PU. M. A. Vol. **23** (2012), No.2, pp. 81 – 102

A construction for a class of binary words avoiding 1^j0^i

Stefano Bilotta
University of Firenze
Dipartimento di Sistemi e Informatica
Viale G.B. Morgagni 65, 50134, Firenze, Italy.
e-mail: bilotta@dsi.unifi.it

and
Elisa Pergola
University of Firenze
Dipartimento di Sistemi e Informatica
Viale G.B. Morgagni 65, 50134, Firenze, Italy.
e-mail: elisa@dsi.unifi.it

and
Renzo Pinzani
University of Firenze
Dipartimento di Sistemi e Informatica
Viale G.B. Morgagni 65, 50134, Firenze, Italy.
e-mail: pinzani@dsi.unifi.it

(Received: May 31, 2012, and in revised form September 14, 2012)

Abstract. In this paper we study the enumeration and the construction, according to the number of ones, of particular binary words avoiding the pattern $\mathfrak{p}(j,i) = 1^j0^i$, $0 < i < j$. The growth of such words can be described by particular jumping and marked succession rules. This approach enables us to obtain an algorithm which constructs all binary words having a fixed number of ones and then kills those containing the forbidden pattern.

Mathematics Subject Classification(2010). 68Q45, 68R15.

Keywords: Binary words, pattern avoidance, succession rule, generating tree.

1 Introduction

The problem of determining the appearance of a fixed *pattern* in long sequences of observation is interesting in many scientific problems.

For example, in the area of computer network security, intrusions are becoming increasingly frequent and their detection is very important. Intrusion detection is primarily concerned with the detection of illegal activities and acquisitions of privileges that cannot be detected with information flow and access control models. There are several approaches to intrusion detection, but recently this subject has been studied in relation to the pattern matching (see [1, 9, 12]).

In the area of computational biology, for example it could be interesting to detect the occurrences of a particular pattern in a genomic sequences over the alphabet $\{A, G, C, T\}$, for instance see [16, 18].

These kinds of applications are interested in the study concerning both the enumeration and the construction of particular words avoiding a given pattern over an alphabet Σ.

In particular, binary words avoiding a fixed pattern $\mathfrak{p} = p_0...p_{h-1} \in \{0,1\}^h$ constitute a regular language and can be enumerated in terms of the number of bits 1 and 0 by using classical results (see, e.g., [10, 11, 17]). Recently, in [2, 13], this subject has been studied in relation to the theory of Riordan arrays.

In [5], the authors study the enumeration and the construction, according to the number of ones, of the class $F^{[\mathfrak{p}(j)]}$, that is, the class $F \subset \{0,1\}^*$ of binary words w excluding the fixed pattern $\mathfrak{p}(j) = 1^{j+1}0^j$, $j \geq 1$, such that $|w|_0 \leq |w|_1$ for any $w \in F$, $|w|_0$ and $|w|_1$ are the number of zeroes and of ones in the word w, respectively. The enumeration problem, according to the number of ones, is solved algebraically by means Riordan arrays theory. This approach gives a *jumping and marked succession rule* describing the growth of such words. Moreover, in [5] was introduced an algorithm for constructing all binary words having a fixed number of ones and excluding those containing the forbidden pattern $\mathfrak{p}(j) = 1^{j+1}0^j$, $j \geq 1$.

In this paper, we focus on the generalization of the fixed forbidden pattern $\mathfrak{p}$, passing from $\mathfrak{p}(j) = 1^{j+1}0^j$, $j \geq 1$ to $\mathfrak{p}(j,i) = 1^j0^i$, $0 < i < j$.

In this case the theory of Riordan arrays is not applicable, while it is possible to adapt the algorithm constructing the class $F^{[\mathfrak{p}(j)]}$ with $\mathfrak{p}(j) = 1^{j+1}0^j$, $j \geq 1$, to the class $F^{[\mathfrak{p}(j,i)]}$ for any $\mathfrak{p}(j,i) = 1^j0^i$, $0 < i < j$.

The paper is organized as follows. In Section 2 we give some basic definitions and notations related to the notions of succession rule and generating tree. In particular, we introduce the concept of *jumping and marked succession rules* (see [7, 8]) which are succession rules acting on the combinatorial objects of a class and producing sons at different levels where appear marked or non-marked labels. In Section 3, we propose an algorithm to construct the set $F^{[\mathfrak{p}(j,i)]}$ for any fixed forbidden pattern $\mathfrak{p}(j,i) = 1^j0^i$, $0 < i < j$. This algorithm can be synthetically expressed by means of a particular jumping and marked succession rule which contains a parameter s depending on the shape of the word in F which it is applied to.

Conclusion are finally drawn in Section 4, which also includes possible future research lines.

2 Basic definitions and notations

A *succession rule* Ω is a system constituted by an *axiom* (a), with $a \in \mathbb{N}$, and a set of *productions* of the form:

$$(k) \rightsquigarrow (e_1(k))(e_2(k))\dots(e_k(k)), \qquad k \in \mathbb{N},\ e_i : \mathbb{N} \to \mathbb{N}.$$

A production constructs, for any given label (k), its *successors* $(e_1(k)), (e_2(k)), \dots, (e_k(k))$. In most of the cases, for a succession rule Ω, we use the more compact notation:

$$\Omega : \begin{cases} (a) & \\ (k) & \rightsquigarrow (e_1(k))(e_2(k))\dots(e_k(k)) \end{cases} \tag{1}$$

The rule Ω can be represented by means of a *generating tree*, that is a rooted tree whose vertices are the labels of Ω; where (a) is the label of the root and each node labelled (k) produces k sons labelled $(e_1(k)), (e_2(k)), \dots, (e_k(k))$, respectively. As usual, the root lies at level 0, and a node lies at level n if its parent lies at level $n-1$. If a succession rule describes the growth of a class of combinatorial objects, then a given object can be coded by the sequence of labels met from the root of the generating tree to the object itself. We refer to [3] for further details and examples.

The concept of succession rule was introduced in [6] by Chung et al. to study reduced Baxter permutations, and was later applied to the enumeration of permutations with forbidden subsequences [4, 19].

We remark that, from the above definition, a node labelled (k) has precisely k sons. A succession rule having this property is said to be *consistent*. However, we can also consider succession rules, introduced in [7], in which the value of a label does not necessarily represent the number of its sons, and this will be frequently done in the sequel.

Regular succession rules are not sufficient to handle all enumeration problems and so we consider a slight generalization called *jumping succession rule*. Roughly speaking, the idea is to consider a set of succession rules acting on the objects of a class and producing sons at different levels.

The usual notation to indicate a jumping succession rule is the following:

$$\begin{cases} (a) \\ (k) \stackrel{1}{\rightsquigarrow} (e_1(k))(e_2(k))\dots(e_k(k)) \\ (k) \stackrel{j}{\rightsquigarrow} (d_1(k))(d_2(k))\dots(d_k(k)) \end{cases} \tag{2}$$

The generating tree associated to (2) has the property that each node labelled (k) lying at level n produces two sets of sons, the first set lies at level $n+1$ and the labels are $(e_1(k)), (e_2(k)), \dots, (e_k(k))$, respectively, and the second one lies at level $n+j$, with $j > 1$, and the labels are $(d_1(k)), (d_2(k)), \dots, (d_k(k))$, respectively. For example, the jumping succession rule (3) counts the number of *2-generalized Motzkin paths* and Figure 1 shows some levels of the associated generating tree. For more details about these topics, see [8].

$$\begin{cases} (1) \\ (k) \stackrel{1}{\rightsquigarrow} (1)(2)\cdots(k-1)(k+1) \\ (k) \stackrel{2}{\rightsquigarrow} (k) \end{cases} \tag{3}$$

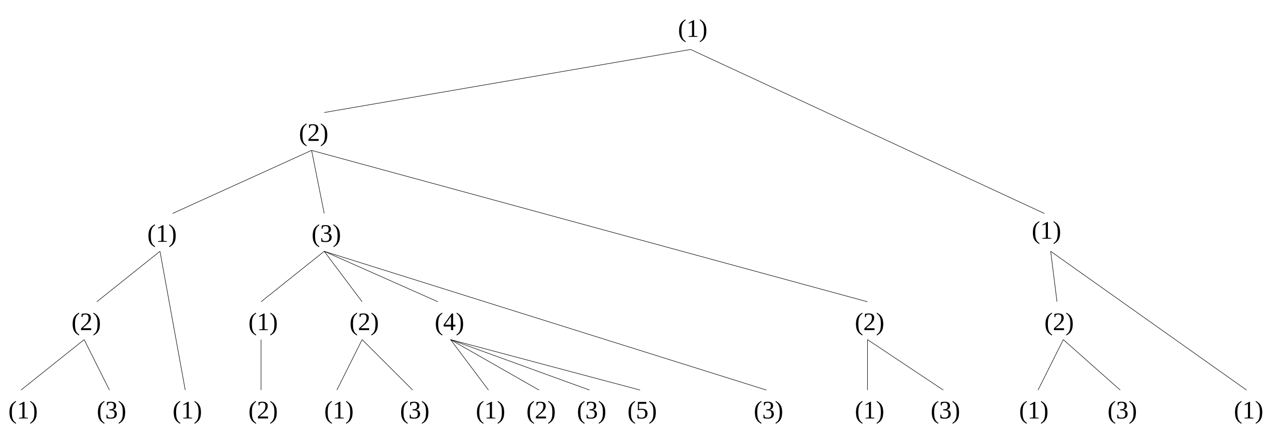

Figure 1: Four levels of the generating tree associated to the succession rule (3)

Another generalization is introduced in [14], where the authors deal with *marked succession rules*. In this case the labels appearing in a succession rule can be marked or not. In this way a generating tree can support negative values if we consider a node labelled $(\overline{k})$ as opposed to a node labelled (k) lying on the same level.

A *marked generating tree* is a rooted labelled tree where marked or non-marked labels appear according to the corresponding succession rule. The main property is that, on the same level, marked labels kill or annihilate the non-marked ones with the same label value. In particular, the enumeration of the combinatorial objects in a class is the difference between the number of non-marked and marked labels lying on a given level.

For any label (k), we introduce the following notation for generating tree specifications:

$$(\overline{\overline{k}}) = (k);$$

$$(k)^n = \underbrace{(k)\dots(k)}_{n}, \quad n > 0.$$

Each succession rule (1) can be trivially rewritten as:

$$\left\{ \begin{array}{ll} (a) & \\ (k) & \rightsquigarrow (e_1(k))(e_2(k))\dots(e_k(k))(k) \\ (k) & \rightsquigarrow (\overline{k}) \end{array} \right. \tag{4}$$

For example, the classical succession rule for Catalan numbers can be rewritten as:

$$\left\{ \begin{array}{ll} (2) & \\ (k) & \rightsquigarrow (2)(3)\dots(k)(k+1)(k) \\ (k) & \rightsquigarrow (\overline{k}) \end{array} \right. \tag{5}$$

Figure 2 shows some levels of the associated generating tree.

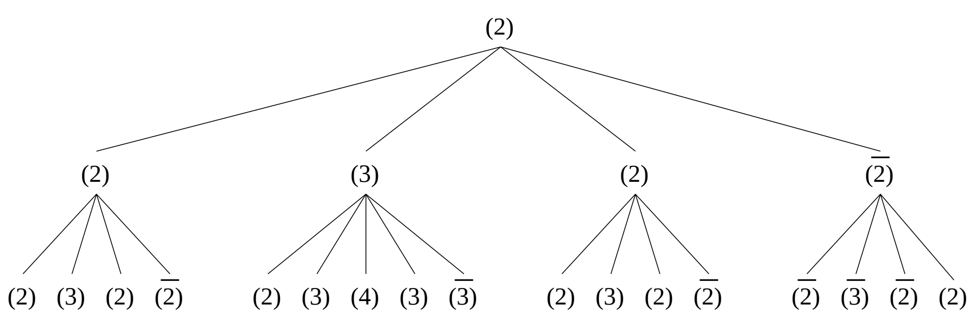

Figure 2: Three levels of the generating tree associated to the succession rule (5)

The concept of marked labels has been implicity used for the first time in [15], then in [7] in relation with the introduction of the signed ECO-systems. In Section 4, we show how marked succession rules appear in the enumeration of a class of particular binary words according to the number of ones. Let

$F \subset \{0,1\}^*$ be the class of binary words ω such that $|\omega|_0 \leq |\omega|_1$ for any $\omega \in F$, $|\omega|_0$ and $|\omega|_1$ are the number of zeroes and ones in ω, respectively.

In this paper we are interested in studying the subclass $F^{[\mathfrak{p}]}$ of F of binary words avoiding a given pattern $\mathfrak{p} = p_0 \dots p_{h-1} \in \{0,1\}^h$, i.e. the word $\omega \in F^{[\mathfrak{p}]}$ that does not admit a sequence of consecutive indices $i, i+1, \dots, i+h-1$ such that $\omega_i\omega_{i+1}\dots\omega_{i+h-1} = p_0p_1\dots p_{h-1}$. Each word $\omega \in F$ can be naturally represented as a lattice path on the Cartesian plane by associating a *rise step*, defined by $(1,1)$ and denoted by x, to each 1's in F, and a *fall step*, defined by $(1,-1)$ and denoted by $\overline{x}$, to each 0's in F. From now on, we refer interchangeably to words or their graphical representations on the Cartesian plane, that is paths.

3 A construction for the class $F^{[\mathfrak{p}(j,i)]}$

In this section, we propose an algorithm to construct the set $F^{[\mathfrak{p}(j,i)]}$, where $\mathfrak{p}(j,i) = x^j\overline{x}^i = 1^j0^i$, $0 < i < j$. The growth of the set, according to the number of rise steps or equivalently the number of ones, can be synthetically expressed by means of a jumping and marked succession rule which is sensible to the shape of the path in F which is applied to.

First of all, we define a *marked forbidden pattern* $\mathfrak{p}(j,i)$ as a pattern $\mathfrak{p}(j,i) = x^j\overline{x}^i$, $0 < i < j$, whose steps cannot be divided, that is, they must lie always in that defined sequence. Therefore, a cut operation is not possible within a marked forbidden pattern.

We denote a marked forbidden pattern by marking its peak. We say that a point is strictly contained in a marked forbidden pattern if it is between two consecutive steps of the pattern itself.

In order to study the enumeration and the construction for the class $F^{[\mathfrak{p}(j,i)]}$, we have to distinguish two cases depending on the shape of the paths in F.

Definition 3.1 A path ω in F is a Δ-*path* if:

- it ends on the x-axis (see Figure 3.a));
- the ordinate of its endpoint is greater than 0 and its rightmost suffix φ begins from the x-axis by a rise step and strictly remains above the x-axis itself. The suffix φ can contain marked forbidden patterns $\mathfrak{p}(j,i)$ (see Figure 3.b)) or not (see Figure 3.c)). If φ contains marked forbidden patterns $\mathfrak{p}(j,i)$, then their marked points have ordinate $b \geq j$.

Definition 3.2 A path ω in F is a Γ-*path* if the ordinate of its endpoint is greater that 0 and its rightmost suffix φ^* begins from the x-axis by a fall step and contains at least one marked forbidden pattern $\mathfrak{p}(j,i)$ such that its marked point has ordinate b with $i < b < j$ (see Figure 3.d)).

3.1 Δ-paths in F

For each Δ-path ω in F having k as ordinate of its endpoint, we apply the following succession rule (6), for each $k \geq 0$:

$$\begin{cases} (0) & \\ (k) & \overset{1}{\rightsquigarrow} (0)^2(1)(2)\cdots(k)(k+1) \\ (k) & \overset{j}{\rightsquigarrow} (\overline{0})^{s+1}(\overline{1})^s(\overline{2})^{s-1}\dots(\overline{s-1})^2(\overline{s})(\overline{s+1})\dots(\overline{k+j-i}) \end{cases} \qquad (6)$$

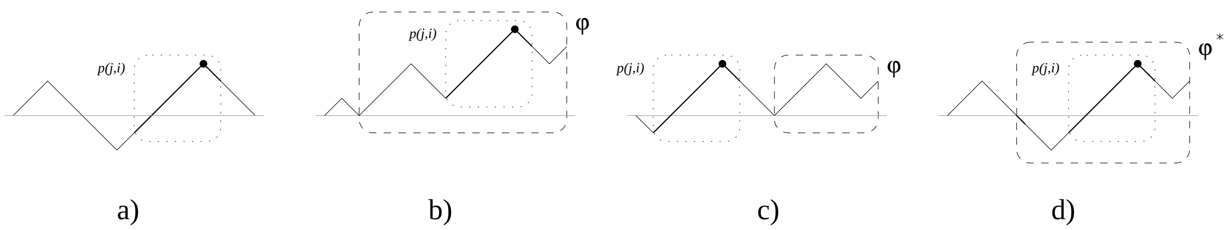

Figure 3: Some examples of paths in F

In the second production of (6), the parameter s, with $s \geq 0$, is related to the shape of the Δ-path ω and the way to find s will be described later in this section.

We define an algorithm which associates a Δ-path in F to a sequence of labels obtained by means of the succession rule (6).

The axiom (0) is associated to the empty path ε.
A Δ-path $\omega \in F$, with n rise steps and such that its endpoint has ordinate k, provides $k+3$ lattice paths, with $n+1$ rise steps, according to the first production of (6) having $0, 0, 1, \ldots, k+1$ as endpoint ordinate, respectively.

The last $k+2$ labels are obtained by adding to ω a sequence of steps consisting of one rise step followed by $k+1-h$ fall steps for each h, $0 \leq h \leq k+1$, (see Figure 4).

Each lattice path so obtained has the property that its rightmost suffix beginning from the x-axis, either remains strictly above the x-axis itself or ends on the x-axis by a fall step. Note that in this way, the paths ending on the x-axis by a rise step are never obtained. These paths are bound to the first label (0) of the first production in (6).

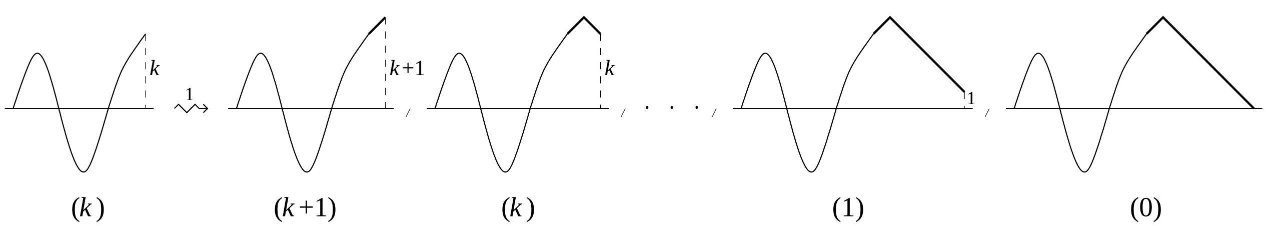

Figure 4: The mapping associated to $(k) \overset{1}{\rightsquigarrow} (0)(1)(2)\ldots(k+1)$ of (6)

In order to obtain the first label (0) according to the first production of (6), we consider the lattice path ω' obtained from ω by adding a sequence of steps consisting of one rise step followed by k fall steps. By applying the previous actions, a path ω' can be written as $\omega' = v\varphi'$, where φ' is the rightmost suffix in ω' beginning from the x-axis and strictly remaining above the x-axis.

We distinguish two cases: in the first one φ' does not contain any marked point and in the second one φ' contains at least one marked point.

If the suffix φ' does not contain any marked point, then the desired label (0) is associated to the path $v(\varphi')^c x$, where $(\varphi')^c$ is the path obtained from φ' by switching rise and fall steps (see Figure 5).

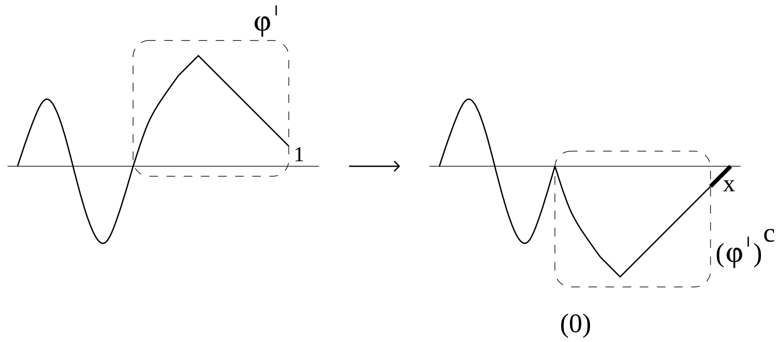

Figure 5: A graphical representation of the actions giving the first label (0) in case of no marked points in φ'

If the suffix φ' contains marked points, let $z = (x_z, y_z)$ be the leftmost point in φ' having highest ordinate, and not strictly contained in a marked forbidden pattern.

The desired label (0) is associated to the path obtained by applying cut and paste actions which consist on the concatenation of a fall step $\overline{x}$ with the path in φ' running from z to the endpoint of the path, say α, and the path running from the initial point in φ' to z, say β (see Figure 6 and 7).

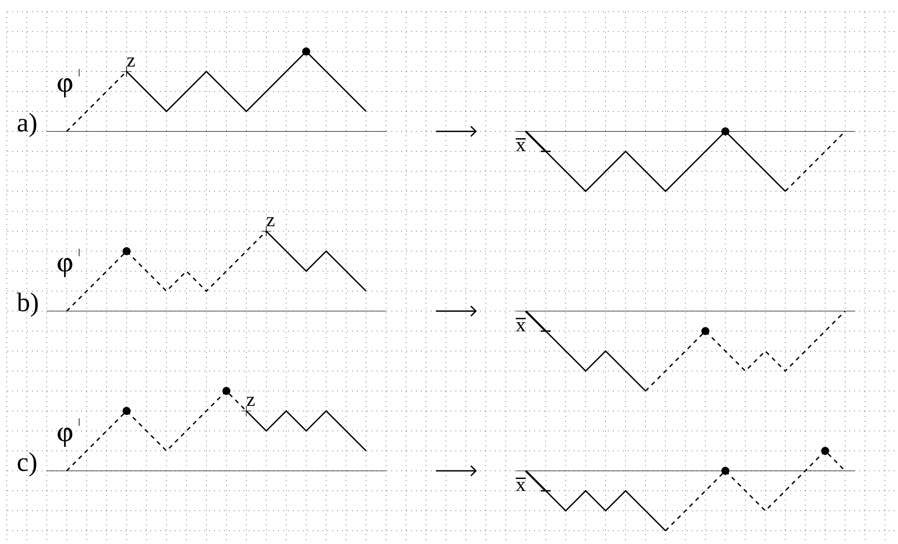

Figure 6: Some examples of the actions giving the first label (0) in the case of marked points in φ', $\mathfrak{p}(j,i) = x^2\overline{x}$

This last mapping can be inverted as follows. Let d be the rightmost fall step in a path ω^* labelled (0) beginning from the x-axis and such that each marked point, on its right, has ordinate less than j. Let us $\omega^* = v\varphi^*$, where φ^* is the rightmost suffix in ω^* beginning with d and let l be the rightmost point in φ^* having lowest ordinate. The inverted lattice path of ω^* is given by $v\beta\alpha$, where β is the path in φ^* running from l to the endpoint of the path and α is the path φ^* running from the endpoint

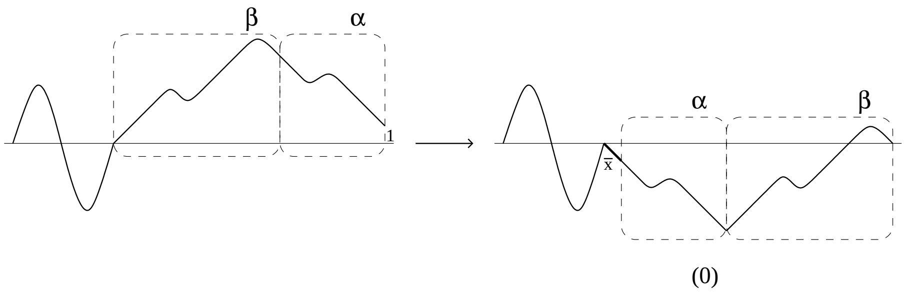

Figure 7: A graphical representation of the cut and paste actions giving the first label (0) in case of marked points in φ'

of d to l (see Figure 8).

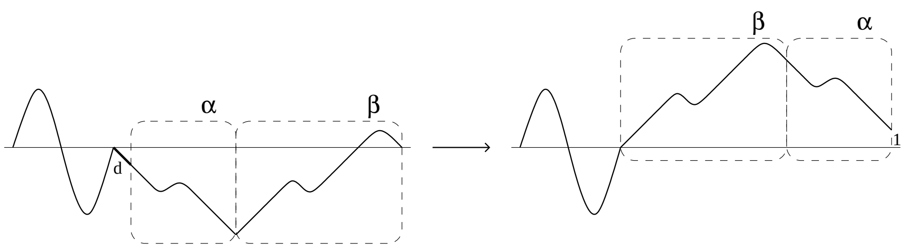

Figure 8: A graphical representation of the lattice path obtained by means of the inverted mapping related to the first label (0) in case of marked points in φ'

Let the parameter s be fixed, a Δ-path $\omega \in F$, with n rise steps and such that its endpoint has ordinate k, provides $1+k+j-i+\sum_{m=1}^{s} m$ lattice paths, with $n+j$ rise steps, according to the second production of (6). The first $1+k+j-i$ lattice paths have $0,1,2,\ldots,k+j-i$ as endpoint ordinate, respectively, and concerning the remaining $\sum_{m=1}^{s} m$ lattice paths each m of them has $s-m$ as endpoint ordinate, for each m, $1 \leq m \leq s$.

The first $1+k+j-i$ lattice paths are obtained by adding to ω a sequence of steps consisting of the marked forbidden pattern $\mathfrak{p}(j,i) = x^j\overline{x}^i$ followed by $k+j-i-h$ fall steps for each h, $0 \leq h \leq k+j-i$, (see Figure 9).

Each lattice path so obtained has the property that its rightmost suffix beginning from the x-axis, either remains strictly above the x-axis itself or ends on the x-axis by a fall step. The $\sum_{m=1}^{s} m$ marked labels according to the second production of (6), must give lattice paths having the rightmost marked point with ordinate less than j.

In order to obtain the $\sum_{m=1}^{s} m$ marked labels according to the second production of (6), we consider the paths ω'' obtained from $\omega = v\varphi$, where φ is the rightmost suffix in ω beginning from the x-axis and strictly remaining above the x-axis, by adding a sequence of steps consisting of the marked forbidden pattern $\mathfrak{p}(j,i) = x^j\overline{x}^i$ followed by $k+j-i-m$ fall steps, for each m, $1 \leq m \leq s$. Therefore, we consider the just obtained paths labelled with $(\overline{m})$, for each m, $1 \leq m \leq s$, which are represented in

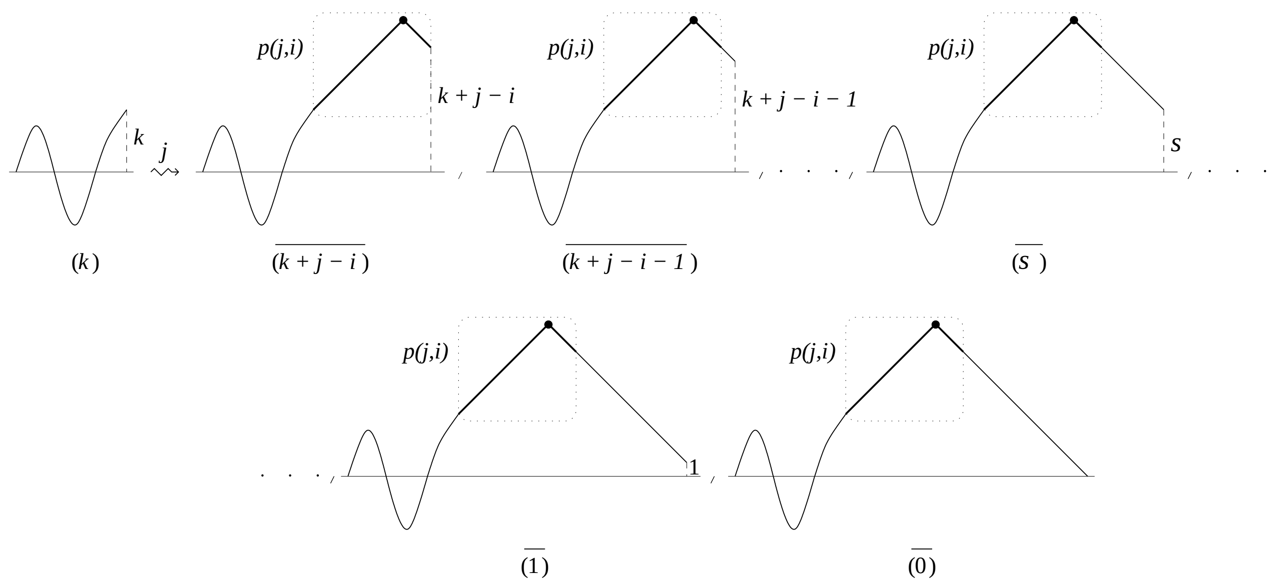

Figure 9: The mapping associated to$(k) \overset{j}{\rightsquigarrow} (\overline{0})(\overline{1})(\overline{2})\dots(\overline{s})\dots(\overline{k+j-i})$ of (6)

Figure 9.

By applying the previous actions, a path ω'' can be written as
$\omega'' = \omega\mathfrak{p}(j,i)\overline{x}^{k+j-i-m} = v\varphi\mathfrak{p}(j,i)\overline{x}^{k+j-i-m} = v\varphi''$, $1 \leq m \leq s$, where φ'' is the rightmost suffix in w'' beginning from the x-axis and strictly remaining above the x-axis (see Figure 10).

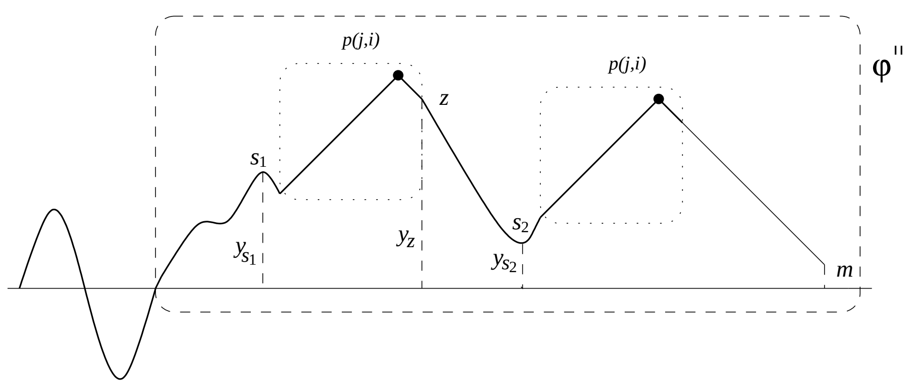

Figure 10: A graphical representation of the path$\omega'' = \omega\mathfrak{p}(j,i)\overline{x}^{k+j-i-m} = v\varphi''$, $1 \leq m \leq s$

Let $z = (x_z, y_z)$ be the leftmost point in φ'' having highest ordinate and not strictly contained in a marked forbidden pattern. Let $s_1 = (x_{s_1}, y_{s_1})$ be the point in φ'' on the left of z, having highest ordinate and not strictly contained in a marked forbidden pattern. Let $s_2 = (x_{s_2}, y_{s_2})$ be the point in φ on the right of z, having lowest ordinate and not strictly contained in a marked forbidden pattern. Then the parameter s in the second production of (6) is $s = min\{y_z - y_{s_1}, y_{s_2}\}$. When z is contained in the suffix $\mathfrak{p}(j,i)\overline{x}^{k+j-i-m}$ of ω'', $1 \leq m \leq s$, s_2 does not exist and then $s = y_z - y_{s_1}$.

By setting $s = min\{y_z - y_{s_1}, y_{s_2}\}$ we assure that, in the reverse of the cut and past actions, the point which must be taken into consideration is exactly l (see Figure 11).

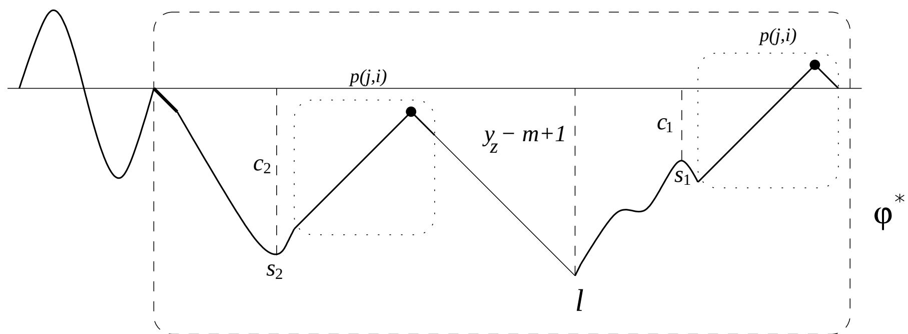

Figure 11: A graphical representation of the path $v\varphi^*$ obtained by applying the cut and paste actions to the path ω''

Remind that, from the reverse of the cut and paste actions, the point l is defined as the rightmost point in φ^* having lowest ordinate. This means that two conditions must be verified: the former one establishes that the ordinate of l must be the lowest in the path φ^* and the latter condition establishes that, if there are two or more points in φ^* having the same lowest ordinate then l is the rightmost one. In order to verify the former condition, the absolute value of the ordinate of the point s_1 in φ^*, that is $c_1 = y_z - m + 1 - y_{s_1}$, must be greater than 0, that is $m < y_z - y_{s_1} + 1$. Moreover in order to verify the latter condition, the ordinate of the point s_2, that is $c_2 = y_z - y_{s_2} + 1$, must be less than or equal to $y_z - m + 1$, that is $m \leq y_{s_2}$. So $s = min\{y_z - y_{s_1}, y_{s_2}\}$ assures that the two conditions are verified as s is the upper value which can get m.

By performing the cut and paste actions on each ω'', we obtain s paths labelled $\overline{(m-1)}$ for each m, $1 \leq m \leq s$. By adding g fall steps for each g, $0 < g \leq m-1$, to each of such paths (see Figure 12), we obtain the complete mapping associated to the second production of (6).

Note that, we apply the cut and paste actions to the paths ω'' exclusively. Indeed, by performing the cut and paste actions to the paths obtained from ω by adding a sequence of steps consisting of the marked forbidden pattern $\mathfrak{p}(j,i) = x^j\overline{x}^i$ followed by m' fall steps, for each m', $0 \leq m' < k+j-i-s$, we have already obtained paths.

Figure 13 shows the complete mapping according to the rule (6) on an example with the pattern $\mathfrak{p}(j,i) = x^5\overline{x}^2$.

3.2 Γ-paths in F

For each Γ-path ω in F having k as ordinate of its endpoint, we apply the following succession rule, for each $k \geq 1$:

$$\begin{cases} (k) & \stackrel{1}{\rightsquigarrow} (0)(1)(2)\cdots(k)(k+1) \\ (k) & \stackrel{j}{\rightsquigarrow} (\overline{0})(\overline{1})(\overline{2})\cdots(\overline{k+j-i-1})(\overline{k+j-i}) \end{cases} \quad (7)$$

A Γ-path $\omega \in F$, with n rise steps and such that its endpoint has ordinate k, provides $k+2$ lattice

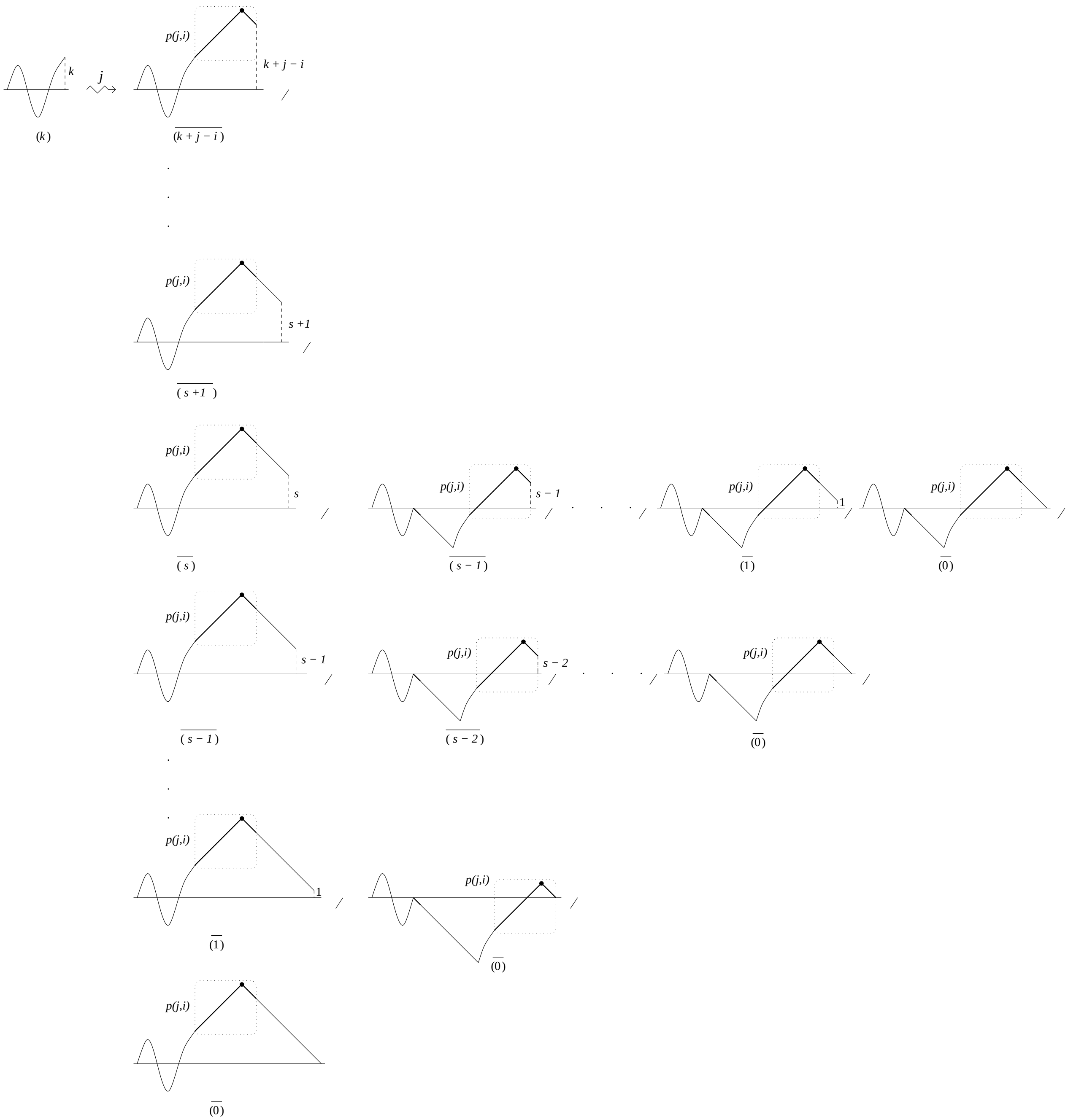

Figure 12: The mapping associated to$(k) \overset{j}{\rightsquigarrow} (\overline{0})^{s+1}(\overline{1})^{s}(\overline{2})^{s-1}\dots(\overline{s-1})^{2}(\overline{s})\dots(\overline{k+j-i})$ of (6)

paths, with $n+1$ rise steps, according to the first production of (7) having $0, 1, \dots, k+1$ as endpoint ordinate, respectively. These labels are obtained by adding to ω a sequence of steps consisting of one rise step followed by $k+1-h$ fall steps for each h, $0 \leq h \leq k+1$.

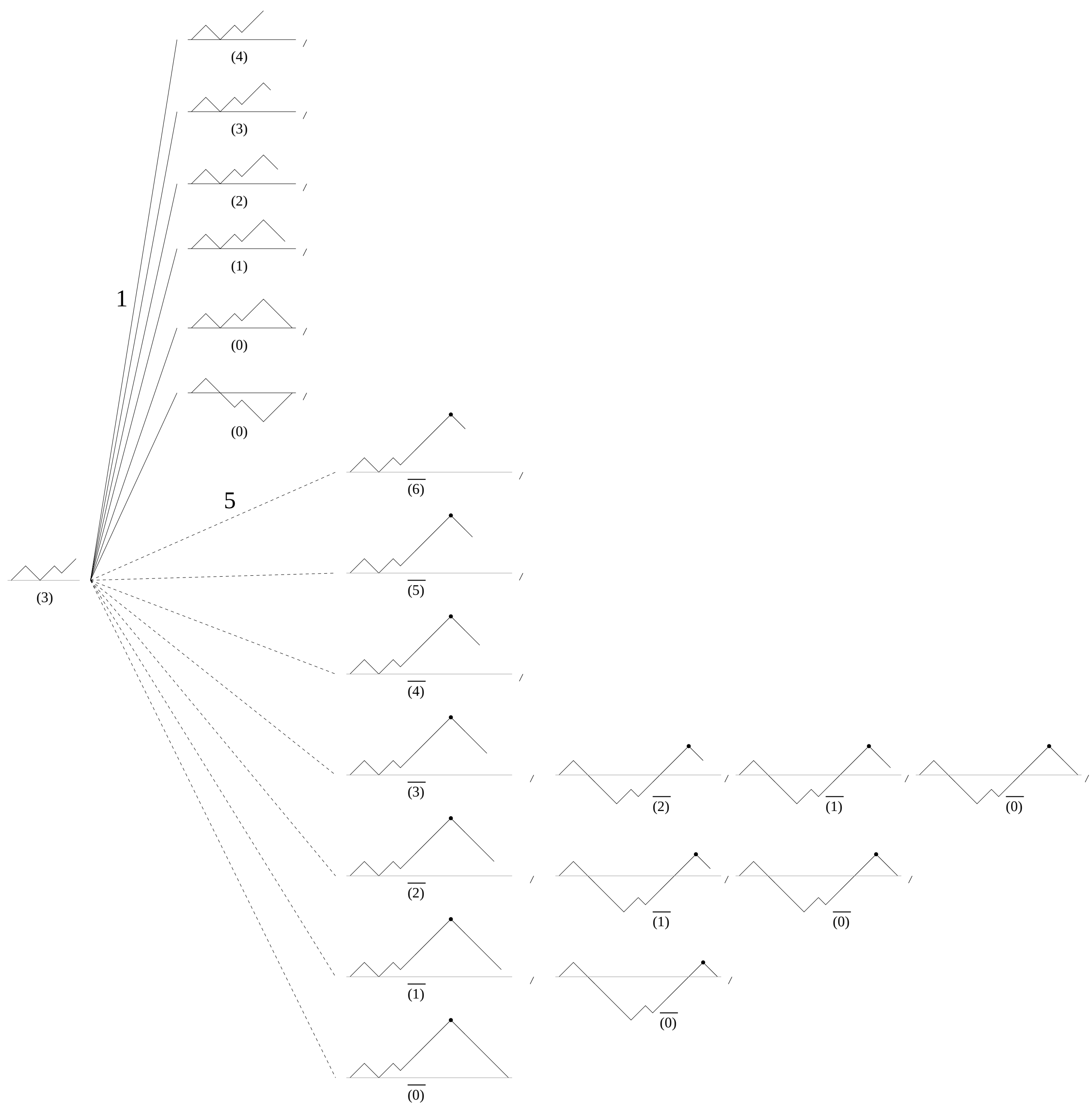

Figure 13: The mapping associated to the succession rule (6), being $\mathfrak{p}(j,i) = x^5\overline{x}^2$

Moreover, a Γ-path $\omega \in F$, with n rise steps and such that its endpoint has ordinate k, provides $1+k+j-i$ lattice paths, with $n+j$ rise steps, according to the second production of (7) having $0,1,2,\ldots,k+j-i$ as endpoint ordinate, respectively. These labels are obtained by adding to ω a sequence of steps consisting of the marked forbidden pattern $\mathfrak{p} = x^j\overline{x}^i$ followed by $k+j-i-h$ fall steps, $0 \leq h \leq k + j - i$.

Figure 14 shows the complete mapping according to the rule (7) on an example with the pattern $\mathfrak{p}(j,i) = x^5\overline{x}^2$.

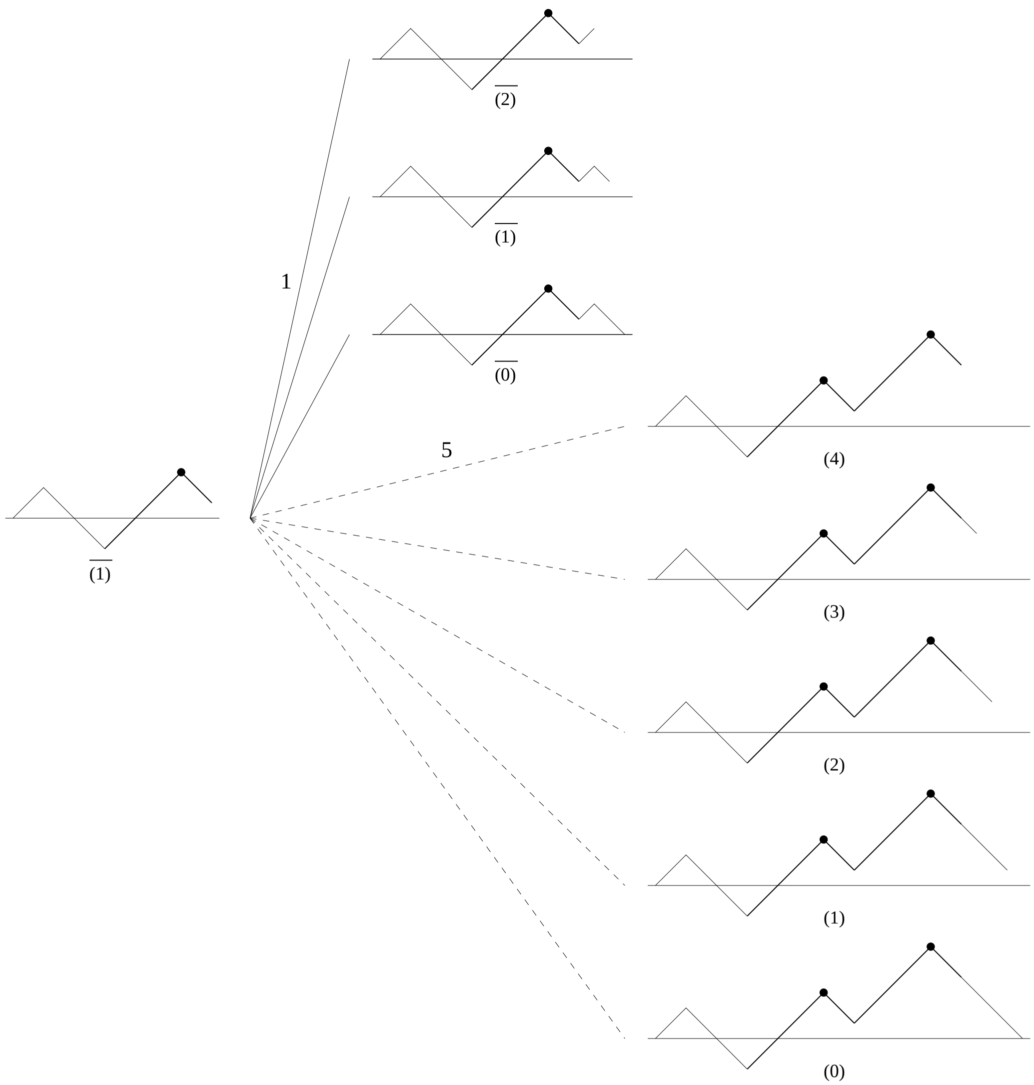

Figure 14: The mapping associated to the succession rule (7), being $\mathfrak{p}(j,i) = x^5\overline{x}^2$

3.3 Proving the construction

The just described construction, both for Δ-paths and Γ-paths in F, generates 2^C copies of each path having C forbidden patterns such that 2^{C-1} are coded by a sequence of labels ending by a marked label, say $(\overline{k})$, and contain an odd number of marked forbidden pattern, and 2^{C-1} are coded by a sequence of labels ending by a non-marked label, say (k), and contain an even number of marked forbidden pattern. For example, Figure 15 shows the 4 copies of a given path having two forbidden patterns $\mathfrak{p}(j,i) = x^5\overline{x}^2$. The sequence of labels below each path is obtained by descending into the generating tree associated to the construction from the root to the path itself.

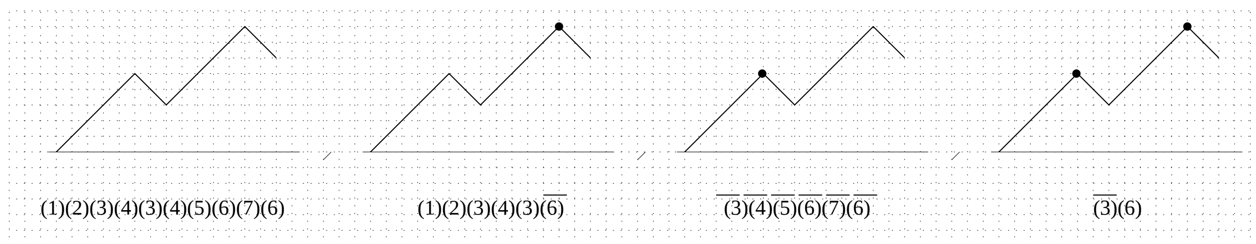

Figure 15: The 4 copies of a given path having 2 forbidden patterns, $\mathfrak{p}(j,i) = x^5\overline{x}^2$

This observation is due to the fact that when a path is obtained either according to the first production of (6) or according to the first production of (7) then no marked forbidden pattern is added. Moreover when a path is obtained either according to the second production of (6) or according to the second production of (7) then exactly one marked forbidden pattern is added. In any case, the actions performed to obtain either the first label (0) according to the first production of (6) or the $\sum_{m=1}^{s} m$ marked labels, according to the second production of (6), do not change the number of marked forbidden patterns in the path.

THEOREM 3.3 *The generating tree of the lattice paths in* $F^{[\mathfrak{p}(j,i)]}$, *where* $\mathfrak{p}(j,i) = x^j\overline{x}^i$, $0 < i < j$, *according to the number of rise steps, is isomorphic to the tree having the root labelled* (0) *and recursively defined by the succession rule (6), related to the shape of the path* $\omega \in F$, *and the succession rule (7).*

Proof. We have to show that the algorithm described in the previous pages is a construction for the set $F^{[\mathfrak{p}(j,i)]}$, according to the number of rise steps. Therefore, all the paths in F with n rise steps must be obtained and for each obtained path ω in $F \backslash F^{[\mathfrak{p}(j,i)]}$ having n rise steps, containing C forbidden patterns and having height of its endpoint equal to k, is also generated a path $\overline{\omega}$ in $F \backslash F^{[\mathfrak{p}(j,i)]}$ having n rise steps, containing C forbidden patterns, having height of its endpoint equal to k and having the same shape as ω but such that the last forbidden pattern is marked if it is not in ω and vice-versa. This means that the last label of the code associated to ω is (k) while the one associated to $\overline{\omega}$ is $(\overline{k})$.

The first assertion is an immediate consequence of the construction according to the first production of (6).

In order to prove the second assertion we have to distinguish two cases (which in their turn are subdivided in 4 and 2 subitems respectively) depending on the fact that the last forbidden pattern is marked or not. For sake of completeness we report the entire proof, which is indeed rather cumbersome. Anyhow, the reader could skip all the subitems, except the first ones. In fact, all the other ones, are slight modifications of the first subitem.

We denote by h the ordinate of the peak of the last forbidden pattern.

First case: the last forbidden pattern in ω is marked. We consider the following subcases: $h \geq j$, $0 < h < j$, $h = 0$ and $h < 0$.

1) $h \geq j$: Each path ω in $F \backslash F^{[\mathfrak{p}(j,i)]}$ can be written as $\omega = \mu x^j \overline{x}^f \nu$, where $\mu \in F$, $\nu \in F^{[\mathfrak{p}(j,i)]}$ and $i \leq f \leq d + j$ where $d \geq 0$ is the ordinate of the endpoint of μ (see Figure 16).

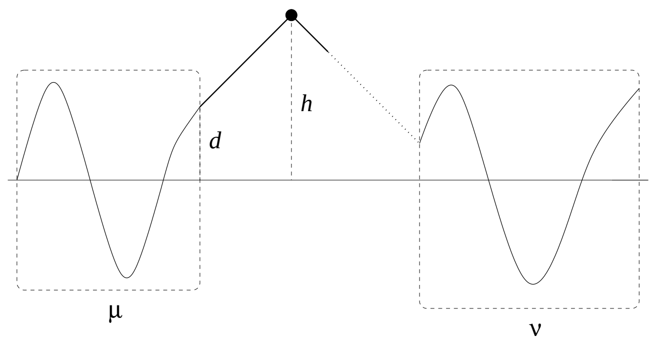

Figure 16: A graphical representation of the path ω in the case $h \geq j$

The path $\overline{\omega}$ which kills ω is obtained by performing on μ the following: add the path x^{j-1} by applying $j-1$ times the mapping associated to $(k) \overset{1}{\rightsquigarrow} (k+1)$ of the first production of (6); add the path $x\overline{x}^f$ by applying the mapping associated to $(k) \overset{1}{\rightsquigarrow} (d+j-f)$ of the first production of (6). The path ν in $\overline{\omega}$ is obtained as in ω.

2) $0 < h < j$: We consider the following subcases: $h > i$, $h = i$ and $h < i$.

2.1) $h > i$: Each path ω in $F \backslash F^{[\mathfrak{p}(j,i)]}$ can be written as $\omega = \mu \overline{x} \gamma x^j \overline{x}^f \nu$, where $\mu, \gamma \in F$, $\nu \in F^{[\mathfrak{p}(j,i)]}$ and $i \leq f \leq h$ (see Figure 17). We observe that the path γ can contain marked points, with ordinate $b < i$, or not. If the path γ contains no marked point, then it remains strictly under the x-axis, otherwise the marked forbidden patterns intersect the x-axis when $0 \leq b < i$.

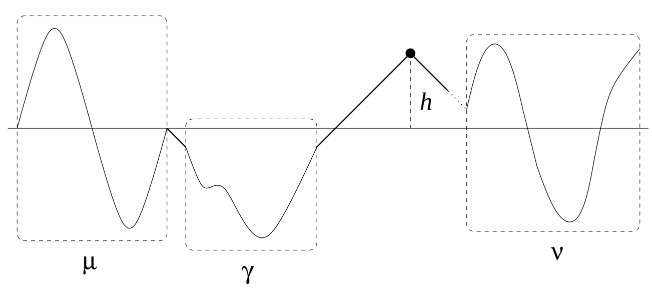

Figure 17: A graphical representation of the path ω in the case $0 < h < j$ with $h > i$

The path $\overline{\omega}$ which kills ω is obtained by performing on $\mu \overline{x} \gamma x^{j-h}$ the following: add the path x^{h-1} by applying $h-1$ times the mapping associated to $(k) \overset{1}{\rightsquigarrow} (k+1)$ of the first production of (6); add the path $x\overline{x}^f$ by applying the mapping associated to $(k) \overset{1}{\rightsquigarrow} (h-f)$ of the first production of (6). The path ν in $\overline{\omega}$ is obtained as in ω.

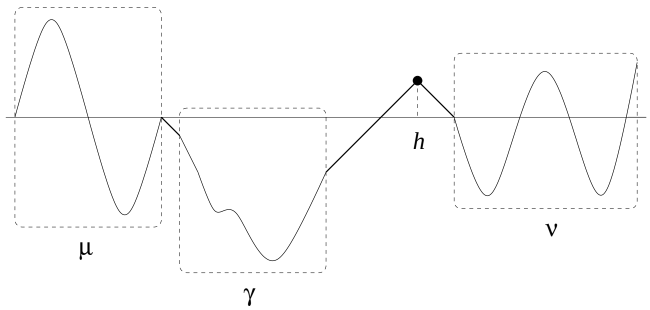

Figure 18: A graphical representation of the path ω in the case $0 < h < j$ with $h = i$

2.2) $h = i$: Each path ω in $F \backslash F^{[\mathfrak{p}(j,i)]}$ can be written as $\omega = \mu\overline{x}\gamma x^j \overline{x}^i \nu$, where $\mu, \gamma \in F$ and $\nu \in F^{[\mathfrak{p}(j,i)]}$ (see Figure 18).

The path $\overline{\omega}$ which kills ω is obtained by performing on $\mu\overline{x}\gamma x^{j-i}$ the following: add the path x^{i-1} by applying $i-1$ times the mapping associated to $(k) \overset{1}{\rightsquigarrow} (k+1)$ of the first production of (6); add the path $x\overline{x}^i$ by applying the mapping associated to $(k) \overset{1}{\rightsquigarrow} (0)$ of the first production of (6) for the second label (0). The path ν in $\overline{\omega}$ is obtained as in ω.

2.3) $h < i$: Each path ω in $F \backslash F^{[\mathfrak{p}(j,i)]}$ can be written as $\omega = \mu\overline{x}\gamma x^j \overline{x}^i \eta x \nu$, where $\mu, \gamma \in F$ and $\eta, \nu \in F^{[\mathfrak{p}(j,i)]}$ (see Figure 19). We observe that the path η remains strictly under the x-axis.

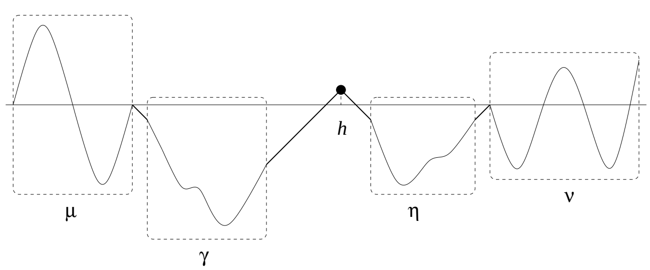

Figure 19: A graphical representation of the path ω in the case $0 < h < j$ with $h < i$

The path $\overline{\omega}$ which kills ω is obtained by performing on $\mu\overline{x}\gamma x^{j-h}$ the following: add the path x^{h-1} by applying $h-1$ times the mapping associated to $(k) \overset{1}{\rightsquigarrow} (k+1)$ of the first production of (6); add the path $x\overline{x}^h$ by applying the mapping associated to $(k) \overset{1}{\rightsquigarrow} (0)$ of the first production of (6) for the second label (0); add the path $\overline{x}^{i-h}\eta x$ by applying consecutive and appropriate mappings associated to $(k) \overset{1}{\rightsquigarrow} (1)(2)\ldots(k+1)$ of the first production of (6) and these mappings must be completed by performing the actions giving the first label (0) in case of no marked points. The path ν in ω' is obtained as in ω.

3) $h = 0$: Each path ω in $F \backslash F^{[\mathfrak{p}(j,i)]}$ can be written as $\omega = \mu\overline{x}\gamma x^j \overline{x}^i \eta x \nu$, where $\mu, \gamma \in F$ and $\eta, \nu \in F^{[\mathfrak{p}(j,i)]}$ (see Figure 20).

The path $\overline{\omega}$ which kills ω is obtained by performing on $\mu\overline{x}\gamma x^j$ the following: add the path $\overline{x}^i \eta x$

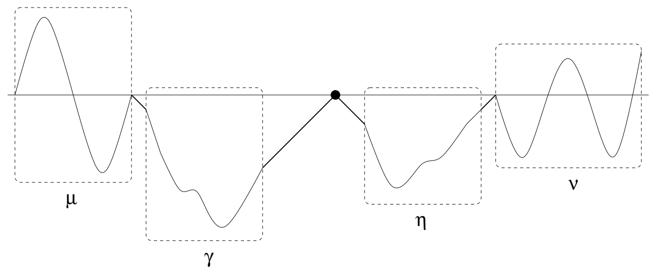

Figure 20: A graphical representation of the path ω in the case $h = 0$

by applying consecutive and appropriate mappings associated to $(k) \stackrel{1}{\rightsquigarrow} (1)\dots(k+1)$ of the first production of (6) and these mappings must be completed by performing the actions giving the first label (0) in case of no marked points. The path ν in ϖ is obtained as in ω.

4) $h < 0$: Each path ω in $F \backslash F^{[\mathfrak{p}(j,i)]}$ can be written as $\omega = \mu\overline{x}\gamma x^j\overline{x}^i\eta x\nu$, where $\mu, \gamma \in F$ and $\eta, \nu \in F^{[\mathfrak{p}(j,i)]}$ (see Figure 21).

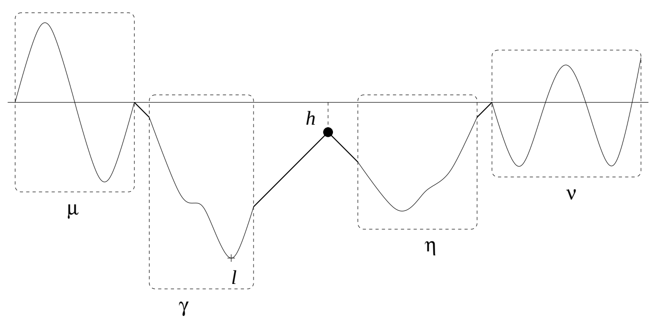

Figure 21: A graphical representation of the path ω in the case $h < 0$

We distinguish two subcases: in the first one the path γ contains no marked points and remains strictly under the x-axis and in the second one the path γ contains at least a marked point.

In the first subcase, the path ϖ which kills ω is obtained by performing on μ the following: add the path $\overline{x}\gamma x^j\overline{x}^i\eta x$ by applying consecutive and appropriate mappings associated to $(k) \stackrel{1}{\rightsquigarrow} (1)(2)\dots(k+1)$ of the first production of (6) and these mappings must be completed by performing the actions giving the first label (0) in case of no marked points. The path ν in ϖ is obtained as in ω.

In the second subcase, we consider the rightmost point l of the path $\overline{x}\gamma x^j\overline{x}^i\eta x$ with lowest ordinate. The path ϖ which kills ω is obtained by performing on μ the following: add the path in $\gamma x^{j+1}\overline{x}^j\eta x$ running from l to the endpoint of the path by applying consecutive and appropriate mappings associated to $(k) \stackrel{1}{\rightsquigarrow} (1)(2)\dots(k+1)$ of the first production of (6) and by applying consecutive and appropriate mappings associated to $(k) \stackrel{j}{\rightsquigarrow} (\overline{1})(\overline{2})\dots(\overline{k+j-i})$ of the second production of (6); add the path in $\gamma x^j \overline{x}^i \eta x$ running from its initial point to l by applying consecutive and appropriate mappings associated to $(k) \overset{1}{\rightsquigarrow} (1)(2)\ldots(k+1)$ of the first production of (6) and by applying consecutive and appropriate mappings associated to $(k) \overset{j}{\rightsquigarrow} (\overline{1})(\overline{2})\ldots(\overline{k+j-i})$ of the second production of (6); apply the cut and paste actions giving the first label (0) in case of marked points. Obviously the last forbidden pattern in the path must be generated by applying consecutive and appropriate mappings of the first production of (6). The path ν in $\overline{\omega}$ is obtained as in ω.

Second case: the last forbidden pattern in ω is not a marked forbidden pattern. We consider the following subcases: $h \geq j$ and $h < j$.

1) $h \geq j$: Each path ω in $F \backslash F^{[\mathfrak{p}(j,i)]}$ can be written as $\omega = \mu x^j \overline{x}^f \nu$, where $\mu \in F$, $\nu \in F^{[\mathfrak{p}(j,i)]}$ and $i \leq f \leq d+j$ where $d \geq 0$ is the ordinate of the endpoint of μ (see Figure 22).

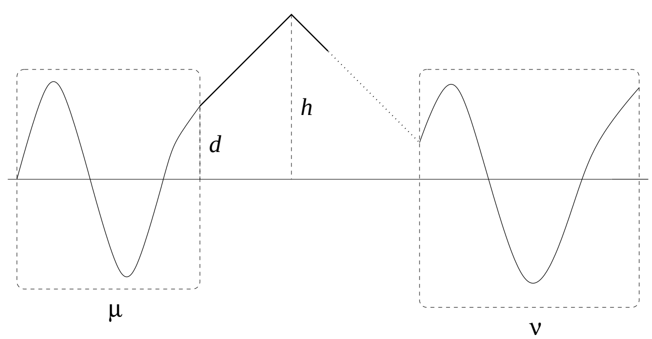

Figure 22: A graphical representation of the path ω in the case $h \geq j$

The path $\overline{\omega}$ which kills ω is obtained by performing on μ the following: add the path $x^j \overline{x}^f$ by applying the mapping associated to $(k) \overset{j}{\rightsquigarrow} (\overline{d+j-f})$ of the second production of (6). The path ν in $\overline{\omega}$ is obtained as in ω.

2) $h < j$: We consider the following subcases: $h > i$, $h = i$ and $h < i$.

2.1) $h > i$: Each path ω in $F \backslash F^{[\mathfrak{p}(j,i)]}$ can be written as $\omega = \mu \overline{x} \gamma x^j \overline{x}^f \nu$, where $\mu, \gamma \in F$, $\nu \in F^{[\mathfrak{p}(j,i)]}$ and $i \leq f \leq h$ (see Figure 23). We observe that the path γ can contain marked points, with ordinate $b < i$, or not. If the path γ contains no marked point, then it remains strictly under the x-axis, otherwise the marked forbidden patterns intersect the x-axis when $0 \leq b < i$. Let l be the rightmost point of the path $\overline{x} \gamma x^j \overline{x}^i$ with lowest ordinate. The path $\overline{\omega}$ which kills ω is obtained by performing on μ the following: add the path in $\gamma x^j \overline{x}^i$ running from l to the endpoint of the path by applying consecutive and appropriate mappings associated to $(k) \overset{1}{\rightsquigarrow} (1)(2)\ldots(k+1)$ of the first production of (6) and by applying consecutive and appropriate mappings associated to $(k) \overset{j}{\rightsquigarrow} (\overline{1})(\overline{2})\ldots(\overline{k+j-i})$ of the second production of (6); add the path in $\gamma x^j \overline{x}^i$ running from its initial point to l by applying consecutive and appropriate mappings associated to $(k) \overset{1}{\rightsquigarrow} (1)(2)\ldots(k+1)$ of the first production of (6) and

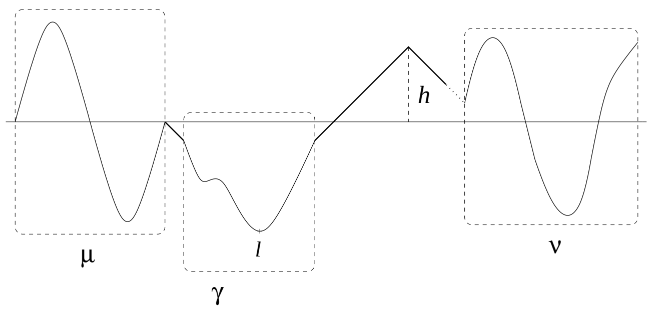

Figure 23: A graphical representation of the path ω in the case $h < j$ with $h > i$

by applying consecutive and appropriate mappings associated to $(k) \stackrel{j}{\rightsquigarrow} (\overline{1})(\overline{2})\dots(\overline{k+j-i})$ of the second production of (6); apply the cut and paste actions in case of marked points and add the path $\overline{x}^{f-i}$ according to the second production of (6). Obviously the last forbidden pattern in the path must be generated by applying the mapping of the second production of (6). The path ν in $\overline{\omega}$ is obtained as in ω. Note that, in case of $i \leq f < h$, any prefix of ν in $\overline{\omega}$ which running from the end of the path $\mu\overline{x}\gamma x^j\overline{x}^f$ to the x-axis is obtained by applying the mapping associated to the first production of (7).

2.2) $h = i$: Each path ω in $F \backslash F^{[\mathfrak{p}(j,i)]}$ can be written as $\omega = \mu\overline{x}\gamma x^j\overline{x}^i\nu$, where $\mu, \gamma \in F$ and $\nu \in F^{[\mathfrak{p}(j,i)]}$ (see Figure 24).

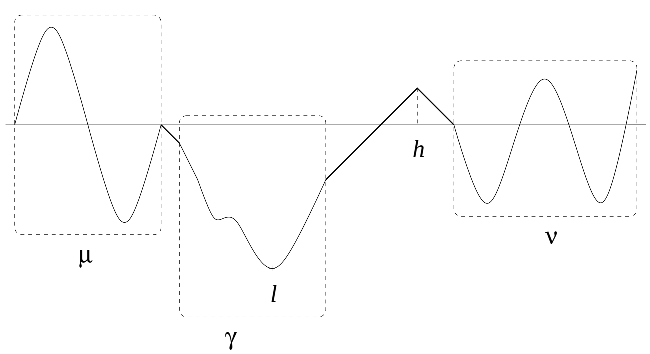

Figure 24: A graphical representation of the path ω in the case $h < j$ with $h = i$

Let l be the rightmost point of the path $\overline{x}\gamma x^j\overline{x}^i$ with lowest ordinate. The path $\overline{\omega}$ which kills ω is obtained by performing on μ the following: add the path in $\gamma x^j\overline{x}^i$ running from l to the endpoint of the path by applying consecutive and appropriate mappings associated to $(k) \stackrel{1}{\rightsquigarrow} (1)(2)\dots(k+1)$ of the first production of (6) and by applying consecutive and appropriate mappings associated to $(k) \stackrel{j}{\rightsquigarrow} (\overline{1})(\overline{2})\dots(\overline{k+j-i})$ of the second production of (6); add the path in $\gamma x^j\overline{x}^i$ running from its initial point to l by applying consecutive and appropriate mappings associated to $(k) \stackrel{1}{\rightsquigarrow} (1)(2)\dots(k+1)$ of the first production of (6) and by applying consecutive and appropriate mappings associated to $(k) \stackrel{j}{\rightsquigarrow} (\overline{1})(\overline{2})\dots(\overline{k+j-i})$

of the second production of (6); apply the cut and paste actions giving the label (0) in case of marked points. Obviously the last forbidden pattern in the path must be generated by applying the mapping of the second production of (6). The path ν in $\overline{\omega}$ is obtained as in ω.

2.3) $h < i$: Each path ω in $F \backslash F^{[\mathfrak{p}(j,i)]}$ can be written as $\omega = \mu\overline{x}\gamma x^j \overline{x}^i \eta x \nu$, where $\mu, \gamma \in F$ and $\eta, \nu \in F^{[\mathfrak{p}(j,i)]}$ (see Figure 25).

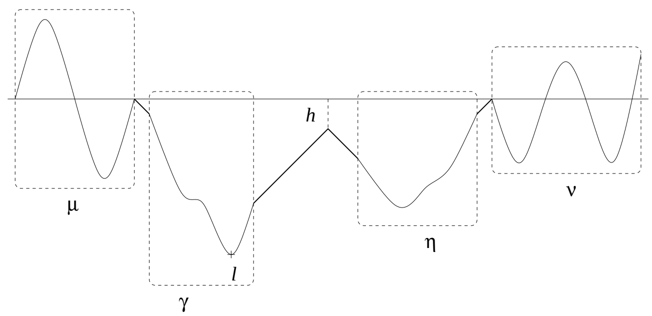

Figure 25: A graphical representation of the path ω in the case $h < j$ with $h < i$

Let l be the rightmost point of the path $\overline{x}\gamma x^j \overline{x}^i \eta x$ with lowest ordinate. The path $\overline{\omega}$ which kills ω is obtained by performing on μ the following: add the path in $\gamma x^j \overline{x}^i \eta x$ running from l to the endpoint of the path by applying consecutive and appropriate mappings associated to $(k) \overset{1}{\rightsquigarrow} (1)(2)\dots(k+1)$ of the first production of (6) and by applying consecutive and appropriate mappings associated to $(k) \overset{j}{\rightsquigarrow} (\overline{1})(\overline{2})\dots(\overline{k+j-i})$ of the second production of (6); add the path in $\gamma x^j \overline{x}^i \eta x$ running from its initial point to l by applying consecutive and appropriate mappings associated to $(k) \overset{1}{\rightsquigarrow} (1)(2)\dots(k+1)$ of the first production of (6) and by applying consecutive and appropriate mappings associated to $(k) \overset{j}{\rightsquigarrow} (\overline{1})(\overline{2})\dots(\overline{k+j-i})$ of the second production of (6); apply the cut and paste actions giving the label (0) in case of marked points. Obviously the last forbidden pattern in the path must be generated by applying the mapping of the second production of (6). The path ν in $\overline{\omega}$ is obtained as in ω.

We observe that for each path ω in $F \backslash F^{[\mathfrak{p}(j,i)]}$ having n rise steps, containing C forbidden patterns and having last label (k) of the associated code, there exists one and only one path $\overline{\omega}$ in $F \backslash F^{[\mathfrak{p}(j,i)]}$ having n rise steps, containing C forbidden patterns and having last label $(\overline{k})$ of the associated code. The paths ω and $\overline{\omega}$ have the same shape, exactly the same number and positions of the forbidden patterns except for the last one which is marked in $\overline{\omega}$ if it is not in ω and vice-versa.

This assertion is consequence of the constructions in the proof, as the described actions are univocally determined. Therefore, it is not possible to obtain a path $\overline{\omega}$ which kills a given path ω by applying two distinct procedures. □

4 Conclusions and further developments

In this paper we study the construction and the enumeration, according to the number of ones, of particular binary words avoiding the pattern $\mathfrak{p}(j,i) = 1^j 0^i$, $0 < i < j$. The proposed constructive algorithm for the class $F^{[\mathfrak{p}(j,i)]}$ represents a generalization of [5] whose the class $F^{[\mathfrak{p}(j)]}$ with $\mathfrak{p}(j) = 1^{j+1}0^j$, $j \geq 1$ is studied.

Afterwards, it should be interesting to study words avoiding patterns having a different shape, that is not only patterns consisting of a sequence of rise steps followed by a sequence of fall steps. This could be the first step to investigate a possible uniform constructive algorithm for pattern avoidance in words.

Successive studies could take into consideration binary words avoiding a set of different shape forbidden patterns both from an enumerative and a constructive point of view.

A further step could consider pattern avoidance in words defined on an arbitrary alphabet having cardinality greater than two.

References

[1] A. Apostolico and M. Atallah, *Compact recognizers of episode sequences*, Inform. and Comput., 174 (2002) 180–192.

[2] D. Baccherini, D. Merlini and R. Sprugnoli, *Binary words excluding a pattern and proper Riordan arrays*, Discrete Math., 307 (2007) 1021–1037.

[3] E. Barcucci, A. Del Lungo, E. Pergola and R. Pinzani, *ECO: a methodology for the Enumeration of Combinatorial Objects*, J. Difference Equ. Appl., 5 (1999) 435–490.

[4] E. Barcucci, A. Del Lungo, E. Pergola and R. Pinzani, *From Motzkin to Catalan permutations*, Discrete Math., 217 (2000) 33–49.

[5] S. Bilotta, D. Merlini, E. Pergola and R. Pinzani, *Pattern $1^{j+1}0^j$ avoiding binary words*, Fund. Inform., 117 (2012) 35-55.

[6] F. R. K. Chung, R. L. Graham, V. E. Hoggatt and M. Kleimann, *The number of Baxter permutations*, J. Combin. Theory Ser. A, 24 (1978) 382–394.

[7] S. Corteel, *Séries génératrices exponentielles pour les ECO-systèmes signés*, Proceedings of the 12-th International Conference on Formal Power Series and Algebraic Combinatorics 2000, Moscow, Russia.

[8] L. Ferrari, E. Pergola, R. Pinzani and S. Rinaldi, *Jumping succession rules and their generating functions*, Discrete Math., 271 (2003) 29–50.

[9] P. Flajolet, W. Szpankowski and B. Vallée, *Hidden word statistics*, J. ACM, 53 (2006) 147–183.

[10] L. J. Guibas and M. Odlyzko, *Long repetitive patterns in random sequences*, Zeitschrift für Wahrscheinlichkeitstheorie, 53 (1980) 241–262.

[11] L. J. Guibas and M. Odlyzko, *String overlaps, pattern matching, and nontransitive games*, J. Combin. Theory Ser. A, 30 (1981) 183–208.

[12] S. Kumar and E.H. Spafford, *A pattern matching model for misuse intrusion detection*, Proceedings of the 17th National Computer Security Conference 1994, 11–21.

[13] D. Merlini and R. Sprugnoli, *Algebraic aspects of some Riordan arrays related to binary words avoiding a pattern*, Theoret. Comput. Sci., 412 (2011) 2988–3001.

[14] D. Merlini, R. Sprugnoli and M. C. Verri, *An algebra for proper generating tree*, In Algorithms, trees, combinatorics and probabilities. Trends in Mathematics, Mathematics and Computer Science, (2000) 127–139.

[15] D. Merlini and M. C. Verri, *Generating trees and proper Riordan arrays*, Discrete Math., 218 (2003) 167–183.

[16] I. Rigoutsos, A. Floratos, L. Parida, Y. Gao and D. Platt, *The emergence of pattern discovery techniques in computational biology*, Metabolic Engineering, 2 (2000) 159–177.

[17] R. Sedgewick and P. Flajolet, *An Introduction to the Analysis of Algorithms*, Addison-Wesley, Reading, MA, 1996.

[18] M. Waterman, *Introduction to Computational Biology*, Chapman & Hall, London, 1995.

[19] J. West, *Generating trees and the Catalan and Schröder numbers*, Discrete Math., 146 (1995) 247–262.