DENSITY OF SOME SEQUENCES MODULO 1

ARTŪRAS DUBICKAS

Abstract. Recently, Cilleruelo, Kumchev, Luca, Rué and Shparlinski proved that for each integer \( a \geq 2 \) the sequence of fractional parts \( \{a^n/n\}_{n=1}^{\infty} \) is everywhere dense in the interval \([0, 1]\). We prove a similar result for all Pisot numbers and Salem numbers \( \alpha \) and show that for each \( c > 0 \) and each sufficiently large \( N \) every subinterval of \([0, 1]\) of length \( cN^{-0.475} \) contains at least one fractional part \( \{Q(\alpha^n)/n\} \), where \( Q \) is a nonconstant polynomial in \( \mathbb{Z}[z] \) and \( n \) is an integer satisfying \( 1 \leq n \leq N \).

1. Introduction

Throughout, let \( \{x\} \) be the fractional part of \( x \in \mathbb{R} \). In a recent paper [3] Cilleruelo, Kumchev, Luca, Rué and Shparlinski showed that for each integer \( a \geq 2 \)

\[
\text{the sequence } \{a^n/n\}_{n=1}^{\infty} \text{ is everywhere dense in the interval } [0, 1] \tag{1}
\]

and, furthermore, for any \( c > 0 \) and any sufficiently large integer \( N \) every interval \( J \subseteq [0, 1] \) of length \( |J| \geq cN^{-0.475} \) contains an element of this sequence with the index \( n \) satisfying \( 1 \leq n \leq N \). In the proof of (1) they considered a subsequence \( A \) of the sequence \( \{a^n/n\}_{n=1}^{\infty} \) with indices \( n = pq \), where both \( p \) and \( q \) are primes satisfying \( q \leq \log p/\log a \). Using exponential sums and other tools from analytic number theory they first proved an upper bound for the discrepancy of the sequence \( A \) which implies (1) (see Theorem 1 in [3]) and then gave an alternative (much shorter) argument which implies (1) as well (see Theorem 2 in [3]). The main result of this note (see Theorem 2 below) generalizes Theorem 2 of [3].

A reader familiar with the literature in analytic number theory from the constant 0.475 and the fact primes numbers are involved in \( A_1 \) may guess that the authors of [3] used some results on gaps between consecutive primes. A well-known result of Baker, Harman and Pintz [1] asserts there is a constant \( \theta < 0.525 \) such that for each sufficiently large \( x \) the interval \((x - x^\theta, x)\) contains a prime number. (Note that 0.475 = 1 − 0.525.) We shall use a version of this result which follows from a more general Lemma 2 of [3] (which itself is extracted from Theorem 10.8 in [7]):

2000 Mathematics Subject Classification. 11K06, 11K31, 11R06.

Key words and phrases. Distribution modulo 1, gaps between primes.
Lemma 1. If $C$ is a positive constant and $h$ is a positive integer satisfying $h \leq (\log x)^C$ then for each sufficiently large $x$ the interval $(x - x^\theta, x)$, where $\theta < 0.525$ is some constant, contains a prime number which is equal to 1 modulo $h$.

Before stating our result we recall that an algebraic integer $\alpha > 1$ is a Pisot number (resp. a Salem number) if all of its conjugates over $\mathbb{Q}$ (if any) lie strictly inside the unit circle $|z| = 1$ (resp. in the disc $|z| \leq 1$ with at least one conjugate lying on the circle $|z| = 1$). See [2] for some basic properties of Pisot and Salem numbers. For example, all rational integers greater than or equal to 2, the golden section $(1 + \sqrt{5})/2 = 1.61803\ldots$ and the number 1.32471\ldots which is a root of the polynomial $z^3 - z - 1$ are Pisot numbers. (Siegel proved that the latter is the smallest Pisot number [9].) The smallest known Salem number 1.17628\ldots is a root of the Lehmer polynomial $z^{10} + z^9 - z^7 - z^6 - z^5 - z^4 - z^3 + z + 1$.

We can now state the main result of this paper.

Theorem 2. If $\alpha$ is a Pisot number or a Salem number and $Q(z)$ is a nonconstant polynomial with integer coefficients then the sequence $\{Q(\alpha^n)/n\}_{n=1}^{\infty}$ is everywhere dense in $[0, 1]$. Furthermore, for any $c > 0$ and any sufficiently large integer $N$ every interval $J \subseteq [0, 1]$ of length $|J| \geq cN^{-0.475}$ contains at least one element of this sequence with the index $n$ in the range $1 \leq n \leq N$.

By the same method Theorem 2 can be proved for nonconstant polynomials $Q$ with rational coefficients. It would be of interest to extend this result to sequences of the form $\{Q(\alpha^n)/P(n)\}_{n=1}^{\infty}$, where $P \in \mathbb{Q}[z]$ is a polynomial of degree at least 2, e.g., to the sequence $\{2^n/(n^3 + 1)\}_{n=1}^{\infty}$.

2. Preparation for the proof of Theorem 2

We begin with a short proof of (1) (following [3], i.e. taking $n = pg$, although without assuming that $g$ is a prime) and then continue the proof of Theorem 2 along the same lines with a more subtle choice of $g$ (see (4) and (11)) and $p$.

To prove (1) it suffices to show that the sequence $\{a^n/n\}_{n=1}^{\infty}$, where $a > 2$ is an integer, is everywhere dense in the open interval $(0, 1)$. Fix any $\lambda$ in the interval $(0, 1)$. We will show that for each $\varepsilon$ satisfying $0 < \varepsilon < \lambda$ there is $n \in \mathbb{N}$ of the form $n = pg$, where $g$ is a large integer and $p$ is a prime number, such that $\lambda - \varepsilon < \{a^n/n\} < \lambda$. Indeed, for each sufficiently large integer $g > g_0(a, \lambda, \varepsilon)$ (which is assumed to be relatively prime to $a$) there is a prime number $p > g$ which satisfies

$$\frac{a^g}{g\lambda} < p < \frac{a^g}{g(\lambda - \varepsilon)}$$

(2)
and $\varphi(g)|(p - 1)$, where $\varphi(g)$ is Euler's function. With this choice of $p$ and $g$, by Euler's theorem, we see that the difference $a^{(p-1)g} - 1$ is divisible by $p$ and by $g$. Hence their product $pg$ divides $a^{pg} - a^g$. Using (2) we find that for $n = pg$

$$\{a^n/n\} = \{a^{pg}/pg\} = \{a^g/pg\} = a^g/pg \in (\lambda - \varepsilon, \lambda),$$

as claimed.

Let $\alpha_1 = \alpha, \alpha_2, \ldots, \alpha_d$ be the full set of conjugates of $\alpha$ over $\mathbb{Q}$ with minimal polynomial

$$F(z) = (z - \alpha_1) \cdots (z - \alpha_d) = z^d + b_{d-1}z^{d-1} + \cdots + b_0 \in \mathbb{Z}[z].$$

Put

$$S_n := \alpha_1^n + \cdots + \alpha_d^n \quad \text{and} \quad R_n := S_n - \alpha_1^n = S_n - \alpha^n. \quad (3)$$

Note that, by the Newton formula,

$$S_n + b_{d-1}S_{n-1} + \cdots + b_0S_{n-d} = 0$$

for each integer $n \geq d + 1$.

Suppose that $g$ is a positive integer satisfying

$$\gcd(b_0, g) = 1. \quad (4)$$

Then $(S_n)_{n=1}^\infty$ is a sequence of integers which is purely periodic modulo $g$ with period $h$ in the range $1 \leq h \leq g^d$. (This result is known and can be easily proved in few lines; see, for instance, Lemma 2 in [5]). In particular, it implies that

$$g|(S_l - S_k) \quad \text{if} \quad h|(l - k). \quad (5)$$

Another useful result concerning $S_n$ is that

$$p|(S_{pk} - S_k) \quad (6)$$

for every $k \in \mathbb{N}$ and every prime number $p$. This is an old result of Schönemann proved in 1839 (see [8]) and then several times rediscovered by different authors. See, e.g., [4] and also [6], [10] for some generalizations; e.g., the latter paper contains the proof of $n|\sum_{t|n} \mu(n/t)S_{tk}$ for each $n \in \mathbb{N}$, where $\mu$ is the Möbius function, which gives (6) when $n$ is a prime number. We remark that the properties (5) and (6) hold for all algebraic integers $\alpha$ (and not just for Pisot and Salem numbers).

Let

$$Q(z) = a_tz^t + \cdots + a_0 \in \mathbb{Z}[z],$$
where $t \in \mathbb{N}$ and $a_t \neq 0$. Without restriction of generality we may assume that $a_t > 0$, since otherwise one can consider the polynomial $-Q$ instead of $Q$. Put

$$D_n := Q(\alpha^n) - \sum_{j=1}^{t} a_j S_{jn}. \quad (7)$$

From (3) and (7) it follows that

$$D_n = a_0 - \sum_{j=1}^{t} a_j R_{jn}. \quad (8)$$

As we already observed above, for any positive integer $g$ as in (4), the sequence $(S_n)_{n=1}^\infty$ is purely periodic modulo $g$ with period $h \leq g^t$. Assume that $p > g$ is a prime which is equal to 1 modulo $h$. Take $n = pg$. Then $p|(S_{jg} - S_{jg})$, by (6). Also, $g|(S_{jg} - S_{jg})$, by (5), because $jg - jg = jg(p - 1)$ is divisible by the period $h$. Hence $pg|(S_{jg} - S_{jg})$, because $\gcd(p, g) = 1$. It follows that $pg$ divides the difference between $\sum_{j=1}^{t} a_j S_{jg}$ and $\sum_{j=1}^{t} a_j S_{jg}$. Thus, if $g < p$ is a positive integer satisfying (4) then in view of (7) we obtain

$$\{Q(\alpha^{pg})/pg\} = \{(pg)^{-1} D_{pg} + (pg)^{-1} \sum_{j=1}^{t} a_j S_{jg}\} = \{y(p)\}, \quad (9)$$

where

$$y(p) := (pg)^{-1}(D_{pg} + \sum_{j=1}^{t} a_j S_{jg}). \quad (10)$$

In the next section we will select appropriate prime numbers $p$ and using (9) complete the proof of Theorem 2.

### 3. Proof of Theorem 2

Fix a large positive integer $N$ and take the largest $g \in \mathbb{N}$ satisfying (4) for which

$$\sum_{j=1}^{t} a_j S_{jg} - K \leq N \quad (11)$$

with $K$ given in (8). Observe that the main term of the expression on the left hand side of (11) is $a_t \alpha^{tg}$ and at least one of $|b_0|$ consecutive integers $g$ satisfies the condition (4). Hence there are two positive constants $c_1 \leq 1$ and $c_2$ (depending on $t, a_t, \alpha$ and $b_0$ only and
not on \( N \) such that
\[
c_1 N \leq \sum_{j=1}^{t} a_j S_{jg} - K \tag{12}
\]
and
\[
g \leq c_2 \log N \tag{13}
\]
for \( N \) large enough. In particular, in view of \( h \leq g^d \) the inequality (13) implies that
\[
h \leq (\log N)^{d+1} \tag{14}
\]
for each sufficiently large \( N \).

For \( g \) chosen as in (11) we set
\[
L_1 := (g^{-1} \sum_{j=1}^{t} a_j S_{jg} + g^{-1} K)/2, \quad L_2 := g^{-1} \sum_{j=1}^{t} a_j S_{jg} - g^{-1} K. \tag{15}
\]
Clearly, by (11), (12) and (15),
\[
c_1 N/g \leq L_2 \leq N/g \tag{16}
\]
and, since \( 2L_1 = L_2 + 2K/g \),
\[
c_1 N/2g \leq L_1 \leq (N + 2K)/2g. \tag{17}
\]
Let \( p_1 < \cdots < p_s \) be all the primes which are equal to 1 modulo \( h \) and are greater than \( L_1 \) and smaller than \( L_2 \). Then, by (16), we have \( p_s < L_2 \leq N/g \) and, by (17), \( p_1 > L_1 \geq c_1 N/2g \). Hence
\[
c_1 N/2 < p_1 g < \cdots < p_s g < N. \tag{18}
\]
Note that \( p_1 > g \), by (13) and (18), so the formula (9) holds for the primes \( p_1, \ldots, p_s \).

Now, for each \( p \in \{p_1, \ldots, p_s \} \) using (8), (10) and (15) we find that
\[
y(p) \geq (pg)^{-1}(-K + \sum_{j=1}^{t} a_j S_{jg}) = L_2/p \geq L_2/p_s > 1.
\]
Similarly,
\[
y(p) \leq (pg)^{-1}(K + \sum_{j=1}^{t} a_j S_{jg}) = 2L_1/p \leq 2L_1/p_1 < 2.
\]
Hence (9) yields
\[
\{Q(\alpha^{pg})/pg\} = y(p) - 1.
\]
for each \( p \in \{p_1, \ldots, p_s\} \).

By (18), all the integers \( p_1 g, \ldots, p_s g \) are smaller than \( N \). We will show that for any \( c > 0 \) and any sufficiently large integer \( N \) every interval \( J \subseteq [0, 1] \) of length \( |J| > cN^{-0.475} \) contains at least one number \( \{Q(\alpha^n)/pg\} = y(p) - 1 \) with \( p \in \{p_1, \ldots, p_s\} \). For a contradiction, suppose that there is an interval \( J \subseteq [0, 1] \) of length \( cN^{-0.475} \) which contains no numbers of the form \( y(p) - 1 \) with \( p \in \{p_1, \ldots, p_s\} \). Our aim is to show that the number \( y(p_s) - 1 \) is ‘very close’ to 0, the number \( y(p_1) - 1 \) is ‘very close’ to 1 and, moreover, the difference between two consecutive values \( y(p_i) - 1 \) and \( y(p_{i+1}) - 1 \) is ‘very small’ too. If this is the case then moving from \( i = 1 \) (with \( y(p_1) - 1 \) being almost the right endpoint of the interval \([0,1]\)) to \( i = s \) (with \( y(p_s) - 1 \) being almost the left endpoint of the interval \([0,1]\)) step by step we will get values all over the interval \([0,1]\) lying in every interval of length \( cN^{-0.475} \).

Indeed, observe first that, by (8), (10) and (15),

\[
y(p_s) = (p_s g)^{-1} (D_{p_s g} + \sum_{j=1}^{t} a_j S_{jg}) \leq (p_s g)^{-1} (K + \sum_{j=1}^{t} a_j S_{jg}) = 2L_1/p_s = L_2/p_s + 2K/p_s g.
\]

By Lemma 1, we have \( L_2 - L_2^\theta < p_s < L_2 \) with \( \theta < 0.525 \). Using (13) and (16) we find that

\[
0 < y(p_s) - 1 < \frac{L_2 + 2K/g}{L_2 - L_2^\theta} - 1 = \frac{L_2^\theta + 2K/g}{L_2 - L_2^\theta} < cN^{-0.475} \tag{19}
\]

in view of \( \theta < 0.525 \). Similarly, as

\[
y(p_1) = (p_1 g)^{-1} (D_{p_1 g} + \sum_{j=1}^{t} a_j S_{jg}) \geq (p_1 g)^{-1} (-K + \sum_{j=1}^{t} a_j S_{jg}) = L_2/p_1,
\]

and, by Lemma 1, \( L_1 < p_1 < L_1 + L_1^\theta \), applying (13) and (17) we find that

\[
2 - y(p_1) < 2 - \frac{L_2}{L_1 + L_1^\theta} = 2 - \frac{2L_1 - 2K/g}{L_1 + L_1^\theta} = \frac{2L_1^\theta + 2K/g}{L_1 + L_1^\theta} < cN^{-0.475}.
\]

Thus

\[
1 - cN^{-0.475} < y(p_1) - 1 < 1. \tag{20}
\]

From (19) and (20) it follows that if such an interval \( J \) of length \( cN^{-0.475} \) (which contains no numbers of the form \( y(p) - 1 \) with \( p \in \{p_1, \ldots, p_s\} \)) exists then \( J = [u, v] \) with \( y(p_s) - 1 < u \) and \( v < y(p_1) - 1 \). Moreover, for some \( i \in \{1, \ldots, s - 1\} \) the distance between two consecutive points \( y(p_i) - 1 \) and \( y(p_{i+1}) - 1 \) must be greater than \( cN^{-0.475} \). So for a
contradiction it suffices to show that

$$|y(p_{i+1}) - y(p_i)| < cN^{-0.475}$$

for each $i \in \{1, \ldots, s - 1\}$.

Since, by (10),

$$y(p_{i+1}) - y(p_i) = (p_{i+1}g)^{-1}(D_{p_{i+1}g} + \sum_{j=1}^{t} a_j S_{jg}) - (p_i g)^{-1}(D_{p_i g} + \sum_{j=1}^{t} a_j S_{jg}),$$

from $|D_{p_{i+1}g}|, |D_{p_i g}| \leq K$ it follows that

$$|y(p_{i+1}) - y(p_i)| \leq \frac{K}{p_{i+1}g} + \frac{K}{p_i g} + \frac{(p_{i+1} - p_i)|\sum_{j=1}^{t} a_j S_{jg}|}{p_{i+1}p_i g} < \frac{2K}{p_i g} + \frac{(p_{i+1} - p_i)|\sum_{j=1}^{t} a_j S_{jg}|}{p_i^2 g}.$$

From (18) we see that the first term, $2K/p_i g$, is less than $c_3/N$. Using $p_{i+1} - p_i < p_i^\theta$ (see Lemma 1) and (11), (12) we can bound the second term:

$$\frac{(p_{i+1} - p_i)|\sum_{j=1}^{t} a_j S_{jg}|}{p_i^2 g} < \frac{p_i^\theta (N + K)}{p_i^2} = \frac{N + K}{p_i^{2-\theta}} \leq \frac{N + K}{p_1^{2-\theta}}.$$

In view of (13) and (18) this second term is less than

$$\frac{N + K}{(c_1N/2g)^{2-\theta}} < \frac{(\log N)^2}{N^{1-\theta}}.$$

Therefore, as $\theta < 0.525$, we conclude that for $N$ large enough

$$|y(p_{i+1}) - y(p_i)| < \frac{c_3}{N} + \frac{(\log N)^2}{N^{1-\theta}} < cN^{-0.475},$$

as claimed. This completes the proof of Theorem 2.

Acknowledgment. I thank the referee for recommending various improvements in exposition.

References


Department of Mathematics and Informatics, Vilnius University, Naugarduko 24, Vilnius LT-03225, Lithuania

E-mail address: arturas.dubickas@mif.vu.lt