $\rho^P$. In view of Theorems 1.27 and 1.30 in [17], we have the following result.

Proposition 6.6. For every prime number $p > 2$, for every $a \in R_p \setminus \{0\}$, where $N = 2n$ or $2n + 1$, we have
\[
\rho(p^n, a) = \begin{cases} 
2 & \text{if } a = a' \mod p \text{ for } a' \in R_p \setminus \{0\} \\
2p^r & \text{if } a = p^r a' \text{ and } a' = a'' \mod p \text{ for } a'' \in R_p \setminus \{0\} \\
p^n & \text{if } a = 0.
\end{cases}
\]

\footnote{Note that compared to notation from [17], we have: $\rho^P(n; n, a) = \lambda_{p^n}(n)$, $\rho^P(n) = \max_{a \in R_p^n} \lambda_{p^n}(n)$; the estimate on $\lambda_P$ in [17] depends only on the degree of the polynomial.}
Moreover, we have
\[ \psi(p^{2n+1}) = \frac{p^{2n+2} + 2p + 1}{2(p+1)} \quad \text{and} \quad \psi(p^{2n}) = \frac{p^{2n+1} + p + 2}{2(p+1)}. \]

Furthermore, If \( p = 2 \) then
\[ \rho(2, a) = 1 \text{ for all } a \in R_2, \quad \rho(4, a) = 2 \text{ for all } a \in R_4 \]
and for any \( N \geq 3 \), where \( N = 2n \) or \( 2n + 1 \), for every \( a \in R_{2N} \), we have
\[ \rho(2N, a) = \begin{cases} 
4 & \text{if } a = 1 \mod 8 \\
4 \cdot 2^r & \text{if } a = 2^r a', 2r \leq N - 3, a' = 1 \mod 8 \\
2 \cdot 2^r & \text{if } a = 2^r a', 2r = N - 2, a' = 1 \mod 4 \\
2^r & \text{if } a = 2^r a', 2r = N - 1, a' = 1 \mod 2 \\
2^n & \text{if } a = 0.
\end{cases} \]

Moreover,
\[ \psi(2^{2n}) = \frac{2^{2n-1} + 4}{3} \quad \text{and} \quad \psi(2^{2n+1}) = \frac{2^{2n} + 5}{3}. \]

**Corollary 6.7.** For every natural \( n \geq 2 \), we have \( \rho(n) \leq 4\sqrt{n} \). Moreover, if \( n \) is square-free, then \( \rho(n) \leq 2^{\omega(n)} \).

**Proof.** By a direct inspection of the formulas in Proposition 6.6, we obtain:
\[ \rho(2N) \leq 2\sqrt{2N}, \quad \rho(3N) \leq 2\sqrt{3N} \]
but for all \( p \geq 5 \), we have
\[ \rho(p^n) \leq \sqrt{pN}. \]

Indeed, for the cases \( a = a' \mod p \) (for \( a' \in R_p \{0\} \)) and \( a = 0 \), it is direct. For the case \( \rho(p^n, a) = 2p^r \), we have \( a = p^r a' < p^n \), so \( 2r \leq N - 1 \) and then indeed \( 2p^r \leq p^{N/2} \).

The second inequality follows directly from \( \rho(p) \leq 2 \). \( \square \)

For some future purposes, we are interested in cases (in Proposition 6.6) which gives possibly smallest values for the function \( \rho \), hence, for every prime number \( p \) and any natural \( n \), let
\[ \tilde{R}_{p^n} := \begin{cases} 
\{ 0 \leq a < p^n : a = a' \mod p \text{ for } a' \in R_p \{0\} \} & \text{if } p > 2 \\
R_2 & \text{if } p^n = 2 \\
R_4 & \text{if } p^n = 4 \\
\{ 0 \leq a < 2^n : a = 1 \mod 8 \} & \text{if } n \geq 3.
\end{cases} \]

By Proposition 6.6 \( \tilde{R}_{p^n} \subset R_{p^n} \).

Let \( n = p_1^{m_1} p_2^{m_2} \ldots p_k^{m_k} \) be the canonical representation of \( n \). Let
\[ \Phi : \mathbb{Z}/n\mathbb{Z} \to \mathbb{Z}/p_1^{m_1} \mathbb{Z} \times \ldots \times \mathbb{Z}/p_k^{m_k} \mathbb{Z} \]

\[ \psi(p^{2n}) = \frac{p - 1}{2} p^{2n-1} + \sum_{r=1}^{n-1} \frac{p - 1}{2} p^{2n-2r-1} + 1 = \]

\[ 1 + \frac{p - 1}{2} \sum_{r=0}^{n-1} p^{2(n-r-1)} = 1 + \frac{p(p-1)}{2} \frac{(p^2)^{n-1} - 1}{p^2 - 1} = \frac{p^{2n+1} + p + 2}{2(p+1)}. \]
be the canonical ring isomorphism. Recall (cf. Lemma 6.1 and Remark 6.2) that \( \Phi \) establishes a one-to-one correspondence between \( \tilde{R}_n \) and \( R_{p_1^{m_1}} \times \ldots \times R_{p_k^{m_k}} \). Set

\[ \tilde{R}_n := \Phi^{-1}(\tilde{R}_{p_1^{m_1}} \times \ldots \times \tilde{R}_{p_k^{m_k}}) \]

and

\[ \tilde{\psi}(n) := \# \tilde{R}_n. \]

Then, clearly, \( \tilde{\psi} \) is a multiplicative function. Moreover, by Proposition 6.6 for each \( a \in \tilde{R}_n \), we have

\[ \rho(p^N, a) = \begin{cases} 1 & \text{if } p^N = 2 \\ 2 & \text{if } p^N = 2 \text{ or } p > 2 \\ 4 & \text{if } p = 2 \text{ and } N \geq 3. \end{cases} \]

Hence, in view of (29), for every \( a \in \tilde{R}_n \), we have

\[ \frac{1}{2} \cdot 2^{\omega(n)} \leq \rho(n, a) \leq 2 \cdot 2^{\omega(n)}. \]

Moreover, by definition,

\[ \tilde{\psi}(p^n) := \begin{cases} \frac{p^n - 1}{2} & \text{if } p > 2 \\ 2 & \text{if } p^n = 2 \\ 2^{n-3} & \text{if } p = 2 \text{ and } n \geq 3. \end{cases} \]

It follows that

\[ \frac{1}{2} \prod_{p^n} \left( 1 - \frac{1}{p} \right) \leq \frac{2^{\omega(n)} \tilde{\psi}(n)}{n} \leq 4 \prod_{p^n} \left( 1 - \frac{1}{p} \right). \]

(To obtain these inequalities, for \( n = p_1^{m_1} p_2^{m_2} \ldots p_k^{m_k} \), write \( \frac{2^{\omega(n)} \tilde{\psi}(n)}{n} = \prod_{i=1}^{k} \frac{2 \psi(p_i^{m_i})}{p_i^{m_i}} \) and apply the formula above.)

6.1. Non-conventional ergodic theorem.

**Theorem 6.8.** Suppose that \((X_x, S)\) is a Toeplitz system such that

\[ \omega_k = o(n_k/\rho^P(n_k)). \]

Then, for every continuous map \( F : X_x \to \mathbb{C} \) and \( y \in X_x \), the limit

\[ \lim_{N \to \infty} \frac{1}{N} \sum_{m \leq N} F(S^{P(m)}y) \]

exists.

**Proof.** To show (33), we need to prove that for every \( \varepsilon > 0 \) there exists \( N_\varepsilon \) so that for every \( N, M \geq N_\varepsilon \) and every \( r \in \mathbb{Z} \), we have

\[ \left| \frac{1}{N} \sum_{m \leq N} F(S^{P(m)+r}x) \right| < \varepsilon. \]

We first assume that \( F : X_x \to \mathbb{R} \) depends only on the zero coordinate, i.e. \( F(y) = f(y(0)) \) for some \( f : A \to \mathbb{R} \).

Fix \( \varepsilon > 0 \). Choose \( k \geq 1 \) so that

\[ \omega_k < \frac{\varepsilon}{8 \rho^P(n_k)}. \]
Next, choose $N_\varepsilon \geq 8n_k^2/\varepsilon$. Then, in view of Remark 6.5 (and the choice of $N_\varepsilon$), for every $N \geq N_\varepsilon$ and $a \in R^P_{n_k}$, we have

$$\left| \rho^P(N; n_k, a) - \rho^P(n_k, a) \right| N n_k \leq \rho^P(n_k) \leq \varepsilon N \frac{N}{n_k}. \tag{36}$$

From now on, we write that an integer number $v$ belongs to $R^P_{n_k}$ if there exists $0 \leq v' < n_k$ such that $v' = v \mod n_k$ and $v' \in R^P_{n_k}$. We will show that for all $N \geq N_\varepsilon$ and $r \in \mathbb{Z}$, we have

$$\left| \frac{1}{N} \sum_{m \leq N} F(S^{P(m)+r}x) - \frac{1}{n_k} \sum_{0 \leq a < n_k, \ a-r \in R^P_{n_k}, x_k(a) \neq ?} \rho^P(n_k, a-r) F(S^a x) \right| \leq \varepsilon \|F\|_{\sup}, \tag{37}$$

and this implies (34).

Recall that $x_k \in (A \cup \{?\})^\mathbb{Z}$ is an $n_k$-periodic sequence (used to construct $x$ at stage $k$). Note that for every $a \in \mathbb{Z}$, we have

$$x_k(a) \neq ? \Rightarrow x(a + j \cdot n_k) = x_k(a) \text{ for every } j \in \mathbb{Z}.$$  

This implies that if $m \leq N$ and $x_k(P(m) + r \mod n_k) \neq ?$, then

$$F(S^{P(m)+r}x) = F(S^{P(m)+r \mod n_k}x). \tag{38}$$

Therefore,

$$\#\{m \leq N : x_k(P(m) + r \mod n_k) = ?\} = \sum_{0 \leq a < n_k, \ a-r \in R^P_{n_k}, x_k(a) = ?} \#\{m \leq N : P(m) = a-r \mod n_k\} = \sum_{0 \leq a < n_k, \ a-r \in R^P_{n_k}, x_k(a) = ?} \rho^P(N; n_k, a-r). \tag{39}$$

Assume that $N \geq N_\varepsilon$. By (36), for every integer $v \in R^P_{n_k}$, we have

$$\rho^P(N; n_k, v) \leq 2\rho^P(n_k) \frac{N}{n_k}. \tag{35}$$

In view of (35), it follows that

$$\#\{m \leq N : x_k(P(m) + r \mod n_k) = ?\} \leq \#\{0 \leq a < n_k : a-r \in R^P_{n_k}, x_k(a) = ?\} 2\rho^P(n_k) \frac{N}{n_k} \leq 2^2 k \rho^P(n_k) \frac{N}{n_k} \leq \varepsilon \frac{N}{4}.$$  

Let

$$U_N := \{m \leq N : x_k(P(m) + r \mod n_k) \neq ?\}.$$  

Then by the above, for every $N \geq N_\varepsilon$,

$$\left| \frac{1}{N} \sum_{m \leq N} F(S^{P(m)+r}x) - \frac{1}{N} \sum_{m \in U_N} F(S^{P(m)+r}x) \right| \leq \frac{\varepsilon}{4} \|F\|_{\sup}. \tag{39}$$


But by (38),
\[ \sum_{m \in U_N} F(S^{P(m)+r} x) = \sum_{0 \leq a < n_k \atop a-r \in R^P_{nk}} \sum_{m \in N} F(S^a x) \]
\[ = \sum_{0 \leq a < n_k \atop a-r \in R^P_{nk}} F(S^a x) \# \{ m \leq N : P(m) = a-r \mod n_k \} \]
\[ = \sum_{0 \leq a < n_k \atop a-r \in R^P_{nk}} F(S^a x) \rho^P(N; a-r). \]

By (39), we have
\[ | \rho^P(N; n_k, a-r) - \rho^P(n_k, a-r) | \leq \frac{\varepsilon N}{8 n_k}. \]

It follows that
\[ \left| \frac{1}{N} \sum_{m \in U_N} F(S^{P(m)+r} x) - \frac{1}{n_k} \sum_{0 \leq a < n_k \atop a-r \in R^P_{nk}} \rho^P(n_k, a-r) F(S^a x) \right| \]
\[ \leq \frac{1}{N} \sum_{0 \leq a < n_k \atop a-r \in R^P_{nk}} |F(S^a x)| | \rho^P(N; a-r) - \rho^P(n_k, a-r) | \frac{N}{n_k} \]
\[ \leq \frac{\varepsilon}{8} \sup_{n_k} \frac{\# \{ 0 \leq a < n_k : x_k(a) \neq ?, a-r \in R^P_{nk} \} }{n_k} \leq \frac{\varepsilon}{8}. \]

Together with (39), this gives (37), which completes the proof in the case of $F$ depending only on the zero coordinate. The rest of the proof runs as in the proof of Theorem 4.1. This is by passing to the Toeplitz sequences $x^{(m)} \in (A^{2m+1})^\mathbb{Z}$ for $m \geq 1$. \[ \square \]

**Remark 6.9.** Denote by $\mathbb{P}(n_k)$ the set of all prime divisors of elements of the sequence $(n_k)_{n \geq 1}$. In view of Corollary 6.3, $\rho_t = o(p(n_k)/d^{(n_k)})$ implies (32). Unfortunately, if $\mathbb{P}(n_k)$ is finite then the sequence $(p(n_k)/d^{(n_k)})_{n \geq 1}$ is bounded, so Theorem 6.8 in the way, is not applicable. Fortunately, if $\mathbb{P}(n_k)$ is infinite then $p(n_k)/d^{(n_k)} \to +\infty$ as $t \to +\infty$, so Theorem 6.8 applies to a non-trivial class of regular Toeplitz shifts, in particular, it applies when the periodic sequences $x_t$ defining $x$ have a bounded number of “?”. However, Theorem 6.8 applies to a much wider class of regular Toeplitz shifts when $P(n) = n^2$. Then, by Corollary 6.4, $\rho_t = o(\sqrt{n_k})$ implies (32). Here the finiteness or the infinity of the set $\mathbb{P}(n_k)$ does not matter.
The assumption (32) about the growth of the sequence $(?_t)_{t \geq 1}$ is the least restrictive when all $n_t$ are square-free. Then, by the second part of Corollary 6.1, $?_t = o(n_t/2^\omega(n_t))$ implies (32). Therefore, $?_t = O(n_t^{1 - \frac{1}{\log_2 \log_2 n_t}})$ also implies (32). Indeed, it suffices to show that $2^\omega(n) = o(n^{1 - \frac{1}{\log_2 \log_2 n}})$ for square-free numbers $n \to +\infty$.

Suppose that $\omega(n) = k$ and denote by $(p_t)_{t \geq 1}$ the increasing sequence of all prime numbers. Since

$$\ln n \geq \sum_{l=1}^k \ln p_l \geq k \ln k,$$

we have

$$\frac{2^\omega(n)}{n^{\frac{1}{\log_2 \log_2 n}}} \leq \frac{2^k}{2^k \log_2 k} = \frac{1}{2^k \log_2 \log_2 k}.$$ 

As $\frac{k \log_2 \log_2 k}{\log_2 k + \log_2 \log_2 k} \to +\infty$ when $k \to +\infty$, this gives $2^\omega(n) = o(n^{1 - \frac{1}{\log_2 \log_2 n}})$.

6.2. Counter-examples. We will show that there exists a regular Toeplitz sequence $x \in \{0, 1\}^{\mathbb{Z}}$ with the period structure $(n_t)_{t \geq 1}$ satisfying

$$(40) \quad n_{t+1} = k_{t+1} n_k \text{ with } (k_{t+1}, n_t) = 1, \ n_{t+1} \geq 2^4 n_t^2 \text{ and } \sum_{p \in \mathbb{P}(n_t)} \frac{1}{p} < +\infty$$

and such that

$$\lim_{t \to \infty} \frac{1}{\sqrt{n_t}} \sum_{0 \leq m < \sqrt{n_t}} F(S^{m^2} x) \text{ does not exist,}$$

where $F(y) = (-1)^{\nu(0)}$. Let

$$0 < \beta := \frac{1}{16} \prod_{p \in \mathbb{P}(n_t)} \frac{p - 1}{p}.$$

By (31), for every $t \geq 1$, we have

$$(41) \quad \frac{2^\omega(n_t) \widetilde{\psi}(n_t)}{n_t} \geq 8 \beta.$$

Passing to a subsequence of $(n_t)_{t \geq 1}$ (and remembering that $\widetilde{\psi}(m) \to \infty$ when $m \to \infty$), we can assume that

$$\sum_{t \geq 1} \frac{1}{\psi(k_t)} \leq \frac{1}{2}.$$

Set

$$\gamma_t := \sum_{t=1}^t \frac{1}{\psi(k_t)} \leq \frac{1}{2}.$$

At stage $t$, $x$ is approximated by the infinite concatenation of $x_t[0, n_t - 1] \in \{0, 1, ?\}^{n_t}$ (that is, we see a periodic sequence of 0, 1, ? with period $n_t$). Successive "?" will be filled in in the next steps of construction of $x$. We require that:

$$(42) \quad \{0 \leq i < n_t : x_t(i) = ?\} \subset R_{n_t};$$

$$(43) \quad \# \{a \in \tilde{R}_{n_t} : x_t(a) = ?\} \geq (1 - \gamma_t) \widetilde{\psi}(n_t);$$

$$(44) \quad \# \{0 \leq m < \sqrt{n_t} : x_t(m^2) = ?\} \geq \beta \sqrt{n_t}.$$
Recall that, in view of Lemma 6.4 (remembering that $P^{-1}(n_{t+1}) = \sqrt{n_{t+1}}$, (30) and (40), for each $a \in \tilde{R}_{nt}$, we have
\[
\#(\{a + jm_t : 0 < j < k_{t+1}\} \cap \mathbb{N}^2) \geq \#(\{m^2 = a \mod n_t : m^2 < n_{t+1}\}) - 1
\]
\[
\geq \left(\frac{\sqrt{n_{t+1}}}{n_t} - 1\right)\rho(n_t, a) - 1 \geq \left(\frac{\sqrt{n_{t+1}}}{n_t} - 1\right)\frac{1}{2}2^{\omega(n_t)} - 1
\]
\[
\geq \frac{1}{2}2^{\omega(n_t)}\left(\frac{\sqrt{n_{t+1}}}{n_t} - 2\right) \geq \frac{1}{4}2^{\omega(n_t)}\frac{\sqrt{n_{t+1}}}{n_t},
\]
so
\[
\#(\{a + jm_t : 0 < j < k_{t+1}\} \cap \mathbb{N}^2) \geq \frac{2^{\omega(n_t)}\sqrt{n_{t+1}}}{4n_t}.
\]
By the definition of the sets $R_n$ and $\tilde{R}_n$, we have
\[
R_{nt+1} \subset \bigcup_{a \in R_{nt}} \{a + jm_t : 0 \leq j < k_{t+1}\},
\]
(46)
\[
\tilde{R}_{nt+1} \subset \bigcup_{a \in \tilde{R}_{nt}} \{a + jm_t : 0 \leq j < k_{t+1}\}.
\]
(47)
Moreover, by Lemma 6.1 for every $a \in \tilde{R}_{nt}$, we have
\[
\#\{i \in \tilde{R}_{nt+1} : i = a \mod n_t\} = \#\tilde{R}_{kt+1} = \tilde{\psi}(k_{t+1}).
\]
We need to describe now which and how we fill "?" in $x_{t+1}[0, n_{t+1} - 1]$. This block is divided into $k_{t+1}$ subblocks
\[
x_t[0, n_t - 1]x_t[0, n_t - 1] \ldots x_t[0, n_t - 1].
\]
We fill in all "?" in the first block $x_t[0, n_t - 1]$ in such a way to “destroy” the convergence of averages in (40) for the time $n_t$, namely
\[
\frac{1}{\sqrt{n_t}} \sum_{0 \leq m < \sqrt{n_t}} F(S^{m^2} x) = \frac{1}{\sqrt{n_t}} \left( \sum_{m < \sqrt{n_t} \atop x_t(m^2) = 0} 1 - \sum_{m < \sqrt{n_t} \atop x_t(m^2) = 1} 1 + \sum_{m < \sqrt{n_t} \atop x_t(m^2) = ?} (-1)^{x(m^2)} \right).
\]
And, since the number of $m$ in the last summand is at least $\beta \sqrt{n_t}$ in view of (44), we can fill in these places at stage $t + 1$ to obtain the sum completely different that the known number which we had from stage $t - 1$. We also fill in (in an arbitrary way) the remaining places in $\{0, \ldots, n_t - 1\}$.
We fill in (in an arbitrary way) all places in $\{n_t, \ldots, n_{t+1} - 1\} \setminus R_{nt+1}$ and only these places, so that (42) will be satisfied at stage $t + 1$.
We must remember that for any $a \in R_{nt}$ if $x_t(a) \neq ?$ then for every $0 \leq j < k_{t+1}$, we have $x_{t+1}(a + jm_t) = x_t(a + jm_t) = x_t(a) \neq ?$. Moreover, for any $a \in \tilde{R}_{nt}$ if $x_t(a) = ?$ then for every $0 < j < k_{t+1}$ with $a + jm_t \in \tilde{R}_{nt+1}$ we have $x_{t+1}(a + jm_t) = ?$. In view of (47), this gives
\[
\#\{i \in \tilde{R}_{nt+1} : x_{t+1}(i) \neq ?\} \leq \tilde{\psi}(n_t) + \sum_{a \in \tilde{R}_{nt} : x_t(a) \neq ?} \#\{a + jm_t \in \tilde{R}_{nt+1} : 0 < j < k_{t+1}\}.
\]
\[\text{#} \mathbb{N}^2 \text{ stands for } \{m^2 : m \geq 0\}.\]
In view of (48) and (43), it follows that
\[ \# \{ i \in \tilde{R}_{n_{t+1}} : x_{t+1}(i) \neq ? \} \leq \tilde{\psi}(n_t) + (\tilde{\psi}(k_{t+1}) - 1) \# \{ a \in \tilde{R}_{n_t} : x_t(a) \neq ? \} \leq \tilde{\psi}(n_t) + (\tilde{\psi}(k_{t+1}) - 1) \gamma_t \tilde{\psi}(n_{t+1}) = \left( \gamma_t + \frac{1 - \gamma_t}{\tilde{\psi}(k_{t+1})} \right) \tilde{\psi}(n_{t+1}) \leq \gamma_{t+1} \tilde{\psi}(n_{t+1}). \]

Therefore, at stage \( t + 1 \), also (43) is satisfied.

A similar argument combined with (45), (43) and (41) shows that
\[ \# \{ 0 \leq m^2 < n_{t+1} : x_{t+1}(m^2) = ? \} = \# \{ i \in R_{n_{t+1}} \cap \mathbb{N}^2 : x_{t+1}(i) = ? \} \geq \sum_{a \in R_{n_t} : x_t(a) = ?} 2^{\omega(n_t)} \sqrt{n_{t+1}} = \frac{2^{\omega(n_t)} \sqrt{n_{t+1}}}{4n_t} \# \{ a \in \tilde{R}_{n_t} : x_t(a) = ? \} \geq (1 - \gamma_t) \frac{2^{\omega(n_t)} \sqrt{n_{t+1}}}{4n_t} \tilde{\psi}(n_t) \geq \beta \sqrt{n_{t+1}}. \]

Therefore, at stage \( t + 1 \), also (44) is satisfied. This completes the construction.

**Remark 6.10.** In view of (42), in the constructed example of Toeplitz system \( (X, S) \) we have \( \gamma_t \leq \psi(n_t) \). Moreover, \( \psi(n_t) = o(\varphi(n_t)) \). Indeed, by Proposition 6.6 for every prime number \( p \) we have \( \psi(p^n) \leq p^{n - \frac{p + 2}{2}} \). It follows that
\[ \frac{\psi(p^n)}{\varphi(p^n)} \leq \frac{1}{2} \cdot \frac{p + 2}{p - 1} \leq \frac{3}{4} \]
for all prime \( p \geq 7 \). It follows that
\[ \frac{\psi(n_t)}{\varphi(n_t)} = O \left( \left( \frac{3}{4} \right)^{\omega(n_t)} \right) = o(1). \]

Consequently, we have \( \gamma_t = o(\varphi(n_t)) \). Therefore, in view of Theorem 4.1 \( (X, S) \) satisfies a PNT.

**Appendix A. The Diameter of a Tower**

Let \( x \in \mathcal{A}^\mathbb{Z} \) be a Toeplitz sequence with the periodic structure given by \( (n_t)_{t \geq 1} \). Recall that
\[ \text{Per}_{n_t}(x) = \{ a \in \mathbb{Z} : x(a + jn_t) = x(a) \text{ for all } j \in \mathbb{Z} \}. \]
Let \( \text{Aper}_{n_t}(x) := \mathbb{Z} \setminus \text{Per}_{n_t}(x) \). Then, we define the periodic sequence \( x_t \in (\mathcal{A} \cup \{ ? \})^\mathbb{Z} \) by: \( x_t(k) = x(k) \) if \( k \in \text{Per}_{n_t}(x) \) and \( x_t(k) = ? \) if \( k \in \text{Aper}_{n_t}(x) \). Note that the density of the set \( \text{Aper}_{n_t}(x) \) is equal to \( \frac{n_t}{n} \), where
\[ ?_t = \# \{ 0 \leq k < n_t : x_t(k) = ? \} = \#(\text{Aper}_{n_t}(x) \cap \{ 0, 1, \ldots, n_t - 1 \}). \]

It follows that the regularity of \( (X, S) \) is equivalent to \( ?_t = o(n_t) \).

**Lemma A.1.** For any Toeplitz sequence \( x \in \mathcal{A}^\mathbb{Z} \) we have
\[ ?_t \leq \delta(E_t) \leq 3 ?_t, \text{ for every } t \geq 1. \]

**Proof.** First note that for every \( 0 \leq j < n_t \) we have
\[ E_j^t = \{ y \in X : y(k - j) = x(k) = x_t(k) \text{ for all } k \in \text{Per}_{n_t} \}. \]
Moreover, if \( k \in \text{Aper}_{n_i}(x) \) then we can find \( y, z \in E'_j \), so that \( y(k - j) \neq z(k - j) \). It follows that

\[
\text{diam}(E'_j) = 2^{-\inf\{|n|: n \in \text{Aper}_{n_i}(x) - \{j\}}\).
\]

Suppose that

\[
\text{Aper}_{n_i}(x) \cap \{0, 1, \ldots, n_t - 1\} = \{l_1, l_2, \ldots, l_s\}
\]

with \( 1 \leq l_1 \leq \ldots \leq l_s \leq n_t \) and \( s = \tau_t \). Thus, \( \text{diam}(E'_i) = 1 \) if and only if \( l_{i-1} < j < l_i \) (\( l_0 = l_s - n_t \) and \( l_{s+1} = l_1 + n_t \)) then \( \text{diam}(E'_i) = 2^{-\min(j - l_{i-1}, l_i - j)} \). Therefore,

\[
\delta(E') = \sum_{0 \leq j < n_t} \text{diam}(E'_j) \geq \sum_{i=1}^{s} \text{diam}(E'_i) = s
\]

and

\[
\delta(E') = \sum_{0 \leq j < n_t} \text{diam}(E'_j) = \sum_{i=1}^{s} \sum_{\frac{l_{i-1} + l_i}{2} \leq j < \frac{l_{i} + l_{i+1}}{2}} \text{diam}(E'_j)
\]

\[
= \sum_{i=1}^{s} \left( 1 + \sum_{1 \leq j < \frac{l_{i+1} - l_{i}}{2}} 2^{-j} + \sum_{1 \leq j < \frac{l_{i} - l_{i-1}}{2}} 2^{-j} \right) \leq 3s,
\]

which completes the proof. \( \square \)

As the regularity of \( x \) is equivalent to \( \tau_t = o(n_t) \), we have the following conclusion.

**Corollary A.2.** A Toeplitz sequence is regular if and only if \( \delta(E') = o(n_t) \).

### Appendix B. Sturmian Dynamical Systems Satisfy a PNT

Let \( T : \mathbb{T} \to \mathbb{T} \ (\mathbb{T} := \mathbb{R}/\mathbb{Z}) \) be an irrational rotation on \( \mathbb{T} \) by \( \alpha \). For every non-zero \( \beta \in \mathbb{T} \) let \( \{A_0, A_1\} \) be the partition given by the intervals \( A_0 = [0, \beta) \) and \( A_1 = [\beta, 1) \). For every \( x \in \mathbb{T} \) denote by \( \bar{x} \in \{0, 1\}^\mathbb{Z} \) the code of \( x \) defined by \( \bar{x}(k) = i \) if and only if \( T^k x \in A_i \). Finally, denote by \( X_{\alpha, \beta} \subset \{0, 1\}^\mathbb{Z} \) the closure of the set \( \{\bar{x} \in \{0, 1\}^\mathbb{Z} : x \in \mathbb{T}\} \). Since \( X_{\alpha, \beta} \) is an invariant subset for the left shift \( S \) on \( \{0, 1\}^\mathbb{Z} \), we can focus the topological dynamical system \( S : X_{\alpha, \beta} \to X_{\alpha, \beta} \).

**Theorem B.1.** For the topological dynamical system \( S : X_{\alpha, \beta} \to X_{\alpha, \beta} \) a PNT holds.

**Proof.** For every \( y = (y(n))_{n \in \mathbb{Z}} \in X_{\alpha, \beta} \) the set \( \bigcap_{n \in \mathbb{Z}} \overline{X_{y(n)}} \subset \mathbb{T} \) has exactly one element \( \pi(y) \in \mathbb{T} \). Moreover, \( \pi : X_{\alpha, \beta} \to \mathbb{T} \) is a continuous map intertwining \( S \) and \( T \) and there exists a unique \( S \)-invariant probability measure \( \mu \) on \( X_{\alpha, \beta} \). The \( \pi \)-image of \( \mu \) coincides with Lebesgue measure on \( \mathbb{T} \).

By Vinogradov’s theorem, for any character \( f(x) = e^{2\pi i n x}, n \in \mathbb{Z} \), we have

\[
\lim_{N \to \infty} \frac{1}{\pi(N)} \sum_{p < N} f(T^p x) = \int_{\mathbb{T}} f(x) \, dx \quad \text{for every } x \in \mathbb{T}.
\]

Since every continuous function \( f : \mathbb{T} \to \mathbb{C} \) is uniformly approximated by trigonometric polynomials, (49) holds also for any continuous \( f \). Moreover, (49) holds for any Riemann integrable \( f : \mathbb{T} \to \mathbb{R} \). Indeed, for every \( \varepsilon > 0 \) there are two continuous functions \( f_-, f_+ : \mathbb{T} \to \mathbb{R} \) such that \( f_-(x) \leq f(x) \leq f_+(x) \) for every \( x \in \mathbb{T} \) and
\[ \int_{T}(f_+(x) - f_-(x)) \, dx < \varepsilon. \] It follows that
\[
\limsup_{N \to \infty} \frac{1}{\pi(N)} \sum_{p < N} f(T^px) \leq \lim_{N \to \infty} \frac{1}{\pi(N)} \sum_{p < N} f_+(T^px) = \int_{T} f_+(x) \, dx < \int_{T} f(x) \, dx + \varepsilon
\]
and
\[
\liminf_{N \to \infty} \frac{1}{\pi(N)} \sum_{p < N} f(T^px) \geq \lim_{N \to \infty} \frac{1}{\pi(N)} \sum_{p < N} f_-(T^px) = \int_{T} f_-(x) \, dx > \int_{T} f(x) \, dx - \varepsilon.
\]
As \( \varepsilon > 0 \) can be chosen freely, this gives (49).

Suppose that \( f : X_{\alpha,\beta} \to \mathbb{R} \) depends only on finitely many coordinates. More precisely, assume that \( f(y) = g(y(-n), \ldots, y(n)) \) for some \( g : \{0,1\}^{2n+1} \to \mathbb{R} \). Then there exists \( F : \mathbb{T} \to \mathbb{R} \) such that \( f = F \circ \pi \) and \( F \) is constant on the atoms of the partition \( \bigvee_{n=0}^{\infty} T^{-1}\{A_0, A_1\} \) (for example, if \( n = 0 \) and \( f \) is the characteristic function of \( \{y \in X_{\alpha,\beta} : y(0) = 0\} \) then \( F \) is \( I_{A_0} \)). It follows that \( F \) is Riemann integrable. Therefore, for every \( y \in X_{\alpha,\beta} \), we have
\[
\frac{1}{\pi(N)} \sum_{p < N} f(S^p y) = \frac{1}{\pi(N)} \sum_{p < N} F(T^p \pi(y)) \to \int_{T} F(x) \, dx = \int_{X_{\alpha,\beta}} f \, d\mu.
\]
Since every continuous function \( f : X_{\alpha,\beta} \to \mathbb{R} \) is uniformly approximated by functions depending on finitely many coordinates,
\[
\frac{1}{\pi(N)} \sum_{p < N} f(S^p y) \to \int_{X_{\alpha,\beta}} f \, d\mu \text{ for any } y \in X_{\alpha,\beta}
\]
holds for every continuous \( f \).

References


P. Sarnak, *Möbius randomness and Dynamics six years later* at CIRM at 1h 08 minute https://library.cirm-math.fr/Record.htm?idlist=1&record=19282918124910001909


Faculty of Mathematics and Computer Science, Nicolaus Copernicus University, ul. Chopina 12/18, 87-100 Toruń, Poland

E-mail address: fraczek@mat.umk.pl

Department of Mathematics, University of Maryland, 4176 Campus Drive, William E. Kirwan Hall, College Park, MD 20742-4015, USA

E-mail address: akanigow@umd.edu

Faculty of Mathematics and Computer Science, Nicolaus Copernicus University, ul. Chopina 12/18, 87-100 Toruń, Poland

E-mail address: mlem@mat.umk.pl