On the fragmentary complexity of symbolic sequences*

P.Diamond[†], P.Kloeden[‡]
and
V.Kozyakin[§], A.Pokrovskii[§]

Abstract

A measure of the ability of a symbolic sequence to be coded by initial fragments of another symbolic sequence — its self-similarity measure — is introduced and its basic properties are investigated. The self-similarity measure of symbolic sequences associated with tori shift mappings corresponding to a special partitioning of a torus are then considered.

Introduction

The classical methods of symbolic dynamics reduce the investigation of a dynamical system to that of a shift operator on a space of infinite symbolic sequences with elements from a finite alphabet. An important characteristic of a dynamical system is the complexity of the symbolic sequences corresponding to its trajectories. A commonly used measure of this complexity

---

*This research has been supported by the Australian Research Council Grant A 89132609.

[†]Department of Mathematics, University of Queensland, Brisbane, Qld 4072, Australia

[‡]Department of Computing and Mathematics, Deakin University, Geelong, Victoria 3217, Australia

[§]Institute of Information Transmission Problems, Russian Academy of Sciences, 19 Ermolovoy str., Moscow 101447, Russia

is the ability of a sequence to be coded by finite words from a universal totality that does not depend on the system being investigated. The classical notions of entropy [9] and linear complexity of sequences [12] find a natural description within this framework.

In this paper a different, but related concept of complexity of infinite sequences based on [6, 7] will be studied. In particular, symbolic sequences $\mathbf{T}$ which can be partioned (perhaps, excluding an initial fragment) as the adjoint union of initial fragments of another sequence $\mathbf{U}$ will be considered and called self-similar. Theorems 1 and 2 below show that they do have a fractal structure as commonly understood. The complexity of such sequence $\mathbf{T}$ will be estimated by the minimal number C of samples of arbitrary long initial fragments of another sequence $\mathbf{U}$ that can cover the sequence $\mathbf{T}$ disjointly. A precise definition and its main properties will be given in Section 1.

Self-similar sequences arise naturally in applications such as the stability analysis of desynchronized systems [7, 10] (see also Theorem 3 below) as well as in description of fractal phenomena [3]. In particular, it has been shown that the self-similarity measure of sturmian sequences [8, 10] with irrational frequencies, such as the symbolic sequences representing shifts of the unit circle, is equal to 2. In Section 3 it will be shown that the properties of individual trajectories of multi-dimensional tori shifts of dimension higher than 2 changes drastically, with the self-similarity measure generally being infinite in such cases (Theorem 4). The situation for 2-dimensional tori shifts is still unclear.

In conclusion of the introduction the application of self-similarity measure of sequences to the stability analysis of frequency desynchronized systems [1, 2, 4, 6] mentioned above will now be briefly described. As seen from [4, 6] this stability problem reduces to that of the nonautonomous difference equation

$$x(n+1) = f[\lambda; n, x(n)], \qquad n = 0, 1, 2, \ldots\ , \tag{1}$$

with the right-hand side $f(\lambda; n, x)$ non-periodic in n and depending on a parameter λ such that the number of different mappings in $\{f[\lambda; n, \cdot]\}$ is finite. Here the order of different mappings in the sequence $\{f[\lambda; n, \cdot]\}$ corresponds to the order of symbols in a symbolic sequence generated by a certain shift mapping of a torus with a special partitioning. Using a concept similar to the fragmentary complexity of these symbolic sequences it was proved in [6, 7] that the asymptotic stability of equation (1) for one particular value of the parameter λ implies its stability for the other values of λ.

1 A measure of fragmentary complexity

1.1 Weakly decomposable texts

Following [12] we shall use linguistic terminology and notation. In particular, elements in symbolic sequences will be not separated by commas. Let $\mathcal{A}$ be a fixed alphabet, that is a set of elements called *letters* or *symbols*. A finite cortege $\mathbf{w} = a_1 \dots a_n$ of letters from $\mathcal{A}$ is called *a word*, for any words $\mathbf{w}^1 = a_1^1 \dots a_{n_1}^1$ and $\mathbf{w}^2 = a_1^2 \dots a_{n_2}^2$ their *product* is the word $\mathbf{w}^1\mathbf{w}^2 = a_1^1 \dots a_{n_1}^1 a_1^2 \dots a_{n_2}^2$, and the *left factor (of the length $j \le n$)* of the word $\mathbf{w} = a_1 \dots a_n$ is the initial fragment $\mathbf{w}(j) = a_1 \dots a_j$ of $\mathbf{w}$. An infinite sequence $\mathbf{T} = a_1 a_2 \dots$ from the alphabet $\mathcal{A}$ is called *an infinite word* or *text*, the word $\mathbf{T}(n) = a_1 a_2 \dots a_n$ its *left factor (of the length n)* and the text $a_{n+1}a_{n+2}\dots$ its *right factor (of the colength n)*, while any word $a_i \dots a_j$ with $i \le j$ is called a *factor* of $\mathbf{T}$.

An ordered finite set

$$\mathbf{S} = \{\mathbf{w}_1, \dots, \mathbf{w}_\nu\} \tag{2}$$

of words of lengths $l_1, \dots, l_\nu$ is said to be *generating* if it satisfies the properties:

P1. $0 < l_1 < \dots < l_\nu$.

P2. The word $\mathbf{w}_\iota$ is a left factor of $\mathbf{w}_\nu$ for each $\iota = 1, \dots, \nu - 1$, that is $\mathbf{w}_\iota$ coincides with the initial segment of $\mathbf{w}_\nu$ of length l_ι.

A finite or infinite word $\mathbf{w}$ is $\mathbf{S}$*–decomposable* if it can be represented as a product of words belonging to a set of words (2), while a text $\mathbf{T}$ is *weakly* $\mathbf{S}$*–decomposable* if it has an $\mathbf{S}$–decomposable right factor. In other words, $\mathbf{T}$ is weakly $\mathbf{S}$–decomposable if there exists an increasing sequence $\mathbf{d} = \{d_0, d_1, d_2, \dots\}$ of natural numbers such that $r_i = d_i - d_{i-1}$ is equal to one of the numbers l_ι, for $\iota = 1, \dots, \nu$ and $\mathbf{w}_i = a_{d_{i-1}} \dots a_{d_i - 1}$; such a sequence $\mathbf{d}$ is *a weak* $\mathbf{S}$*–decomposition* of $\mathbf{T}$.

Now consider two texts $\mathbf{T}$ and $\mathbf{U}$. The text $\mathbf{T}$ will be called a $\mathbf{U}$*–generated* if for any N there exists a finite generating set $\mathbf{S}$ of left factors of the text

$\mathbf{U}$ such that all words $\mathbf{w} \in \mathbf{S}$ are of length greater than N and the text $\mathbf{T}$ is weakly $\mathbf{S}$–decomposable. An periodic text $\mathbf{T}$ is clearly $\mathbf{T}$–generated, or *self-generative.* where $\mathbf{U}$ is a periodic part of $\mathbf{T}$, but as will be seen in Subsection 1.3 below there also exist exist self-generative texts with much more complicated structure. The fact that a text $\mathbf{T}$ is $\mathbf{U}$–generated for a certain text $\mathbf{U}$ can be useful. For example, if a text $\mathbf{U}$ is ergodic in the sense that for all $a \in \mathcal{A}$ the limiting frequencies $q_n(a)$ exist, where $q_n(a)$ is the number of times the letter a occurs in $\mathbf{U}(n)$, then any $\mathbf{U}$–generated text $\mathbf{T}$ is also ergodic with the same limiting frequencies. This was used in the analysis of desynchronized systems [2].

Denote by $\mathcal{S}(\mathbf{T},\mathbf{U})$ the family of all finite generating sets $\mathbf{S}$ of left factors of the text $\mathbf{U}$ for which the text $\mathbf{T}$ is weakly $\mathbf{S}$–decomposable and by by $\mathcal{S}_*(\mathbf{T},\mathbf{U})$ the totality of elements of $\mathcal{S}(\mathbf{T},\mathbf{U})$ of the form (2) which satisfy the additional property:

P3. For each $\iota = 1,\ldots,\nu-1$ the word $\mathbf{w}_\iota$ is not a power, that is cannot be partitioned into repeating fragments.

Theorem 1 *Let a text* $\mathbf{T}$ *be* $\mathbf{U}$*–fractal. Let* $\mathbf{S}^{short} \in \mathcal{S}_*(\mathbf{T},\mathbf{U})$, $\mathbf{S}^{long} \in \mathcal{S}(\mathbf{T},\mathbf{U})$ *and suppose that the shortest word from* $\mathbf{S}^{long}$ *is longer than the longest word from* $\mathbf{S}^{short}$*. Then every word from* $\mathbf{S}^{long}$ *is* $\mathbf{S}^{short}$*–decomposable and each weak* $\mathbf{S}^{long}$*–decomposition* $\mathbf{d}^{long}$ *is a subset of any weak* $\mathbf{S}^{short}$*–decomposition* $\mathbf{d}^{short}$ *satisfying* $d_0^{short} \le d_0^{long}$*.*

Proof. Suppose the opposite. Then there exists a number $d \in \mathbf{d}^{long}$ and an index I such that $d_I^{short} < d < d_{I+1}^{short}$. Write $\mathbf{w}^1 = a_{d_I^{short}} \ldots a_{d-1}$ and $\mathbf{w}^2 = a_d \ldots a_{d_{I+1}^{short}-1}$. By property P1 and the assumptions of the theorem we have $\mathbf{w}^1\mathbf{w}^2 = \mathbf{w}^2\mathbf{w}^1$. Hence, by Proposition 1.3.2 from [12] the word $\mathbf{w} = a_{d_I^{short}} \ldots a_{d_{I+1}^{short}-1}$ is a power. By the construction, this word belongs to a generating set, but this contradicts property P3 of $\mathcal{S}_*(\mathbf{T},\mathbf{U})$. ∎

Informally speaking, Theorem 1 says that every $\mathbf{U}$–decomposition can be considered as the result of a partitioning of some "bigger" $\mathbf{U}$–decomposition.

Example 1 *Let* $\mathcal{A} = \{a,b\}$,

$$\mathbf{U} = abbababb\ldots, \qquad \mathbf{T} = babbabbababbababbabbab\ldots$$

and let $\mathbf{S}^{short} = \{\mathbf{w}_1^{short}, \mathbf{w}_2^{short}\}$, $\mathbf{S}^{long} = \{\mathbf{w}_1^{long}, \mathbf{w}_2^{long}\}$, *where*

$$\mathbf{w}_1^{short} = ab\,, \quad \mathbf{w}_2^{short} = abb\,, \quad \mathbf{w}_1^{long} = abbab\,, \quad \mathbf{w}_2^{long} = abbabab\,.$$

Then the following decomposition of $\mathbf{T}$ *is valid:*

$$\mathbf{T} = b\ \underbrace{abb}_{\mathbf{w}_2^{short}}\ \overbrace{\underbrace{abb}_{\mathbf{w}_2^{short}}\ \underbrace{ab}_{\mathbf{w}_1^{short}}}^{\mathbf{w}_1^{long}}\ \overbrace{\underbrace{abb}_{\mathbf{w}_2^{short}}\ \underbrace{ab}_{\mathbf{w}_1^{short}}\ \underbrace{abb}_{\mathbf{w}_2^{short}}}^{\mathbf{w}_2^{long}}\ \overbrace{\underbrace{abb}_{\mathbf{w}_2^{short}}\ \underbrace{ab}_{\mathbf{w}_1^{short}}}^{\mathbf{w}_1^{long}}\ \ldots$$

The text $\mathbf{U}$ will be called *self-generative* if for any N there exists a finite generating set $\mathbf{S}$ of left factors of $\mathbf{U}$ itself such that all words $\mathbf{w} \in \mathbf{S}$ are of length greater than N and the text $\mathbf{U}$ is $\mathbf{S}$–decomposable. As a the corollary of Theorem 1 we have:

Corollary 1 *A text* $\mathbf{U}$ *is self-generative if and only if there exists a* $\mathbf{U}$*–generated text* $\mathbf{T}$*.*

PROOF. If the text $\mathbf{T}$ is periodic after a certain index N then text $\mathbf{U}$ must be also periodic and there is nothing to prove. Consider the case when the text $\mathbf{T}$ is not eventually periodic. Let $\mathbf{S} \in \mathcal{S}_*(\mathbf{T}, \mathbf{U})$ be a generating set for the text $\mathbf{T}$ consisting of ν left factors $\mathbf{U}(l_1), \ldots, \mathbf{U}(l_\nu)$. The corollary will be proven if we establish that the text $\mathbf{U}$ is $\mathbf{S}$–decomposable. Consider the sequence of originating for $\mathbf{T}$ sets

$$\mathbf{S}_n = \Big(\mathbf{w}_1, \ldots, \mathbf{w}_{\nu(n)}\Big) \tag{3}$$

which satisfy the following conditions:

Q1. Each element of any set $\mathbf{S}_n$ is a left factor of $\mathbf{U}$.

Q2. The length of the shortest word in $\mathbf{S}_n$ is greater than n.

By Theorem 1 for $n \geq l_\nu$ all words from the set (3) are $\mathbf{S}$–decomposable. Denote the corresponding decomposition by

$$\mathbf{d}^{n,\iota} = d_1^{n,\iota}, \ldots, \delta_{m(n,\iota)}^{n,\iota}, \quad \iota = 1, \ldots, \nu(n) \tag{4}$$

and denote by $\mathbf{d}^*$ the sequence that is a limit point of the sequence (4) in the topology of point-wise convergence. By the construction, the sequence $\mathbf{d}^*$ is a $\mathbf{S}$–decomposition of $\mathbf{U}$, and the corollary is proven. ∎

Self-generative texts have an important property of being recurrent. For each text $\mathbf{T}$ denote by $\mathcal{W}(\mathbf{T}, n)$ the totality of words $a_i a_{i+1} \dots a_{i+n-1}$, $i = 1, 2, \dots$. A text $\mathbf{T}$ is said to be *recurrent* [10] if for each natural number m there exists a natural number n such that any word from $\mathcal{W}(\mathbf{T}, m)$ is a factor of words from $\mathcal{W}(\mathbf{T}, n)$.

Lemma 1 *Each self-generative text* $\mathbf{U}$ *is recurrent.*

PROOF. Choose a natural number N such that all words from $\mathcal{W}(\mathbf{U}, m)$ are factors of $\mathbf{U}(N)$. Consider a generating set $\mathbf{S}$ of left factors of $\mathbf{U}$ such that $\mathbf{U}$ is $\mathbf{S}$–decomposable and all words from $\mathbf{S}$ are longer than N. Let L denote the length of longer word in $\mathbf{S}$. By construction every word from $\mathcal{W}(\mathbf{U}, m)$ is a factor of each word from $\mathcal{W}(\mathbf{U}, L)$, and so the lemma is proven. ∎

The general construction of self-generative texts to be presented in Subsection 1.3 thus provides a means of constructing recurrent texts.

1.2 Fragmentary complexity of texts

Let a text $\mathbf{T}$ be $\mathbf{U}$–generated. Denote by $\mathcal{S}(\mathbf{T}, \mathbf{U}; N)$ the subset of $\mathcal{S}(\mathbf{T}, \mathbf{U})$ containing those generating sets $\mathbf{S}$ all words from which are longer than N. For any natural N it is defined the minimal quantity $C_f(\mathbf{T}, \mathbf{U}; N)$ of elements in sets from $\mathcal{S}(\mathbf{T}, \mathbf{U}; N)$. Clearly, the function $C_f(\mathbf{T}, \mathbf{U}; N)$ is increasing in N. It is naturally to characterize the complexity of the text $\mathbf{T}$ with respect to the text $\mathbf{U}$ by the rate of increase of this function.

In particular, of a special interest is the situation when this function is bounded, in which case we will call the number

$$C_f(\mathbf{T}, \mathbf{U}) = \max_N C_f(\mathbf{T}, \mathbf{U}; N) \tag{5}$$

the $\mathbf{U}$–*complexity* of the text $\mathbf{T}$. It is convenient to set $C_f(\mathbf{T}, \mathbf{U}) = \infty$ if the function $C_f(\mathbf{T}, \mathbf{U}; N)$ is unbounded or if $\mathbf{T}$ is not $\mathbf{U}$–generated. If the text $\mathbf{T}$ has a finite $\mathbf{U}$–complexity with respect to at least one text $\mathbf{U}$ then define $C_f(\mathbf{T}) = \min_{\mathbf{U}} C_f(\mathbf{T}, \mathbf{U})$. The quantity $C_f(\mathbf{T})$ will be called the *fragmentary complexity of* $\mathbf{T}$.

1.3 General construction of self-generative texts

We now describe a general construction an self-generative texts with fragmentary complexity not exceeding C. Let $\mathcal{A}_k$ be an alphabet with $k > 1$ letters, say $1, \dots, k$. If to every letter $\kappa \in \mathcal{A}_k$ there corresponds a word $\mathbf{w} = F(\kappa) \in \mathcal{W}(\mathcal{A})$, then to each word $\mathbf{v}$ of the alphabet $\mathcal{A}_k$ we associate a word $F(\mathbf{v})$ obtained by substituting the word $F(\kappa)$ for each letter κ in the word $\mathbf{v}$.

Let us now choose

- a natural number $\nu \le C$,
- a sequence $\mathcal{S}$ of generating sets $\mathbf{S}_n,\ n = 1, 2, \dots$ in the alphabet $\mathcal{A}_{\nu(n-1)}$ containing words $\mathbf{v}^{n,\iota},\ \iota = 1, \dots \nu(n)$, with lengths $l(\mathbf{v}^{n,\iota}) > n$;
- a particular generating set $\mathbf{S}_0^*$ of words of the alphabet $\mathcal{A}$ which contains ν elements.

Then we construct recursively the generating subsets

$$\mathbf{S}_n^* = \{\mathbf{w}^{n,1}, \dots, \mathbf{w}^{n,\nu(n)}\} \qquad n = 1, 2, \dots$$

in the alphabet $\mathcal{A}$. Suppose that $\mathbf{S}_{n-1}^*$ is already defined. Then define $F_n(\kappa) = \mathbf{w}^{n-1,\kappa}$ for $\kappa = 1, \dots, \nu(n-1)$ and set $\mathbf{w}^{n,\iota} = F_n(\mathbf{v}^{n,\iota}),\quad \iota = 1, \dots, \nu(n)$.

Example 2 *Let, for instance,* $S = \{2, 2, 2, \dots\}$, $\mathcal{A} = \{a, b\}$, $\mathbf{S}_0^* = (a, ab)$ *and*

$$\mathbf{S}_1 = \{1,\ 12\}, \quad \mathbf{S}_2 = \{21,\ 211\}, \quad \mathbf{S}_3 = \{121,\ 1211\}.$$

Then

$$F_1(1) = a,\ F_1(2) = ab \quad \text{and} \quad \mathbf{S}_1^* = \{\underbrace{a}_1,\ \underbrace{a}_1\underbrace{ab}_2\}.$$

Analogously,

$$F_2(1) = a,\ F_2(2) = aab \quad \text{and} \quad \mathbf{S}_2^* = \{\underbrace{\underbrace{a}_1\underbrace{ab}_2}_2\underbrace{\underbrace{a}_1}_1,\ \underbrace{\underbrace{a}_1\underbrace{ab}_2}_2\underbrace{\underbrace{a}_1}_1\underbrace{\underbrace{a}_1}_1\}.$$

Further,

$$F_3(1) = aaba,\ F_3(2) = aabaa$$

and

$$\mathbf{S}_3^* = \{\underbrace{\underbrace{aab}_{2}\underbrace{a}_{1}}_{1}\underbrace{\underbrace{aab}_{2}\underbrace{a}_{1}\underbrace{a}_{1}}_{2}\underbrace{\underbrace{aab}_{2}\underbrace{a}_{1}}_{1},\ \underbrace{\underbrace{aab}_{2}\underbrace{a}_{1}}_{1}\underbrace{\underbrace{aab}_{2}\underbrace{a}_{1}\underbrace{a}_{1}}_{2}\underbrace{\underbrace{aab}_{2}\underbrace{a}_{1}}_{1}\underbrace{\underbrace{aab}_{2}\underbrace{a}_{1}}_{1}\}.$$

Clearly, $\mathbf{w}^{n,1}$ is a left factor of $\mathbf{w}^{n+1,1}$ and $\lim_{n\to\infty} l(\mathbf{w}^{n,1}) = \infty$. Therefore there exists a pointwise limit $\mathbf{U} = \mathbf{U}(\nu, \mathbf{S}_0^*, \mathcal{S})$ of the sequence of words $\mathbf{w}^{n,1}$ when $n \to \infty$.

Lemma 2 *Each text $\mathbf{U}(\nu, \mathcal{S}, \mathbf{S}_0^*)$ is self-generative of $\mathbf{U}$–complexity not exceeding ν. Moreover, each self-generative text $\mathbf{U}$ of $\mathbf{U}$–complexity C can be regarded as $\mathbf{U}(C, \mathcal{S}, \mathbf{S}_0^*)$ for appropriate $\mathcal{S}$ and $\mathbf{S}_0^*$.*

Proof. By construction each text $\mathbf{U}(\nu, \mathcal{S}, \mathbf{S}_0^*)$ is self-generative of $\mathbf{U}$–complexity no more than ν. Therefore, we need only prove that each self-generative text $\mathbf{U}$ of $\mathbf{U}$–complexity C coincides with a text $\mathbf{U}(C, \mathcal{S}, \mathbf{S}_0^*)$ for appropriate $\mathcal{S}$ and $\mathbf{S}_0^*$.

Consider the case where the text $\mathbf{U}$ is not periodic. Choose a certain set $\mathbf{S}_0^* = (\mathbf{U}(l_1^0), \ldots, \mathbf{U}(l_C^0)) \in \mathcal{S}_*(\mathbf{U}, \mathbf{U})$ based on $\mathbf{U}$. By definition there exists a sequence of such sets

$$\mathbf{S}_n^* = \{\mathbf{U}(l_1^n), \ldots, \mathbf{U}(l_C^n)\} \in \mathcal{S}_*(\mathbf{U}, \mathbf{U}) \tag{6}$$

for which $l_C^{n-1} \le l_1^n$, $n = 1, 2, \ldots$.

By Theorem 1 each word from $\mathbf{S}_n^*$ is $\mathbf{S}_{n-1}^*$–decomposable. Denote the respective decomposition by

$$\mathbf{d}^{n,\iota} = \{d_0^{n,\iota}, d_1^{n,\iota}, \ldots, d_{m(n,\iota)}^{n,\iota}\}, \quad \iota = 1, \ldots, C,$$

and introduce words $\mathbf{v}^{n,\iota} = v_1^{n,\iota} \ldots v_{m(n,\iota)}^{n,\iota}$, $i = 1, \ldots, C$, in the alphabet $\mathcal{A}_C$ by equalities $v_i^{n,\iota} = \kappa$ if and only if $l_i^{n,\iota} - l_{i-1}^{n,\iota} = l_\kappa^{n-1}$. Define $\mathbf{S}_n = \left(\mathbf{v}^{n,1}, \ldots, \mathbf{v}^{n,C}\right)$ and $\mathcal{S} = \{\mathbf{S}_1, \mathbf{S}_2, \ldots\}$. Then, by construction, $\mathbf{U} = \mathbf{U}(C, \mathbf{S}_0^*, \mathcal{S})$, which is the assertion of lemma. ∎

1.4 Texts with finite fragmentary complexity

For any alphabet $\mathcal{A}_*$ denote by $\mathcal{W}(\mathcal{A}_*)$ the totality of finite words in this alphabet. If there is a word $\mathbf{w} = F(a) \in \mathcal{W}(\mathcal{A}_*)$ for any letter $a \in \mathcal{A}$ then corresponding to the text $\mathbf{T}$ in the alphabet $\mathcal{A}$ denote the text $F(\mathbf{T})$ in the alphabet $\mathcal{A}_*$ be formed by substituting the word $F(a_i)$ for each letter a_i of the text $\mathbf{T}$. The text $\mathbf{T}$ is said to be *eventually periodic*, if it has a right factor which is periodic.

Lemma 3 *The following assertions are true*

(a) the fragmentary complexity of a text is equal to the fragmentary complexity of any of its right factors;

(b) the fragmentary complexity of a text is equal to 1 if and only if this text is eventually periodic;

(c) for any function $F : \mathcal{A} \mapsto \mathcal{W}(\mathcal{A}_*)$ and any text $\mathbf{T}$ in the alphabet $\mathcal{A}$ the complexity inequality $\mathcal{C}_f(\mathbf{T}) \geq C_f(F(\mathbf{T}))$ holds.

Another classical set of "simple" texts is the class of texts with linear complexity for subwords [12]. The text $\mathbf{T}$ is said to be of *linear complexity for subwords* if the number $\#(\mathbf{T}, N)$ of its subwords of the length N satisfies the bound

$$\sup_N \frac{\#(\mathbf{T}, N)}{N} < \infty. \tag{7}$$

Generally speaking, the properties of a text "to have finite fragmentary complexity" and "to be of linear complexity for subwords" do not follow one from another. Note that for texts of fragmentary complexity 2 the estimate

$$\lim_{K\to\infty} \inf_{N\geq K} \frac{\#(\mathbf{T}, N)}{N} < \infty \tag{8}$$

is always true. This is slightly weaker than (7). Note also that texts with the fragmentary complexity 2 always contain squares, i.e. repeated words one immediately next to other. It is not clear to us if there exist cube free words of fragmentary complexity 2 (probably, the well known Thue–Morse words [12] are not fractal).

Let us describe one more property of texts with fragmentary complexity 2. For any integer $\gamma \geq 0$ and any sequence $\mathbf{d}$ denote by $\mathrm{Pr}_\gamma(\mathbf{d})$ the subsequence of $\mathbf{d}$ consisting of elements of d_i with indices $d_i \geq \gamma$. Recall also that $\mathbf{U}(i)$ denotes the left factor of the length i of $\mathbf{U}$. Analogously to the Theorem 1 can be shown that:

Theorem 2 *Let a text* $\mathbf{T}$ *have* $\mathbf{U}$*–fragmentary complexity* 2 *and suppose that* $\mathbf{T}$ *is weakly* $(\mathbf{U}(i), \mathbf{U}(j))$*–decomposable where* $(\mathbf{U}(i), \mathbf{U}(j)) \in \mathcal{S}_*(\mathbf{T}, \mathbf{U})$*. Then for any two weak* $(\mathbf{U}(i), \mathbf{U}(j))$*–decompositions* $\mathbf{d}$ *and* $\mathbf{d}^*$ *the identity* $\mathrm{Pr}_L\, \mathbf{d} = \mathrm{Pr}_L\, \mathbf{d}^*$ *holds for* $L = \max\{d_0, d_0^*\} + i + j$*.*

2 Fragmentary complexity of tori shifts

2.1 The one-dimensional case

Consider the mapping S of the interval $[0, 1)$ onto itself defined by

$$S(x) = x + \varphi(x) \ (\mathrm{mod}\ 1)$$

where φ is a bounded 1–periodic function satisfying $|\varphi(x) - \varphi(y)| < |x - y|,\ x \neq y$ (see Fig. 1).

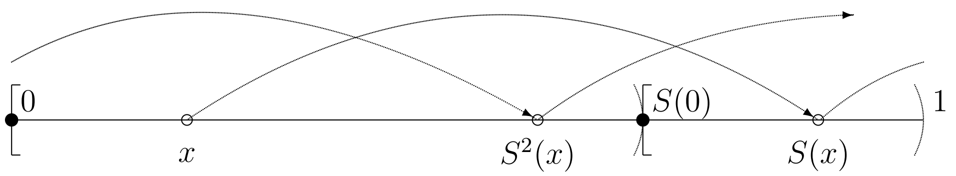

Figure 1: One-dimensional shift mapping

Each point $x \in [0, 1)$ generates a sequence $\{x_n\}$ defined by $x_0 = x$ and the recurrence relation $x_{n+1} = S(x_n)$, $n = 0, 1, \ldots$. The limit

$$\tau(S) = \lim_{n\to\infty} n^{-1} \sum_{k=1}^{k} \psi_k(x)$$

where $\psi_k(x) = \varphi(S^k(0))$, $k = 1, 2, \ldots$, exists and is independent of x. It is called [5] the *rotation number of the mapping* S. If, for instance

$$S(x) = S_\tau(x) = x + \tau \pmod 1 \tag{9}$$

where $\tau \in [0, 1)$ is a fixed real number then $\tau(S) = \tau$.

Suppose that corresponding to each point $x \in [0, 1)$ there is a symbolic sequence (text)

$$\mathbf{T}(x, S) = \sigma_0(x)\sigma_1(x)\ldots\sigma_n(x)\ldots \tag{10}$$

consisting of two letters, say a and b, where

$$\sigma_n(x) = \begin{cases} a, & \text{if } x_n = S^n(x) \in [0, S(0)), \\ b, & \text{if } x_n = S^n(x) \in [S(0), 1). \end{cases} \tag{11}$$

Texts (10) are called as *sturmian beams* with a–frequence $\tau(S)$ in [10]. Note that a different "internal" characterizsation of sturmian beams is proposed in [10].

If the value τ is rational then all texts (10) are, clearly, eventually periodic and by the assertion (b) of Lemma 3 the fragmentary complexity of each text $\mathbf{T}(x, S)$ with $x \in [0, 1)$ is equal to 1. The following result regarding the case of irrational τ is a corollary of Theorem 1 from [7] (see also Theorem 5 from [6]).

Theorem 3 *Let $\tau(S)$ be irrational and $x \in [0, 1)$ with $x \neq S(0)$. Then $\mathbf{T}(x, S)$–fragmentary complexity of each text $\mathbf{T}(y, S)$ with $y \in [0, 1)$ is equal to* 2.

2.2 The multi-dimensional case

The authors attempts to formulate an analogue of Theorem 3 for shift mappings of multi-dimensional tori have not been successful. Nevertheless, some interesting insights into why a direct generalization of this theorem is not possible have been obtained.

Let I^M be the unit multi-dimensional cube $[0, 1) \times [0, 1) \times \ldots \times [0, 1) = [0, 1)^M$ of the space $\mathbb{R}^M$, let $\tau = \{\tau_1, \tau_2, \ldots, \tau_M\}$ be a point in I^M and consider the shift mapping S_τ from the cube I^M onto itself defined by

$$S_\tau(x) = \{x_1 + \tau_1 \pmod 1), x_2 + \tau_2 \pmod 1), \ldots, x_M + \tau_M \pmod 1)\},$$

where $x = \{x_1, x_2, \ldots, x_M\} \in I^M$. In addition, denote by $\mathcal{U}$ the set of all subsets $U_i \subset I^M$, $i = 1, 2, \ldots, 2^M$, of the form $U_i = H_1 \times H_2 \times \ldots \times H_M$

where each H_j coincides either with $[0, \tau_j)$ or with $[\tau_j, 1)$. Finally, let a letter a_i correspond to each subset U_i and denote the text $\sigma_0(x)\sigma_1(x)\ldots\sigma_n(x)\ldots$ defined by the relations

$$\sigma_n(x) = a_{i_n} \quad \text{if} \quad S_{n\tau}(x) \in U_{i_n}$$

by $\mathbf{T}(x, \tau)$.

Note, that if $M = 1$ then introduced texts coincide with the sturmian beams generated by the mapping (9) The principal result to be proved in the paper indicates that a direct analog of Theorem 3 for multi-dimensional tori shifts is not valid:

Theorem 4 *The text* $\mathbf{T}(y, \tau)$ *is not* $\mathbf{T}(x, \tau)$*–fractal for almost all* $x, y \in I^M$ *and* $\tau \in \mathcal{T}$.

This result will be obtained as a corollary to another stronger (but also more cumbersome) result. We shall need some additional definitions in order to formulate this stronger result.

Given $x, \tau \in [0, 1)$, let $\mathcal{D}_m$ denote the set of all words $\mathbf{T}_n(x, \tau)$, $n \geq m$. How well can the text of some point $y \in [0, 1)$ be "coded" by words from $\mathcal{D}_m$? To solve this problem consider the text

$$\mathbf{T}(y, \tau) = \sigma_0(y)\sigma_1(y)\ldots\sigma_i(y)\ldots$$

and denote by $\mathcal{C}_m(y)$ the set of those indices i for which there are integers k_i, n_i with $0 \leq k_i \leq i \leq n_i$, such that the word $w_i = \sigma_{k_i}(y)\ldots\sigma_{n_i}(y)$ belongs to $\mathcal{D}_m$. Set

$$\Delta_{n,m}(y) = \frac{1}{n}\,\#\{\mathcal{C}_m(y) \bigcap [0, n-m)\},$$

where $\#(X)$ is the number of elements of the set X. Then, clearly,

$$(k+n)\Delta_{k+n,m}(y) \geq k\Delta_{k,m}(y) + n\Delta_{n,m}(y)$$

and hence that

$$(k+n)(1 - \Delta_{k+n,m}(y)) \leq k(1 - \Delta_{k,m}(y)) + n(1 - \Delta_{n,m}(y)).$$

From the latter inequality the existence of $\lim_{n\to\infty}(1-\Delta_{n,m}(y))$ follows. Then the limit $\Delta_m(y) = \lim_{n\to\infty} \Delta_{n,m}(y)$ also exists.

Theorem 5 *If $M \geq 3$, then* $\lim_{m\to\infty} \Delta_m(y) = 0$ *for almost all* $x, y \in I^M$, *and* $\tau \in [0, 1)$.

Theorem 4 follows immediately from Theorem 5.

We remark that in view of Theorem 3 $\Delta_m(y) = 1$ for any m in one-dimensional case. In fact, the statement of Theorem 3 is even stronger than this equality.

2.3 Remark

We suspect that a similar result will also hold for the case $M = 2$. If the below proof is any guide, its proof will, however, be complicated by the problem of small denominators.

3 Proof of Theorem 5

3.1 Auxiliary results

To prove Theorem 5 we shall need some auxiliary results. For $i = 1, 2, \ldots, M$ denote by L_i the hyperplanes

$$L_1 = \{x|\ x_1 = \tau_1\}, \quad L_2 = \{x|\ x_2 = \tau_2\}, \quad \ldots \quad , L_M = \{x|\ x_M = \tau_M\}.$$

Let $x \in I^M$ and let Ω be some region in I^M containing the point x and belonging to a particular subset $U_i \in \mathcal{U}$. Denote $\Omega_0 = \Omega$ and define recursively

$$\Omega_n(x) = S_\tau(\Omega_{n-1}) \bigcap U_{i_n},$$

where U_{i_n} is that set in $\mathcal{U}$ which contains the point $S_\tau^n(x)$ (see Fig. 2).

Since $x \in \Omega$, the set $\Omega_n(x)$ is nonempty and belongs to a single set from $\mathcal{U}$ for any n. Writing

$$\Theta_n(x) = S_\tau^{-n}(\Omega_n(x)).$$

it is clear that

$$\Theta_k(x) \subseteq \Theta_l(x) \quad \text{for} \quad k \geq l \tag{12}$$

and that

$$S_\tau^k(\Theta_n(x)) \subseteq \Omega_k(x), \tag{13}$$

Hence the interior of each set $S_\tau^k(\Theta_n(x))$, $k = 0, 1, \ldots, n$, will not intersect with any of the hyperplanes L_i, $i = 1, 2, \ldots, M$.

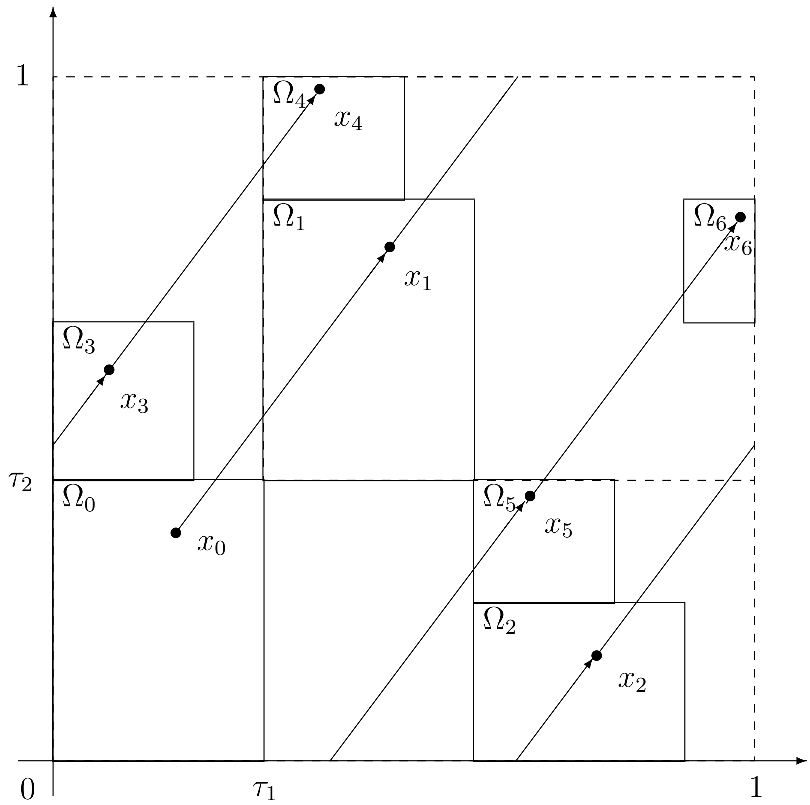

Figure 2: Sets $\{\Omega_i\}$ for multi-dimensional shift mapping

Lemma 4 *For $z \in I^M$*

$$\sigma_0(z)\sigma_1(z)\dots\sigma_k(z) \in \mathcal{D}_m(x) \tag{14}$$

if and only if $k \geq m$ and

$$z \in \Theta_k(x). \tag{15}$$

Proof. Suppose that (15) holds. Since $S_\tau^i(z) \in S_\tau^i(\Theta_k(x)) \subseteq \Omega_i(x)$ (see (13)) and $S_\tau^i(x) \in \Omega_i(x)$, then

$$\sigma_i(z) = \sigma_i(x), \qquad i = 0, 1, \dots, k,$$

and inclusion (14) follows.

Now, suppose that inclusion (14) is valid. Then, by definition of sets $\{\Omega_i(x)\}$, the inclusion $S_\tau^k(z) \in \Omega_k(x)$ holds. Hence $z \in S_\tau^{-k}(\Omega_k(x)) = \Theta_k(x)$, which is inclusion (15). ■

Lemma 5 *If $j \in \mathcal{C}_m(y)$ then*

$$S_\tau^k(y) \in \left\{\bigcup_{i=0}^{m-1} S_\tau^i(\Theta_m(x))\right\} \bigcup \left\{\bigcup_{i=m}^{\infty} \Omega_i(x)\right\}. \tag{16}$$

Proof. If $j \in \mathcal{C}_m(y)$ then by definition of the set $\mathcal{C}_m$ there exist integers k and n with $k \leq j \leq n$ and $n \geq k+m$, such that

$$\sigma_k(y)\dots\sigma_j(y)\dots\sigma_n(y) \in \mathcal{D}_m.$$

In addition for $z = S_\tau^k(y)$ the equalities

$$\sigma_{k+i}(y) = \sigma_i(z), \qquad i = 0, 1, \dots, n-k.$$

are valid. Then, by virtue of Lemma 4, $z \in \Theta_{n-k}(x)$. Since $n-k \geq m$, from this inclusion and (12) follows the inclusion $z \in \Theta_{n-k}(x)$. Therefore

$$S_\tau^j(y) = S_\tau^{j-k}(S_\tau^k(y)) = S_\tau^{j-k}(z) \in S_\tau^{j-k}(\Theta_m(x)).$$

If $0 \leq j-k \leq m$, then

$$S_\tau^j(y) \in \bigcup_{i=0}^{m} S_\tau(\Theta_m(x)) \tag{17}$$

and if $j-k>m$, then $S_\tau^{j-k}(\Theta_m(x)) \subseteq \Omega_{j-k}(x)$ in view of (13) and hence

$$S_\tau^j(y) \in \bigcup_{j=m}^{\infty} \Omega_j(x), \tag{18}$$

(16) then follows from (17) and (18). ∎

Let us now make a crucial observation. As is well known [5] the mapping $S_\tau(\cdot)$ is ergodic for any $\tau = \{\tau_1, \tau_2, \dots, \tau_M\}$ with irrational $\tau_1, \tau_2, \dots, \tau_M$. Hence for almost any $y \in I^M$, the value $\Delta_m(y)$, which by Lemma 5 is the mean absorption time of iterations $S_\tau^i(y)$, $i = 0, 1, 2, \dots$, into the set

$$\left\{ \bigcup_{i=0}^{m-1} S_\tau^i(\Theta_m(x)) \right\} \bigcup \left\{ \bigcup_{i=m}^{\infty} \Omega_i(x) \right\},$$

coincides with the Lebesgue measure of this set, that is,

$$\Delta_m(y) = \sum_{i=0}^{m-1} \text{mes } \Theta_m(x) + \sum_{i=m}^{\infty} \text{mes } \Omega_i(x).$$

Now the mapping S_τ is measure preserving, so

$$\sum_{i=0}^{m-1} \text{mes } \Theta_m(x) = m \,\text{mes } \Theta_m(x) = m \,\text{mes } \Omega_m(x)$$

and hence

$$\Delta_m(y) = m \,\text{mes } \Omega_m(x) + \sum_{i=m}^{\infty} \text{mes } \Omega_i(x). \tag{19}$$

Now we are able to pose the main problem in the proof of Theorem 5:

Show that if $M \geq 3$ then

$$\Delta_m(y) = m \,\text{mes}\, \Omega_m(x) + \sum_{i=m}^{\infty} \text{mes}\, \Omega_i(x) \to 0 \quad \text{as} \quad m \to \infty. \tag{20}$$

3.2 The one-dimensional case revisted

To solve the main problem stated above we shall consider only the case where $0 < x_i < \tau_i$, $i = 1, 2, \ldots, M$, for which we take

$$\Omega = [0, \tau_1) \times [0, \tau_2) \times \ldots \times [0, \tau_M).$$

This set Ω is the maximal set containing x and contained in a single subset from $\mathcal{U}$. It is obvious that for any i the set $\Omega_i(x)$ is parallelepiped, i.e.

$$\Omega_i(x) = [a_{i1}, b_{i1}) \times [a_{i2}, b_{i2}) \times \ldots \times [a_{iM}, b_{iM}).$$

Let us determine upper bounds for the lengths of sides of parallelepiped $\Omega_i(x)$. Clearly, it suffices to do this just for the first side $\omega_{i1} = [a_{i1}, b_{i1})$.

Consider one-dimensional shift mapping $S_\tau(x)$, let $\omega_0 = \omega = [0, \tau)$, and define

$$\omega_n = \begin{cases} S_\tau(\omega_{n-1}) \bigcap [0, \tau), & \text{if} \quad S_{n\tau}(x) \in [0, \tau), \\ S_\tau(\omega_{n-1}) \bigcap [\tau, 1), & \text{if} \quad S_{n\tau}(x) \in [\tau, 1). \end{cases}$$

Then for each n the set ω_n is an interval. If we write $\theta_n = S_\tau^{-n}(\omega_n)$, then $\omega_n = S_\tau^n(\theta_n)$. Let $n_0 = 0$ and successively choose the integer n_i as the smallest integer $n > n_{i-1}$ satisfying the condition $\theta_n \neq \theta_{n_i - 1}$. Then

$$\theta_0 \supset \theta_1 \supset \ldots \supset \theta_i \supset \theta_{i+1} \ldots .$$

Lemma 6 *For any $n_i \leq n < n_{i+1}$ the equalities $\theta_n = \theta_{n_i}$ are valid, one of the endpoints of the interval ω_{n_i} is either 0 or τ and neither of these points belongs to the interior of the intervals $S_\tau^n(\theta_i)$ for $n = 0, 1, \ldots, n_{i+1} - 1$.*

Let $\left\{\frac{p_n}{q_n}\right\}$ denotes the convergent sequence of the simple continued fraction (see, e.g., [11]) of the number τ defined by the condition $p_0 = 0, q_0 = 1$.

Lemma 7 *For almost all τ and for any $\epsilon > 0$ there is an integer $K = K(\tau, \epsilon)$ such that*

$$q_{n+1} < q_n^{1+\epsilon} \quad \text{for} \quad n > K. \tag{21}$$

Proof. According to Theorem 4 on page 164 of [5], for almost all τ there exists $c = c(\tau) > 1$ such that

$$q_n^{\frac{1}{n}} \to c \quad \text{for} \quad n \to \infty.$$

Hence

$$\frac{q_{n+1}^{\frac{1}{n+1}}}{(q_n^{1+\epsilon})^{\frac{1}{n}}} \to c^{-\epsilon} \quad \text{for} \quad n \to \infty,$$

so

$$\frac{q_{n+1}}{q_n^{1+\epsilon}(q_n^{1+\epsilon})^{\frac{1}{n}}} - c^{-\epsilon(n+1)} \to 0 \quad \text{for} \quad n \to \infty$$

and

$$\frac{q_{n+1}}{q_n^{1+\epsilon}} \to 0 \quad \text{for} \quad n \to \infty$$

hold. The required inequality (21) is thus valid for all sufficiently large values of n. ∎

Lemma 8 *Let $\xi = [z, z+\eta) \subseteq [0,1)$ and let N be such that $S_\tau^n(\xi) \bigcap \{0\} \neq \emptyset$ for $0 \leq n < N$. Then*

$$\eta < \frac{2}{N^{\frac{1}{2+\epsilon}}} \tag{22}$$

for almost all τ and for any $\epsilon > 0$.

Proof. Define $\xi_0 = \xi$ and $\xi_i = S_\tau^i(\xi)$ for $i = 1, 2, \dots$. There is an alternative: either all of intervals ξ_i are pairwise non-intersecting or there is a such minimal k for which $\xi_k \bigcap \xi_0 \neq \emptyset$.

In the first case the total length of the intervals ξ_i, $i = 0, 1, \dots, N-1$ does not exceed 1. Since the shift mapping S_τ is measure-preserving, than the lengths of all intervals ξ_i are then identical and equal to η. Therefore

$$\eta \leq \frac{1}{N}$$

and the required estimate (22) holds for any $\epsilon > 0$.

In the second case a more detailed analysis is required. Introduce the intervals $\zeta_i = \xi_i - \{z\}$, $i = 1, 2, \dots$. From the identity

$$S_\tau(x+z) \equiv S_\tau(x) + z \pmod 1 \tag{23}$$

it then follows that

$$\zeta_i = S_\tau^i(\zeta_0) \quad i = 1, 2, \dots ,$$

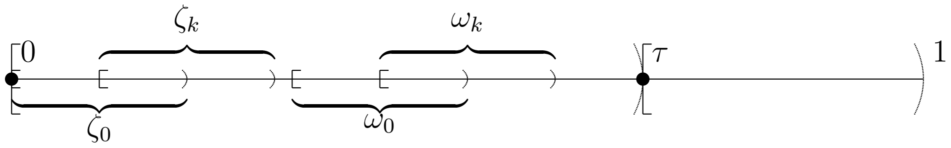

Figure 3: Relation between sets ω_i and ζ_i

with

$$\zeta_k \bigcap \zeta_0 \neq \emptyset, \qquad \zeta_i \bigcap \zeta_0 = \emptyset, \quad \text{for} \quad i = 1, 2, \ldots, k-1. \tag{24}$$

In view of (24)

$$|S_\tau^k(0)| < \eta \quad \text{or} \quad |S_\tau^k(0) - 1| < \eta. \tag{25}$$

(Fig. 3 corresponds to the first case). According to property of the best approximation for convergent sequence of continued fractions (see, e.g. [11]) the integer k coincides with one of numbers $\{q_n\}$, say $k = q_m$. In both cases (25) $|\tau q_m - p_m| < \eta$ and hence

$$\left| \tau - \frac{p_m}{q_m} \right| < \frac{\eta}{q_m}. \tag{26}$$

At the same time (see, e.g. [11])

$$\frac{1}{2 q_m q_{m+1}} < \left| \tau - \frac{p_m}{q_m} \right|. \tag{27}$$

On the other hand for $k < q_m$ there are no points of the form $S_\tau^k(0)$ in the intervals $[0, \eta)$ and $[1 - \eta, 1)$. Since $S_\tau^{q_{m-1}}(0) = \tau q_{m-1} + p_{m-1} \pmod 1$, then $|\tau q_{m-1} + p_{m-1}| > \eta$ and therefore

$$\frac{\eta}{q_{m-1}} < \left| \tau - \frac{p_{m-1}}{q_{m-1}} \right| < \frac{1}{q_{m-1} q_m}. \tag{28}$$

Combining (26), (27) and (28) we obtain

$$\frac{1}{2 q_{m+1}} < \eta < \frac{1}{q_m}, \qquad k = q_m. \tag{29}$$

Now from the definition of the intervals $\{\zeta_n\}$ and from identity (23) it follows that lower endpoints of the intervals ω_0 and ω_k differ by $|\tau q_m - p_m| > \frac{1}{2q_m q_{m+1}}$. Hence, applying the mapping S_τ $2q_m q_{m+1}$ times to the interval ω_0, we can cover the whole interval $[0, 1)$ and in particular the point 0. Therefore

$$N \leq 2q_m q_{m+1} < 2q_{m+1}^2.$$

Now from Lemma 7 it follows $q_{m+1} < q_m^{1+\epsilon}$ for m sufficiently large, so

$$N < 2q_m^{2+\epsilon}.$$

Applying the right inequality (29) we obtain

$$N < \frac{2}{\eta^{2+\epsilon}}$$

and hence

$$\eta < \frac{2^{\frac{1}{2+\epsilon}}}{N^{\frac{1}{2+\epsilon}}} < \frac{2}{N^{\frac{1}{2+\epsilon}}}.$$

which completes the proof of Lemma 8. ∎

3.3 Proof of Theorem 5

As was shown in Section 3.1, from Lemma 5 it follows that in order to prove Theorem 5 we need only establish the relation (20). But from Lemma 8 and the definition of sets $\Omega_i(y)$ for almost all τ the following estimate is valid:

$$\text{mes } \Omega_i(y) \leq \frac{2^M}{i^{\frac{M}{2+\epsilon}}}.$$

Hence from (19)

$$\Delta_m(y) \leq \frac{m2^M}{m^{\frac{M}{2+\epsilon}}} + \sum_{i=m}^{\infty} \frac{2^M}{i^{\frac{M}{2+\epsilon}}}$$

or, what is the same,

$$\Delta_m(y) \leq 2^M m^{1-\frac{M}{2+\epsilon}} + 2^M \sum_{i=m}^{\infty} i^{-\frac{M}{2+\epsilon}}.$$

Note, that the value of ϵ can be chosen arbitrarily small. Hence, the right hand part of the latter inequality clearly tends to 0 as $m \to \infty$ when $M \geq 3$. This completes the proof of Theorem 5. ∎

References

[1] E.Asarin, P.Diamond, I.Fomenko, A.Pokrovskii. Chaotic phenomena in desynchronized systems and stability analysis. Computers Math. Applic., Vol. 25, No. 1, pp. 81-87, 1993.

[2] E.A.Asarin, V.S.Kozyakin, M.A.Krasnoselskii, N.A.Kuznetsov. Stability analysis of desynchronized discrete event systems, Moscow, Nauka, 1992 (in Russian).

[3] M.Barnsley, *Fractals Everywhere*, Academic Press, Boston, 1988.

[4] A.F.Kleptsyn, V.S.Kozyakin, M.A.Krasnoselskii, N.A.Kuznetsov. Desynchronization of linear systems, Mathematics and Computers in Simulation, Vol. 26, 1984 pp. 423-431.

[5] I.P. Cornfeld, S.V. Fomin and Ya.G. Sinai, *Ergodic Theory*, Springer–Verlag, New York, 1982.

[6] V.S.Kozyakin. On stability of phase and frequency desynchronized systems under perturbations of component updating moments, Automatika i Telemekhanika, No. 8, 1990, pp. 35-42 (in Russian).

[7] V.S.Kozyakin. On stability analysis of desynchronized systems by symbolic dynamic methods, Dokl. Akad. Nauk SSSR, Vol. 311, No. 3, 1990, pp. 549-552 (in Russian).

[8] F.Mignosi. On the number of factors of sturmian words. Theoretical Computer Science 82, 1991, 71–84.

[9] N.Martin and J.England. Mathematical theory of entropy. Encyclopedia of Mathematics and its Applications, v. 13. London, Addison–Wesley Publishing Company, 1981.

[10] M.Morse and G.A.Hedlund. Symbolic dynamics II: Sturmian trajectories, American J. Math. 62, 1940, 1–42

[11] C.T. Long. Elementary Introduction to Number Theory, 2nd edition, Lexington, Massachussetts, Toronto, London, D.C.Heath and Company, 1972.

[12] M.Lothaire. Combinatorics of Words. Encyclopedia of Mathematics and its Applications, v. 17. London, Addison–Wesley Publishing Company, 1983, 238p.