Some new counterexamples to context-free substitution in recurrent double sequences over finite sets

Mihai Prunescu[1;2]

(1) Brain Products, Freiburg, Germany; and (2) Simion Stoilow Institute of Mathematics of the Romanian Academy Research unit 5, P. O. Box 1-764, RO-014700 Bucharest, Romania.

Abstract. The recurrent double sequences over finite sets form a Turing complete model of computation. Some recurrent double sequences can be deterministically produced by expansive systems of context-free substitutions. We recall an automatic proof method for statements like [given a recurrent double sequence R and a system of substitutions S, R can be produced by S] and we state a conjecture that has got a lot of evidence by this method. Finally we concentrate over recurrent double sequences that cannot be generated by context-free substitutions. The point is that these particular recurrent double sequences interpret arithmetic progressions.

Key Words: recurrent double sequence, context-free substitution, Turing complete models of computation, Frobenius' automorphism, homomorphisms of finite abelian p-groups. A.M.S.-Classification: 05B45, 28A80, 03D03.

1 Introduction

Definition 1. Let $(A, f, 1)$ be a finite structure with one ternary function f, and one constant 1. The recurrent double sequence $a : \mathbb{N} \times \mathbb{N} \to A$ is a sequence $(a(i,j))$ satisfying the initial conditions $a(i,0) = a(0,j) = 1$ and the recurrence $a(i,j) = f(a(i-1,j), a(i-1,j-1), a(i,j-1))$.

In [7] the author studied the problem to decide if recurrent double sequences are ultimately zero or not, where $0 \in A$ is some other fixed constant. In this article is proved that this problem is undecidable even if restricted to binary functions with the recurrence $a(i,j) = f(a(i-1,j), a(i,j-1))$ which are, moreover, commutative. (With other words, $A \models f(x,y,z) = f(x,1,z) = f(z,1,x)$.) In the next statement we call an instance of the Halting Problem a pair (M, w) where M is a Turing machine and w is an input. In [7] is proven the following:

Theorem 1. *To every instance (M, w) of the Halting Problem one can algorithmically associate a commutative finite algebra $\mathcal{A} = (A, f, 0, 1)$ such that the recurrent double sequence defined by $a(i,0) = a(0,j) = 1$ and $a(i,j) = f(a(i-1,j), a(i,j-1))$ is ultimately zero if and only if for input w: (the machine M stops with cleared tape without having done any step in the negative side of the tape) or (the machine M makes at least one step in the negative side of its tape and the first time when M makes such a step the tape of M is cleared). Consequently, it is undecidable if such (or the more general) recurrent double sequences over finite sets are ultimately zero.*

This result makes clear that recurrent double sequences are Turing complete, so they form a very rich class of objects. This article has the goal to make a step towards a separation of complexity classes inside the set of recurrent double sequences. First of all, is there a set of recurrent double sequences that can be canonically recognized as simpler, easier, less complex?

The starting point of this whole research was an open problem related to a very special form of linear recurrent double sequences over prime fields of finite characteristic, posed by Lakhtakia and Passoja in [6]. The author proved in [8] that if $A = \mathbb{F}_q$ is the finite field with q elements and $f(x,y,z) = x + my + z$, where $m \in \mathbb{F}_q$ is an arbitrary fixed element, $f(x,y,z)$ generates a self-similar pattern $(a(i,j))$. In the case when q is prime and so $\mathbb{F}_q = \mathbb{Z}/q\mathbb{Z}$ as ring of classes of remainders modulo q, the pattern can be also obtained by substitutions of type $x \to xB$, where B is the $q \times q$ matrix occurring as left upper minor in the recurrent double sequence. This fact is not explicitly stated in [8], but is very easy to see it applying the Kronecker product representation theorem from [8] in the case q prime, where the only one automorphism of Frobenius is the identity.

The most classical example of such recurrent double sequence is Pascal's Triangle modulo 2, called also Sierpinski's Gasket. The most easy case of the result proved in [8] is that this recurrent double sequence given by $f(x,y,z) = x + z$ over $\mathbb{F}_2$ can be obtained by substitutions starting with 1 at stage 0 and applying the rules $1 \to \begin{matrix} 1\,1 \\ 1\,0 \end{matrix}$ and $0 \to \begin{matrix} 0\,0 \\ 0\,0 \end{matrix}$, such that one substitutes all elements of stage n in order to achieve the stage $n+1$. The matrix of stage k is a square and has dimension 2^k.

The case analysed above is "regular" in the sense that the substitution rules have the special form element $\to$ matrix. The author has been surprised to observe that a lot of other repetitive phenomena in recurrent double sequences come from a more general kind of context-free substitution, matrix $\to$ matrix. This makes the subject of the next section. *However, we must observe here that the analogy between these recurrent double sequences and the context-free languages as defined by Noam Chomsky in [2] is quite limited, because our generation procedure is* deterministic.

In the second section we present without proof the result on which is based an automatic proof method that a concretely given recurrent double sequence can be also produced by context-free substitutions. In the third section we present an example and a Conjecture concerning this kind of recurrent double sequences. The last section is dedicated to an example of recurrent double sequence for which one can easily prove that it cannot be generated by context-free substitutions.

For other results concerning self-similarity and automata see [5], [3], [10]. To visualize recurrent double sequences we use images obtained by interpreting the values as different colours. The list of colour correspondences will be concretely given here only if is important for understanding some proof.

2 Expansive systems of context-free substitutions

The definitions and the results of this section appeared the first time [9]. In the following definitions x, y, s are positive integers, $y = xs$ and $s \geq 2$. A is a fixed finite set.

Definition 2. Let $\mathcal{X}$ be a finite set of $x \times x$ matrices over A and $\mathcal{Y}$ be a set of $y \times y$ matrices over A such that every $Y \in \mathcal{Y}$ has a $s \times s$ block matrix representation $(X(i,j))_{0 \leq i,j < s}$ and all blocks $X(i,j) \in \mathcal{X}$. We call system of (context-free) substitutions of type $x \to y$ the tuple $(\mathcal{X}, \mathcal{Y}, \Sigma, X_1)$, where $\Sigma : \mathcal{X} \to \mathcal{Y}$ is a fixed function and $X_1 \in \mathcal{X}$ is a fixed element of $\mathcal{X}$, called start-symbol. If a $u \times v$ matrix Z consists only of neighbouring blocks $X(i,j) \in \mathcal{X}$, $Z = (X(i,j))_{0 \leq i < u,\, 0 \leq j < v}$, we define $\Sigma(Z)$ to be the $su \times sv$ matrix with block representation $(\Sigma(X(i,j)))$. We define the sequence of matrices $(S(n))$ by $S(1) = X_1$ and $S(n) = \Sigma^{n-1}(X_1)$. The number s is called scaling factor of the system of substitutions.

Definition 3. We call the system of substitutions $(\mathcal{X}, \mathcal{Y}, \Sigma, X_1)$ expansive if the block representation of the matrix $\Sigma(X_1) = (X(i,j))$ using matrices in $\mathcal{X}$ fulfills the condition $X(0,0) = X_1$. To be more clear: $X(0,0)$ is the $x \times x$ left upper block of $\Sigma(X_1)$.

Lemma 1. *Let $(\mathcal{X}, \mathcal{Y}, \Sigma, X_1)$ be an expansive system of substitutions. Then for all $n > 0$ the matrix $S(n)$ is $xs^{n-1} \times xs^{n-1}$ left upper minor of the matrix $S(n+1)$.*

Definition 4. Let $(A, 1, f)$ be a finite structure with ternary function f. Denote by $R = (a(i,j))$ the recurrent double sequence according to Definition 1. Suppose that two natural numbers $x \geq 1$ and $s \geq 2$ have been fixed. For $n \geq 1$ denote by $R(n)$ the finite matrix $(a(i,j))$ with $0 \leq i,j < xs^{n-1}$.

Definition 5. We say that a $u \times v$ matrix $K = (k(\alpha, \beta))$ occurs in the $w \times z$ matrix $T = (t(a,b))$ in position (i,j) if $0 \leq i < w$, $0 \leq j < z$, $i + u \leq w$, $j + v \leq z$ and for all $0 \leq \alpha < u$ and $0 \leq \beta < v$ one has $t(i + \alpha, j + \beta) = k(\alpha, \beta)$.

Definition 6. Let x be a fixed natural number and T be a $wx \times zx$ matrix over A. We denote $\mathcal{N}_x(T)$ the set of all $2x \times 2x$ matrices occurring in some position (kx, lx) in T.

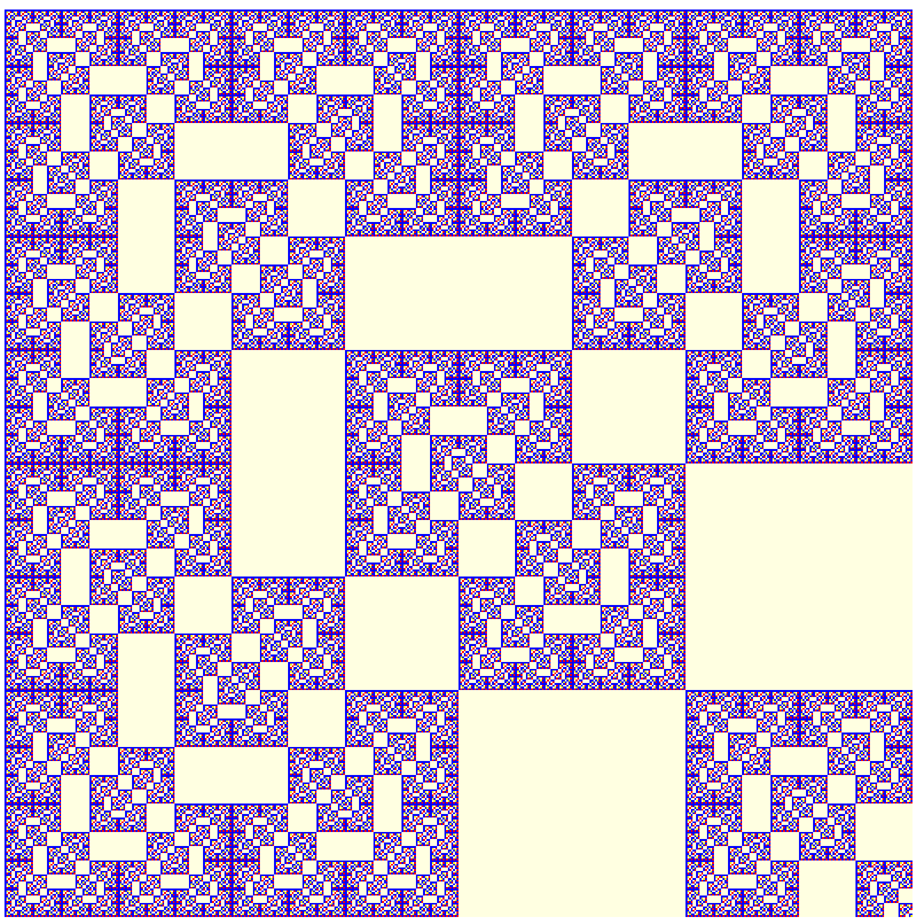
Definition 7. Let x be a fixed natural number and T be a $wx \times zx$ matrix over A. We denote by $\mathcal{J}_x(T)$ the set of all $x \times x$ matrices occurring in T in some position $(0, kx)$. Analogously, we denote by $\mathcal{I}_x(T)$ the set of all $x \times x$ matrices occurring in T in some position $(kx, 0)$.

Theorem 2. *Let $(A, f, 1)$ be a finite structure with ternary function f and let $(\mathcal{X}, \mathcal{Y}, \Sigma, X_1)$ be an expansive system of substitutions of type $x \to y$ over A. We define the matrices $(R(n))_{n\geq 1}$ according to x and $s = y/x$ given by the system of substitutions. Suppose that for some $m > 1$ following conditions hold: (1) $R(m) = S(m)$, (2) $\mathcal{N}_x(R(m-1)) = \mathcal{N}_x(R(m))$, (3) $\mathcal{J}_x(R(m-1)) = \mathcal{J}_x(R(m))$ and $\mathcal{I}_x(R(m-1)) = \mathcal{I}_x(R(m))$. Then for all $n \geq 1$ one has $R(n) = S(n)$.*

This result says essentially that if a recurrent double sequence and a double sequence produced by an expansive system of context-free substitutions are identical in a starting minor, then they are identical everywhere. The theorem can be immediately used to implement an automatic method to prove or disprove that a given recurrent double sequence is produced by an expansive system of substitutions of a given type $x \to sx$.

3 An example and a conjecture

Consider the following rule of recurrence for the Open Peano Curve: $(\mathbb{F}_4, 2x^2 + 2y + 2z^2, 3)$, see Figure 1. Here the elements of the field with 4 elements $\mathbb{F}_4 = \{0, 1, \epsilon, \epsilon^2 = \epsilon + 1\}$ are denoted with $\{0, 1, 2, 3\}$.



(1) Open Peano Curve, 512×512. $(\mathbb{F}_4, 2x^2 + 2y + 2z^2, 3)$.

Corollary 1. *The recurrent double sequence Open Peano Curve given by* $(\mathbb{F}_4, 2x^2 + 2y + 2z^2, 3)$ *can be equally generated by an expansive system of substitutions of type* $2 \to 4$ *with 7 rules.*

Here we describe the expansive system of substitutions generating the Open Peano Curve. The set $\mathcal{X}$ consists of the following matrices $X_1, \ldots, X_7$:

$$\begin{matrix} 3\,3 \\ 3\,1 \end{matrix} \quad \begin{matrix} 3\,3 \\ 0\,2 \end{matrix} \quad \begin{matrix} 3\,0 \\ 3\,2 \end{matrix} \quad \begin{matrix} 2\,0 \\ 0\,3 \end{matrix} \quad \begin{matrix} 0\,3 \\ 3\,0 \end{matrix} \quad \begin{matrix} 1\,3 \\ 3\,2 \end{matrix} \quad \begin{matrix} 0\,0 \\ 0\,0 \end{matrix}$$

The set $\mathcal{Y}$ and the function Σ are defined as follows:

$$X_1 \to \begin{matrix} X_1\, X_2 \\ X_3\, X_4 \end{matrix} \quad X_2 \to \begin{matrix} X_1\, X_2 \\ X_5\, X_6 \end{matrix} \quad X_3 \to \begin{matrix} X_1\, X_5 \\ X_3\, X_6 \end{matrix} \quad X_4 \to \begin{matrix} X_6\, X_7 \\ X_7\, X_1 \end{matrix}$$

$$X_5 \to \begin{matrix} X_7\, X_3 \\ X_2\, X_7 \end{matrix} \quad X_6 \to \begin{matrix} X_4\, X_2 \\ X_3\, X_6 \end{matrix} \quad X_7 \to \begin{matrix} X_7\, X_7 \\ X_7\, X_7 \end{matrix}$$

We observe that the given system of substitutions is indeed expansive: the first rule has X_1 as a left upper minor in its right-hand side.

We also observe that the function $2x^2 + 2y + 2z^2$ is a sum of Frobenius automorphisms with coefficients, so is a homomorphisms of abelian p-groups from $(C_2 \times C_2)^3 \to C_2 \times C_2$. Using the automatic proof method explained above, the author checked positively some hundreds of homomorphisms of finite abelian groups of type $f : G^3 \to G$ and all of them produced recurrent double sequences that can be also produced by context-free substitutions. In particular, the sequences $(\mathbb{Z}/p^2\mathbb{Z}, x + y + z, 1)$ for $p \in \{3, 5, 7, 11\}$ first published as pictures in [6] are of this kind. The author has got enough experimental evidence to state the following:

Conjecture: *Let G be a finite abelian p-group, $f : G^3 \to G$ a homomorphism of p-groups, and $a \in G \setminus \{0\}$ an arbitrary element. Then the recurrent double sequence defined by (G, f, a) can be also constructed using an expansive system of context-free substitutions.*

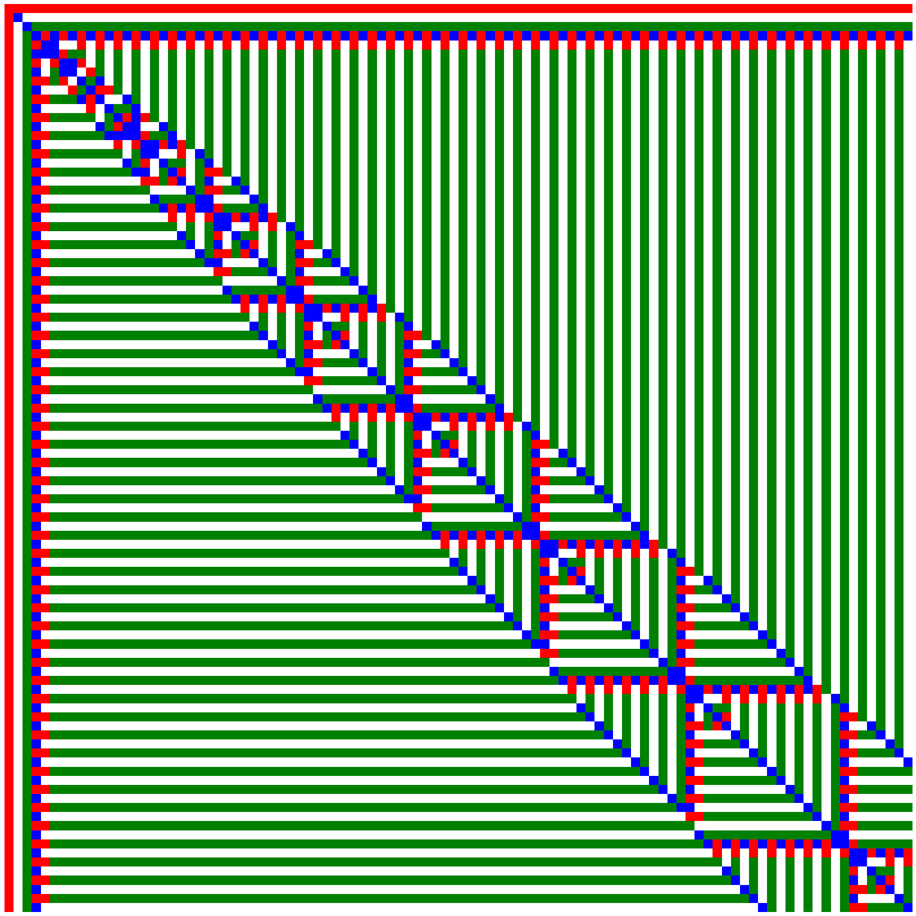
On the other side, a lot of functions $f : \mathbb{F}_5^3 \to \mathbb{F}_5$ produced also extremely complex patterns of substitution without being explicitely homomorphisms of p-groups. The next question will be, if those cases admit or not a representation using homomorphisms of p-groups.

4 A counterexample to substitution

All these examples lead to the following natural question: Is it true that all recurrent double sequences over finite sets are generated by expansive systems of context-free substitutions? The answer is negative, as already observed in [1]. Recall that $(\mathbb{Z}/n\mathbb{Z}, x + z, 1)$ produces Pascal's triangle modulo n. Pascal's triangle modulo 6 is an overlapping of Pascal's triangles modulo 2 and modulo 3, because of the isomorphism $\mathbb{Z}/6\mathbb{Z} \simeq \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}$ given by the Chinese Remainder Theorem. However, this double sequence cannot be generated by a common expansive substitution because of the fact that $\log 3/\log 2$ is irrational. Are there other phenomena leading to counterexamples? As we show here, there exist such phenomena. That one studied here originates from computability and modells the construction of the successor function.

In this section we study the recurrent double sequence Stairway to Heaven SH given by the recurrent rule $(\mathbb{F}_5, 4x^2y^4z^2 + 4x^4y^3 + 4y^3z^4 + 4y^2 + 2, 1)$. We shall prove here that this sequence cannot be generated by expansive systems of context free substitutions of any type. It must be emphasized here that this is not a counterexample to the given conjecture. It only shows that the classes of all recurrent double sequences and of all double sequences obtained by expansive systems of substitution are incomparable.

This recurrent double sequence contains in fact only four elements and accept minimal representations over the field $\mathbb{F}_4$. However, it uses only 52 triples from the whole 64 triples in $\mathbb{F}_4^3$, and it is quite difficult to look for a nice definition in the 4^{12} many possible minimal representations. The author has got a definition with 10 terms. According to Figure (2) we recall the elemets of $\mathbb{F}_5$ occurring in SH as follows: white = 0, red = 1, green = 2, blue = 3.



(2) Stairway to Heaven, 100×100. $(\mathbb{F}_5, 4x^2y^4z^2 + 4x^4y^3 + 4y^3z^4 + 4y^2 + 2, 1)$.

Theorem 3. *SH given by* $(\mathbb{F}_5, 4x^2y^4z^2 + 4x^4y^3 + 4y^3z^4 + 4y^2 + 2, 1)$ *cannot be generated using expansive systems of context-free substitutions.*

Lemma 2. *Suppose that a double sequence can be generated by an expansive system of context-free substitutions of some type* $x \to y$. *Then for all* $k \geq 2$ *the double sequence can be generated by expansive systems of context-free substitutions of type* $kx \to ky$.

Proof. Let $(\mathcal{X}, \mathcal{Y}, \Sigma, X_1)$ be an expansive system of substitutions of type $x \to y$. We define the new expansive system of substitutions $(\mathcal{X}', \mathcal{Y}', \Sigma', X_1')$ in the following way: The set $\mathcal{X}'$ consists of all $kx \times kx$ matrices consisting of k^2 many $x \times x$ blocks, where every element of $\mathcal{X}$ may occur as a block. Σ' is the natural block-wise extension of Σ. Let $\mathcal{Y}'$ be $\Sigma'(\mathcal{X}')$. Let X_1' be the $kx \times kx$ left upper minor of the double sequence generated by $(\mathcal{X}, \mathcal{Y}, \Sigma, X_1)$. Using Lemma 1 one gets that the new system of context-free substitutions is also expansive.

In the next statements the word *minor* will be used for *connected minor of SH*.

Definition 8. For $i \in \mathbb{N}$: α_i is the 5×5 minor starting with $a(c(i), c(i))$, where $c(i) = i^2 + 7i + 15$, $\alpha := \alpha_0$; β_i is the 2×2 minor $a(c(i+1) - 2, c(i+1) - 2)$, $\beta := \beta_0$; D_i is the $(8 + 2i) \times (8 + 2i)$ minor starting with $a(c(i), c(i))$.

Lemma 3. *For* $i \geq 1$ *all elements* $a(i, i)$ *are blue. All minors* α_i *are translated copies of* α. *All minors* β_i *are translated copies of* β. *The squares* D_i *cover the first diagonal for* $i \geq 15$. D_i *has* α_i *as a left upper minor and* β_i *as a right lower minor. Between* α_i *and* β_i *one finds* $i + 1$ *white stripes and* i *green stripes.*

SH interprets the set $\{n \in \mathbb{N} \mid n \geq 2\}$ in the following way: For all $i \geq 0$, the square D_i has exactly $i + 2$ many red unit-squares on every edge.

Proof. By induction. The crucial configuration to look at is the configuration around $a(c(i)-1, c(i)-2i-10)$. This configuration adds 2 to the edge of D_{i-1} to get the edge of D_i.

Lemma 4. *Consider $u \in \mathbb{N}$ even, $u \geq 8$. Let $N \in \mathbb{N}$ even, $N \geq 16$, such that a $u \times u$ minor starting with $a(N, N)$ does not contain any α_i and does not intersect any β_j. Suppose that SH is decomposed in $u \times u$ minors $U(k, h)$, where $U(k, h)$ starts with $a(ku, hu)$. Let M_u be the set of $u \times u$ matrices occurring in SH as minors $U(n, n)$ with $nu \geq N$. Then:*

- *M_u contains $u/2 + 3$ elements.*
- *Every element of M_u occurs infinitely often on the main diagonal of SH, as $U(n, n)$ with $nu \geq N$.*

Proof. $U(n, n)$ may contain at most one starting point for a β_i. Counting the possible starting points, we get $u/2$ cases. Further one has two cases where $U(n, n)$ intersects an α_i and the case where $U(n, n)$ does not intersect any α_i and any β_j.

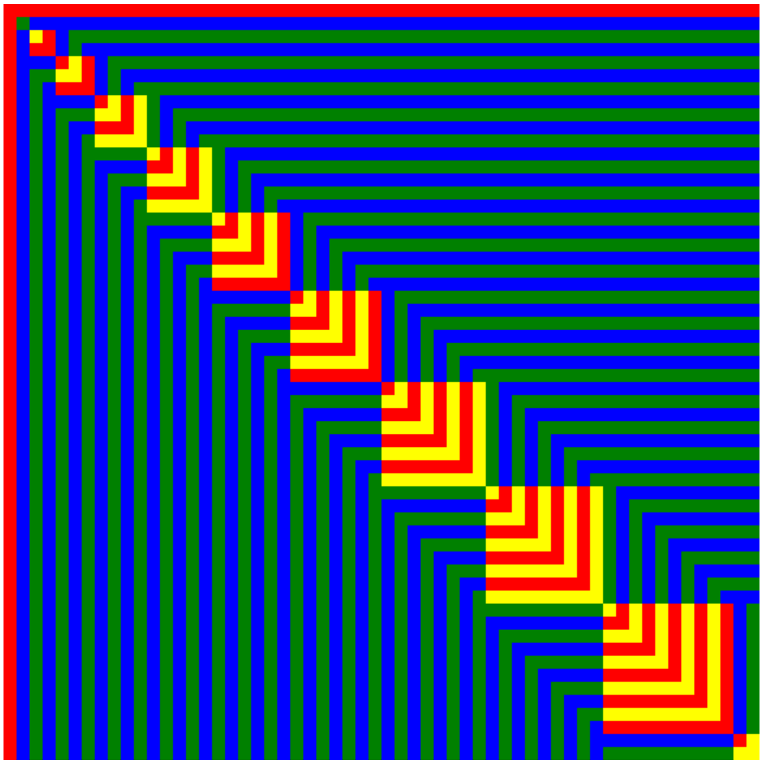
Proof (Theorem 3). Suppose that SH can be produced by some expansive system of context-free substitutions of type $x \to mx$. By Lemma 2 we can suppose that x is even and $x \geq 8$. Choose N good for $u = mx$ in the sense of Lemma 4 and observe that this value N is also good for $u = x$ in the same sense. Every element of M_{mx} occurs in SH infinitely often as minor starting in some $a(kmx, hmx)$, so all these minors must be right hand side in a substitution rule with left hand side in M_x. So $(x/2) + 3 \geq m(x/2) + 3$ which is possible only for $m = 1$. Contradiction.

5 Other counterexamples

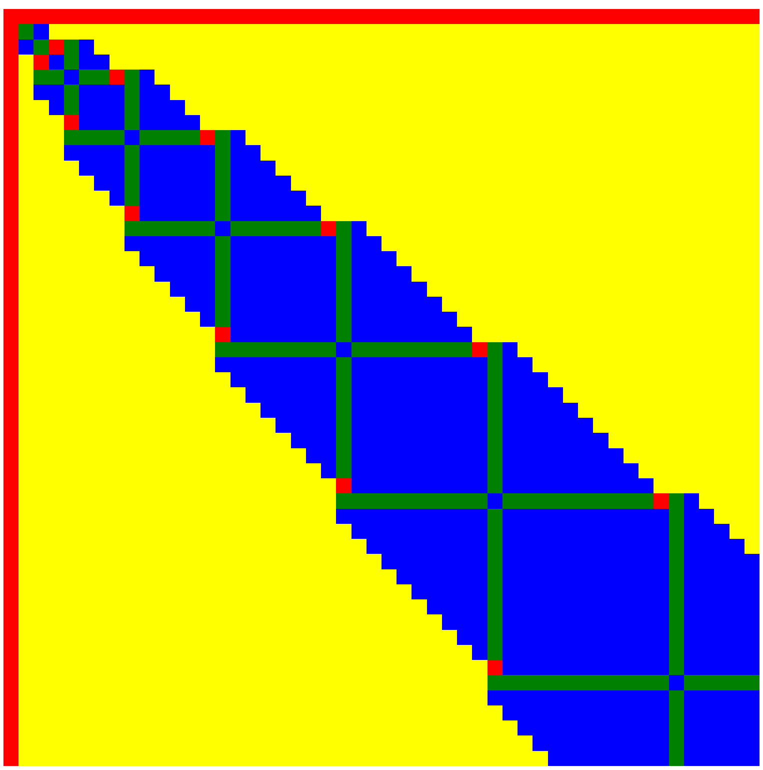
The author found other some examples of recurrent double sequences that cannot be generated by expansive systems of context-free substitutions. Although most of them are more complicated than SH, there are also two simpler examples. The reason, why they cannot be generated by systems of substitution, is the same like for SH, and the proof can be easily adapted. However, as the author believes, those examples deliver a supplementary understanding, because every one keeps an other property of SH: Second Stairway keeps the stripes, Third Stairway keeps the frames.

Second Stairway is given by $(\mathbb{F}_5, 2x^3y^3z^3 + 2xy^2 + 2y^2z + y, 1)$, see Fig. 3. This representation uses the same colours like those of SH. Moreover, yellow represents $4 \in \mathbb{F}_5$. Only four elements occur in the recurrent double sequence and only 32 triplets from 64 possible are really used. The set of all occurrences of red $= 1$ which are not on the border together with the set of all occurrences of yellow $= 4$ define the union of a sequence of growing squares. They grow in arithmetic progression with step $= 1$. The first square starts with $a(2, 2)$ and has the edge of length 2. An interesting property of this sequence of squares is that considering the colour of the element $a(i, i)$ with smalles possible i and the colour of the element $a(i, i)$ with biggest possible i, the combinations yr (yellow-red), rr, ry, yy are possible. In fact, all these types occur periodically in exactly this order in the sequence of squares, starting with an yr.

Third Stairway is given by $(\mathbb{F}_5, 4x^3yz^3 + 4x^4y^2 + 4y^2z^4 + x^2y^2z^2 + 4, 1)$, see Fig. 4. In this example we use the same colour code. Only four elements occur in the recurrent double sequence and only 30 triplets from 64 possible are really used. The set of all occurrences of red $= 1$ which are not on the border, together with green $= 2$, define the union of a sequence of growing square frames, where every frame laks two opposite corners. The first square frame starts with $a(2, 2)$, which is exceptionally green, and has an edge of length 3. The sequence of square frames grows in arithmetic progression with step 2, like SH.



(3) Second Stairway, 58×58. $(\mathbb{F}_5, 2x^3y^3z^3 + 2xy^2 + 2y^2z + y, 1)$.



(3) Third Stairway, 50×50. $(\mathbb{F}_5, 4x^3yz^3 + 4x^4y^2 + 4y^2z^4 + x^2y^2z^2 + 4, 1)$.

References

1. **J-P. Allouche, F. von Haeseler, H.-O. Peitgen, G. Skordev** *Linear cellular automata, finite automata, and Pascal's triangle.* Discrete Applied Mathematics, 66, 1 - 22, 1996.
2. **Noam Chomsky**: *Three models for the description of language.* IRE Transactions on Information Theory, 2, 113 - 124, 1956.
3. **Jarkko Kari**: *Theory of cellular automata: A survey.* Theoretical Computer Science 334, 3 - 33, 205.
4. **Roger Penrose**: *Pentaplexity.* Mathematical Intelligencer 2 (1), 32 - 37, 1979.
5. **Benoit B. Mandelbrot**: *The fractal geometry of nature.* W. H. Freeman and Company, San Francisco, 1977, 1982.
6. **Dann E. Passoja, Akhlesh Lakhtakia**: *Carpets and rugs: an exercise in numbers.* Leonardo, 25, 1, 1992, 69 - 71.
7. **Mihai Prunescu**: *Undecidable properties of recurrent double sequences.* Notre Dame Journal of Formal Logic, 49, 2, 143 - 151, 2008.
8. **Mihai Prunescu**: *Self-similar carpets over finite fields.* European Journal of Combinatorics, 30, 4, 866 - 878, 2009.
9. **Mihai Prunescu**: *Recurrent double sequences that can be produced by context-free substitutions.* Fractals Journal, 18, 1, 65 - 73, 2010.
10. **Stephen J. Willson**: *Cellular automata can generate fractals.* Discrete Applied Mathematics, 8, 1984, 91 - 99.