Записки научных
семинаров ПОМИ
Том 240, 1997 г.

A. N. Livshits

SOME APPLICATIONS OF COMPUTER TO INVESTIGATION OF SUBSTITUTIONAL DYNAMICAL SYSTEMS

In [1] and in [2] the question about the topological mixing of substitutional dynamical systems was considered. For some classes of substitutions with growing excess the proof or the new proof of topological mixing was obtained. In [1] for some substitution with property, considered by Dumont-Thomas, computer was applied to proof of the growth of excess. In [2,3] it was proved without the computer for the general substitution with such property. Still for the topological mixing some additive conditions of combinatorial nature were argued. The example, which does not satisfy them, was considered and the other method was used to prove that the topological mixing still takes place. Here we formulate more weak sufficient condition, which is fulfilled, but without computer it would be difficult to convince oneself in it.

For the investigation of dynamical systems such notion as subword complexity is of importance. If represent it as a sum of four parts, defined by beginning and end of subword (studied in [4]), the natural recurrence for these quadruples takes place. The Qbasic-program was elaborated, giving the Tex-file of such recurrent formulas after entering the words of substitution on two symbols, and the program, writing the program, corresponding to these recurrent formulas. These formulas and the computer calculations by them did help author to see, that for many (possibly all) Dumont-Thomas substitutions, not satisfying considered condition, the topological mixing takes place (even growth of excess for one part of sum and other interesting properties). The method of proof is analogous to one of [1].

§1. Substitution $1 \to 121$, $2 \to 22212$ and new combinatorial condition.

Let $\omega_{\mathcal{A}}$ ($\mathcal{A} = \{A_1, A_2\}$) be the primitive substitution over the alphabeth $Z_* = \{1,2\}$. To it the homeomorphism $T_{\mathcal{A}} : X_{\mathcal{A}} \to X_{\mathcal{A}}$ [5] corresponds, which is called the substitutional dynamical system. General

This work is supported by the grant INTAS-93-0570.

homeomorphism $T : X \to X$ is called topological mixing, if for any open sets X, Y there exists such $N > 0$ that for any $n > N$ $T^n U \bigcap V \neq 0$ Our substitutional homeomorphism is strictly ergodic. Let μ be its invariant measure and $\mu_1 = \mu(\widetilde{1})$, where $\widetilde{1}$ is a cylinder. Formulate our sufficient condition of topological mixing.

Fact. *Let the primitive substitution $\omega_{\mathcal{A}}$ have 3 following properties.*

Property 1. *For all $n > 0$ $|\omega_{\mathcal{A}}^n(1)|$ and $|\omega_{\mathcal{A}}^n(2)|$ are relatively simple.*

Property 2. *Let $v(n)$ stand for $\max\{|A|_1 - \mu_1 n, |A| = n$ and A is admissible$\}$, $u(n)$ stand for $\max\{-|A|_1 + \mu_1 n, |A| = n$ and A is admissible$\}$. Then $v(n) \to \infty$, $u(n) \to \infty$, when $n \to \infty$.*

Property 3. *There exists such k, that $\omega_{\mathcal{A}}^k(1) = A_1 b_1 C_1 = A_2 b_2 C_2$; $\omega_{\mathcal{A}}^k(2) = D_1 e_1 F_1 = D_2 e_2 F_2$, where either $|C_1| + |C_2| = |F_1| + |F_2|$, , $|C_1|_1 + |C_2|_1 = |F_1|_1 + |F_2|_1$, b_i, e_i are the letters from Z_* (and numbers) and $b_1 + b_2 = e_1 + e_2$ or $|C_1| = |F_1|, |C_1|_1 = |F_1|_1$ and $b_1 = e_1$.*

Then it topologically mixes.

The property 3 from [2] is the particular case of this property 3 (if, for example, $e_1 = b_2, e_2 = b_1, |C_2| = |F_2| = \Lambda$).

The proof is mainly analogous to that of [2]. The consideration of second alternative was done in [2]. If the first takes place, we can, without loss of generality, suppose, that $b_1 = e_1, b_2 = e_2$. For any word $1A2$ we have $\omega_{\mathcal{A}}^k(1A2) = A_2 b_2 C_2 \omega_{\mathcal{A}}^k(A) D_2 e_2 F_2$ for any word $2A'1$ we have $\omega_{\mathcal{A}}^k(2A'1) = D_1 e_1 F_1 \omega_{\mathcal{A}}^k(A') A_1 b_1 C_1$. Consider the subword $B = C_2 \omega_{\mathcal{A}}^k(A) D_2 e_2$ of the first word and the subword $B' = F_1 \omega_{\mathcal{A}}^k(A') A_1 b_1$ of second. As $|F_1| - |C_1| = |C_2| - |F_2|$,$|F_1|_1 - |C_1|_1 = |C_2|_1 - |F_2|_1$ we have $|B| = |B'|, |B|_1 = |B'|_1$ if $|A2| = |A'1|, |A2|_1 = |A'1|_1$. Suppose, that for some k' $\omega_{\mathcal{A}}^{k'}(b_2) = U d' V$; $\omega_{\mathcal{A}}^{k'}(b_1) = U' d'' V'; d'' = d' \in Z_*$. Consider the admissible subword $B'^* = d'' V' \omega_{\mathcal{A}}^{k'}(B')(d''V')^{-1}$ of word $d''V'\omega_{\mathcal{A}}^{k'}(B')$ and the admissible subword $B^* = d' V \omega_{\mathcal{A}}^{k'}(B)(d'V)^{-1}$ of word $d'V\omega_{\mathcal{A}}^{k'}(B)$. We have $B'^* d''$, $B^* d'$ are admissible. There exist constants m, n, p, q, t, u such, that $|B'^*| = m|A'| + n|A'|_1 + t; |B'^*|_1 = p|A'| + q|A'|_1 + u$ and, as well $|B^*| = m|A| + n|A|_1 + t; |B^*|_1 = p|A| + q|A|_1 + u$, what follows from property 3. But, it is easy to see, like in [2], that for any admissible G there exists admissible Xw, such that vXw is admissible $w \neq v \in Z_*$ and $|Xw| = |G|, |Xw|_1 = |G|_1$ (it is proved by "moving" the word in admissible sequence). So, we did prove the following: there exist such constants k, k_1, c, d and letter $a \in Z_*$ that given admissible word G we can choose the admissible word H: $|\omega_{\mathcal{A}}^k G| + c = |H|, |\omega_{\mathcal{A}}^k G|_1 + d = |H|_1$ and the word $a\omega_{\mathcal{A}}^{k_1}(H)a$ is admissible. The rest (using the properties 1) and 2)) is analogous to [1] and [2].

In [2] we did prove, that for substitution $\omega_{\mathcal{A}}, \mathcal{A} = (121, 22212)$ the property 3) from [2] does not take place. But it is not the case with property 3) of this work. As calculation did show, such $k, A_i, b_i, C_i, D_i, e_i, F_i$ exist. For example, $\omega_{\mathcal{A}}^4(1) = A_1 2C_1 = A_2 2C_2, \omega_{\mathcal{A}}^4(2) = D_1 2F_1 = D_2 2F_2$, where $|C_1| = 128, |C_2| = 106, |F_1| = 190, |F_2| = 44$.

That first property takes place, is established for general case in [2]. Of course, in the formulation one can change ends and beginnings of words.

§2. Using the recurrence.

Now the other substitution $1 \to 1211$, $2 \to 22212$ with growing excess will be considered, for which this property 3 cannot take place (as calculation did show, too). So, we shall use other method for proving topological mixing. This strategy, possibly, works for the general case of substitutions on two letters with the property of Dumont–Thomas. Let us at first obtain some recurrent relation for general case.

Let $G_{\mathcal{A}} = \begin{pmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{pmatrix}$ be the matrix of substitution $\omega_{\mathcal{A}}$. It is evident, that if for word $A \in Z_*, |A| > \max(|\omega_{\mathcal{A}}(1)|, |\omega_{\mathcal{A}}(2)|)$, then it is admissible iff there exist letters u, v and admissible words $A_1, A_2, A_3, B_1, C_1, D_1, E_1, F_1$, such that $A = A_1F_1B_1; C_1A_1F_1B_1D_1 = \omega_{\mathcal{A}}(A_2), C_1A_1 = \omega_{\mathcal{A}}(u), B_1D_1 = \omega_{\mathcal{A}}(v), A_2 = uA_3v$, the words C_1, D_1 can be empty. If for given substitution such representation is unique, we have the following recurrence: if $N(i,j,m,k)$ is the number of admissible words of sort $iAj, |iAj|_1 = m, |iAj|_2 = k$, then for m, k large enough we have

$$N(i,j,m,k) = \sum_{\substack{i_1 \in Z_* \\ i_2 \in Z_*}} \sum_{\substack{A,B,C,D \\ AiB = \omega_{\mathcal{A}}(i_1) \\ CjD = \omega_{\mathcal{A}}(i_2)}} N(i_1, i_2, m(A,B,C,D), k(A,B,C,D)), \tag{1}$$

where $G_{\mathcal{A}} * (m(A,B,C,D),\ k(A,B,C,D)) = (m + |A|_1 + |D|_1,\ k + |A|_2 + |D|_2)$ as vectors.

Here is the fragment of the corresponding TeX-file, made by program. The formulas for all possible values of $(m * t_{22} - k * t_{21}) \mod (t_{11} * t_{22} - t_{12} * t_{21})$ are listed separately.

$y(1) = 1211, y(2) = 22212$
$N(1,1,m,k) =$
$+N(1,1,\frac{m*4-k*1+10}{11},\frac{-m*1+k*3+3}{11}) + N(1,2,\frac{m*4-k*1-1}{11},\frac{-m*1+k*3+3}{11})$

$N(1,1,m,k) =$
$N(1,1,m,k) =$
$+N(2,1,\frac{m*4-k*1-3}{11},\frac{-m*1+k*3+9}{11})$
$N(1,1,m,k) =$
$+N(1,1,\frac{m*4-k*1+7}{11},\frac{-m*1+k*3+1}{11}) + N(1,1,\frac{m*4-k*1+7}{11},\frac{-m*1+k*3+1}{11})$
$+N(1,1,\frac{m*4-k*1+7}{11},\frac{-m*1+k*3+1}{11}) + N(2,2,\frac{m*4-k*1-4}{11},\frac{-m*1+k*3+12}{11})$
$N(1,1,m,k) =$
$+N(1,2,\frac{m*4-k*1+6}{11},\frac{-m*1+k*3+4}{11})$
$N(1,1,m,k) =$
$N(1,1,m,k) =$
$+N(1,1,\frac{m*4-k*1+4}{11},\frac{-m*1+k*3-1}{11}) + N(2,1,\frac{m*4-k*1+4}{11},\frac{-m*1+k*3+10}{11})$
$N(1,1,m,k) =$
$+N(1,1,\frac{m*4-k*1+3}{11},\frac{-m*1+k*3+2}{11}) + N(1,1,\frac{m*4-k*1+14}{11},\frac{-m*1+k*3+2}{11})$
$N(1,1,m,k) =$
$+N(1,2,\frac{m*4-k*1+2}{11},\frac{-m*1+k*3+5}{11})$
$N(1,1,m,k) =$
$+N(2,1,\frac{m*4-k*1+1}{11},\frac{-m*1+k*3+8}{11})$
$N(1,1,m,k) =$
$+N(1,1,\frac{m*4-k*1+0}{11},\frac{-m*1+k*3+0}{11}) + N(1,1,\frac{m*4-k*1+11}{11},\frac{-m*1+k*3+0}{11})$

$N(1,2,m,k) =$
$+N(1,2,\frac{m*4-k*1+10}{11},\frac{-m*1+k*3+3}{11}) + N(2,2,\frac{m*4-k*1-1}{11},\frac{-m*1+k*3+14}{11})$
$N(1,2,m,k) =$
$+N(1,2,\frac{m*4-k*1+9}{11},\frac{-m*1+k*3+6}{11}) + N(2,2,\frac{m*4-k*1-2}{11},\frac{-m*1+k*3+17}{11})$
$N(1,2,m,k) =$
$+N(1,1,\frac{m*4-k*1+8}{11},\frac{-m*1+k*3-2}{11}) + N(1,2,\frac{m*4-k*1+8}{11},\frac{-m*1+k*3+9}{11})$
$+N(2,2,\frac{m*4-k*1-3}{11},\frac{-m*1+k*3+9}{11})$
$N(1,2,m,k) =$
$+N(1,2,\frac{m*4-k*1+7}{11},\frac{-m*1+k*3+1}{11})$
$N(1,2,m,k) =$
$+N(1,2,\frac{m*4-k*1+6}{11},\frac{-m*1+k*3+4}{11})$
$N(1,2,m,k) =$
$+N(1,2,\frac{m*4-k*1+5}{11},\frac{-m*1+k*3+7}{11}) + N(2,1,\frac{m*4-k*1+5}{11},\frac{-m*1+k*3+7}{11})$
$N(1,2,m,k) =$
$+N(1,1,\frac{m*4-k*1+15}{11},\frac{-m*1+k*3-1}{11}) + N(1,2,\frac{m*4-k*1+4}{11},\frac{-m*1+k*3+10}{11})$
$N(1,2,m,k) =$
$+N(1,2,\frac{m*4-k*1+3}{11},\frac{-m*1+k*3+2}{11}) + N(1,2,\frac{m*4-k*1+3}{11},\frac{-m*1+k*3+2}{11})$
$N(1,2,m,k) =$
$+N(1,2,\frac{m*4-k*1+2}{11},\frac{-m*1+k*3+5}{11})$

$N(1,2,m,k) =$
$+N(1,2,\frac{m*4-k*1+1}{11},\frac{-m*1+k*3+8}{11})$
$N(1,2,m,k) =$
$+N(1,1,\frac{m*4-k*1+11}{11},\frac{-m*1+k*3+0}{11}) + N(1,2,\frac{m*4-k*1+0}{11},\frac{-m*1+k*3+0}{11})$
$+N(2,2,\frac{m*4-k*1+0}{11},\frac{-m*1+k*3+11}{11})$

$N(2,1,m,k) =$
$+N(2,1,\frac{m*4-k*1-1}{11},\frac{-m*1+k*3+3}{11}) + N(2,2,\frac{m*4-k*1-1}{11},\frac{-m*1+k*3+3}{11})$
$N(2,1,m,k) =$
$+N(2,1,\frac{m*4-k*1-2}{11},\frac{-m*1+k*3+6}{11}) + N(2,2,\frac{m*4-k*1-2}{11},\frac{-m*1+k*3+6}{11})$

If there are any ambiguities, we may take them into account in any way or use the formalism of structurized sequences, which is described in [2,6] and in some other works of author. Here we can use some variant of it. Considering for simplicity only such substitutions, that $\forall i \in Z_* \; \omega_{\mathcal{A}}(i) = i...i$, we shall mean under the word of length $r > 2$ either the word of first type, or the word of second type: word of first type is defined as the collection $(i \in Z_*, j \in Z_*, m, l, k), k > 0, 0 < m \leqslant |\omega^k_{\mathcal{A}}(i)|, ij-$admissible word$, 0 < l \leqslant |\omega^k_{\mathcal{A}}(j)|, r = m + l, m > |\omega^{k-1}_{\mathcal{A}}(i)|$ or $l > |\omega^{k-1}_{\mathcal{A}}(j)|$, word of second type is the collection $(i \in Z_*, A, B, m, l, k)$, where A, B-words, $k > 0$ is such that either $|\omega^k_{\mathcal{A}}(1)| \geqslant r > |\omega^{k-1}_{\mathcal{A}}(1)|$ or $|\omega^k_{\mathcal{A}}(2)| \geqslant r > |\omega^{k-1}_{\mathcal{A}}(2)|$ and $\omega_{\mathcal{A}}(i) = AuBvC, B \neq \Lambda$, $0 < m \leqslant |\omega^{k-1}_{\mathcal{A}}(u)|, 0 < l \leqslant |\omega^{k-1}_{\mathcal{A}}(v)|; r = m + l + |\omega^{k-1}_{\mathcal{A}}(B)|$. Of course to the word of any type the concrete admissible word naturally corresponds, so we can assign to our "words" the $|\quad|_1$ and $|\quad|_2$, first and last letter.

Introduce instead of $N(i,j,m,k)$ the other number N^{st} – number of collections of both types, i,j,m,k being of the same sence, where the words of first type (i_1, j_1, m, l, k) are counted with the coefficient

$$c(i_1, j_1) = \sum_{i_2} \sum_{\substack{A,C \\ Ai_1j_1C= \\ \omega_{\mathcal{A}}(i_2)}} 1 + \sum_{\substack{\{i_2,j_2\} \\ i_2j_2-\text{admissible} \\ \omega_{\mathcal{A}}(i_2)=\ldots i_1 \\ \omega_{\mathcal{A}}(j_2)=j_1\ldots}} 1,$$

then one can prove, that $N^{st}(i,j,m,k)$ satisfies this recurrence. In every case the results of computer calculations agree with it (for many substitutions) .

We are going to prove for our substitution, using this recurrence, that the diapasons of excesses of appropriate (though other then previous) sets of words with fixed length, are growing too with growth of this length. We study some subsets of sets of admissible words of sort $1A2$, so, if on this set the sufficient diapasons can be realized, we don't need the property 3. What tells us about such words our recurrence? We have

$$N(1,2,m,k) > \sum_{\substack{A,B,C,D \\ A1B=\omega_{\mathcal{A}}(1) \\ C2D=\omega_{\mathcal{A}}(2)}} N(1,2,m(A,B,C,D),k(A,B,C,D)) \quad (2)$$

Given m, k we apriory are not sure in that the A, D from the list of formula (1) exist, such that $\det(G_{\mathcal{A}}) | l.c.d\,(t_{22} * (m + |A|_1 + |D|_1) - t_{21} * (k + |A|_2 + |D|_2), -t_{12} * (m + |A|_1 + |D|_1) + t_{11} * (k + |A|_2 + |D|_2))$, what is necessary for us. Still in fact for $N(1, 2, m, k)$ it is so even for the members of (2) (and it was computer to prompt it), as is seen from our fragment of TeX-file. Of course, in general case, properties $\det(G_{\mathcal{A}}) \neq 0$, $(t_{11}, t_{12}) = 1$ and $(t_{21}, t_{22}) = 1$ imply that contemporarily $\det(G_{\mathcal{A}}) | (t_{22} * (m+|A|_1+|D|_1)-t_{21}*(k+|A|_2+|D|_2)) = dd$ or not and $\det(G_{\mathcal{A}}) | (-t_{12} * (m + |A|_1 + |D|_1) + t_{11} * (k + |A|_2 + |D|_2)) = uu$ or not. Because, for any integer a, b we have, noting $a * t_{11} + b * t_{12}$ as e and $a * t_{21} + b * t_{22}$ as f, $\det(G_{\mathcal{A}}) | (e * dd + f * uu)$. It did allow us to find for every n the large interval $I_n = [u(n), v(n)]$, such that $k \in I_n$ implies $N(1, 2, n-k, k) > 0$. We define them in the following way. For $n < 18$ we calculate the sets of all k, such that $N(1, 2, n-k, k) > 0$ (they turned to be the intervals). For $n > 17$ we get the intervals, using the lists from recurrence (2). The quickness of calculations was founded on inequality $0 \leqslant u(n+1)-u(n) \leqslant 1, 0 \leqslant v(n+1)-v(n) \leqslant 1$, proved by induction, using special tables, made by computer.

Further strategy is analogous to that of [1]. Though in [1] we consider all the admissible words, and not special subset, the method (described below) can be applied. Denote as ω the number $t_{11} + t_{12} - t_{21} - t_{22}$ (we suppose that it is > 0, here it is 1). Let $\theta > \theta_1 > 1$ be the eigenvalues of $G_{\mathcal{A}}$. For any natural N we denote $d_1^N = \frac{\mu(\tilde{2})*N-u(N))}{N^\beta}$, (where $\beta = \frac{\log_2 \theta_1}{\log_2 \theta}$), $d_2^N = \frac{v(N)-\mu(\tilde{2}*N)}{N^\beta}$. For any natural $N_1 < N_2$ we denote $c_1(N_1, N_2) = \min\limits_{N_1 \leqslant n \leqslant N_2} d_1^n, c_2(N_1, N_2) = \min\limits_{N_1 \leqslant n \leqslant N_2} d_2^n$ We choose some large natural N_1, N_2, such that $N_2 > \theta * N_1 + \omega * N_1^\beta c_2(N_1, N_2)$, further we denote as N^* the number $[\theta * N_2 - \omega c_1(N_1, N_2) N_2^\beta]$ and prove, that using only the words of sort $\omega_{\mathcal{A}}(A), N_1 \leqslant |A| \leqslant N_2$, which are present in the list of formula (2) for our substitution, we can estimate from below $c_1(N_2+1, N^*)$ as $c_1(N_1,N_2) - \lambda N_2^{-\beta}$, $c_2(N_2+1, N^*)$ as $c_2(N_1,N_2)(1-\omega c_2(N_1,N_2)(\frac{N_2}{\theta})^{\beta-1}\frac{1}{\theta})^{\beta} - \lambda N_2^{-\beta}$, where $\lambda = \det(G_{\mathcal{A}}) = t_{11}t_{22} - t_{12}t_{21}$. Further, iterating the transformation $(N_1,N_2) \to (N_1,N^*)$ we evaluate $c_1(N_1,\infty), c_2(N_2,\infty)$. So, if for some initial N_1, N_2 we shall have such c_1, c_2, that $c_1(N_1,\infty) > 0$ and $c_2(N_1,\infty) > 0$, then we obtain the proof of the fact, that excess tends to infinity. We took $N_1 = 535$. Resulting estimates of [1] are the following ones. Denote $\omega c_2(N_1,N_2)(\frac{N_2}{\theta})^{\beta-1}\frac{1}{\theta}$ as α, $\theta - \frac{1}{N_2} - \omega c_1(N_1,N_2)N_1^{\beta-1}$ as ν. Then $c_1(N_1,\infty > c_1(N_1,N_2) - \frac{\lambda}{N_2^{\beta}}\frac{1}{1-\nu^{-\beta}}, c_2(N_1,\infty > c_1(N_1,N_2) - \frac{\lambda}{N_2^{\beta}}\frac{1}{1-\nu^{-\beta}} + c_2(N_1,N_2)\beta\ln(1-\alpha)\frac{1}{1-\nu^{\beta-1}}$. Calculations did show, that $c_1(N_1,\infty) > 0, c_2(N_1,\infty) > 0$.

References

1. A. N.Livshits, *The method of proving of topological mixing for certain class of substitutions.* — Preprint POMI **P-10-91** (1991).
2. A. N.Livshits, *Some properites of substitutional dynamical systems.* — Preprint POMI **13/1994** (1994).
3. A. N.Livshits, *About the limit distributions and asymptotic of extremal values for some sequences of random variables.* — Zap. nauchn. semin. POMI **216:13** (1994), 86–103.
4. F. M. Dekking, M.Keane, *Mixing properties of substitutions.* — ZFW **B.42, H.1** (1978), 375–389.
5. A. N. Livshits, *A sufficient condition for weak mixing of substitutions and stationary adic transformations.* — Mat. Zametki **44:6** (1988), 785–793.
6. A. N. Livshits, *About some properties of decoding for primitive substitutions.* — Zap. nauchn. semin. POMI **223** (1995), 127–136.