Università degli Studi di Firenze

DOTTORATO DI RICERCA IN
Ingegneria Informatica e dell'Automazione

CICLO XXV

COORDINATORE Prof. Luigi Chisci

Pattern avoiding words and cross-bifix-free sets: a general approach

Settore Scientifico Disciplinare INF/01

Dottorando
Dott. Stefano Bilotta

Tutori
Prof.ssa Elisa Pergola, Prof. Renzo Pinzani

Anni 2010/2012

Acknowledgments

First of all, I would like to deeply thanks my tutors, Professors Elisa Pergola and Renzo Pinzani, for the continuous support in my studies during these years. I am very grateful to them for the opportunity to work on such interesting topics. By means of their passion, expertise and advices, I grew up both from professional and personal side.

I would also like to express my sincere gratitude to my research group in Firenze and in Siena: Gianni Aguzzi, Elena Barcucci, Antonio Bernini, Stefano Brocchi, Filippo Disanto, Andrea Frosini, Elisabetta Grazzini, Luca Ferrari, Maddalena Poneti and Simone Rinaldi.

Finally, thanks to my family and my friends for supporting me through all my life.

Contents

Introduction

This work is placed in the area of the Combinatorics, which plays a fundamental role in an ever and ever increasing number of fields.

The problems which Combinatorics must face arise from, for example, Theoretical Computer Science (average or worst case analysis of algorithms and data structures [33, 34, 69], sorting problems [46, 64, 80], analysis of regularities in words [16, 52], particular instances of pattern matching algorithms optimization [20], decidability problems concerning tiling of the plane [14, 24]), Mathematics (representations of symmetric group [36, 65], basis of symmetric functions [37], etc.), Hydrogeology (representation of the morphological structure of river networks [44, 75]), Anatomy (representation of bronchial networks or arterial vessel [82]), Molecular Biology (theoretical considerations about secondary structures of single nucleic acids [58, 77, 78]), Botany [53], Neurophysiology [61] and Statistical Physics (percolation phenomena [21]).

In general Combinatorics find application every time a phenomena can be modelled by discrete structures.

Enumerative Combinatorics is one of the main area of Combinatorics, its scope is counting the number of elements in a finite set in an exact or approximate way.

Let us consider a given class $\mathcal{O}$ of object, and a parameter p on this class, $p : \mathcal{O} \to \mathbb{N}$, $\mathbb{N}$ being the set of non negative integers. We focus on the set $\mathcal{O}_n$ of the elements of $\mathcal{O}$ such that the value of the parameter is equal to n. Enumerative Combinatorics answers the question to know the cardinality f_n of each set $\mathcal{O}_n$ for all possible n.

The most satisfactory form of f_n is a completely explicit closed formula involving only well known functions, and free from summation symbols. Only in a few case such a formula exists. A recurrence for f_n may be given in terms of previously calculated values f_k, thereby giving a simple procedure for calculating f_n for any $n \in \mathbb{N}$.

Another method to evaluate f_n is to give the formal power series $f(x) = \sum_n f_n x^n$ which is said to be the *generating function* of the class of object $\mathcal{O}$, according to the parameter p. Notice that the n-th coefficient of the Taylor series of $f(x)$ is just the term f_n.

In enumeration problems the generating function of a class of objects often appears as the solution of an equation, and it is classified following the type of the equation from which it rises from; that is, rational, algebraic, differential, holonomic, etc.

A common approach to enumeration consists in searching for a construction of the class of objects under consideration and later of translating it into a recursive relation, or an equation (usually called *functional equation*) satisfied by the generating function of the objects.

Schützenberger [68] developed a methodology which consists in codifying the objects with words of a language such that the size of the objects corresponds to the to the length of the words of the language. If the language is generated by an unambiguous context-free grammar, then it is effectively possible to translate the productions of the grammar into a system of functional equations whose solution is unique and is the generating function of the words of the language.

Finally, The ECO method [12] is a unifying method for Enumerating some classes of Combinatorial Objects, based on the recursive construction of the objects of a given size realized by means a local expansions performed on the objects of immediately lower size. Enumeration is not the only role played by ECO method in Combinatorics, since it allows us to perform exhaustive and random generation algorithms. *Succession rules* are a fundamental tool applied in the context of ECO method.

More recently, succession rules have been considered as remarkable objects to be studied independently from of their applications, and they have been treated by several points of view. In [8], Banderier et al. explore in detail the relationship between the form a succession rule and the related generating function. Besides, some algebraic properties of succession rules – represented in terms of a *rule operator* – have been determined and studied in [30].

Furthermore, some extensions of the concept of succession rule have been proposed. In [31] are introduced the so called *jumping succession rules*, while in [25] the author presents *marked succession rules*. These extensions empower the “original” tool.

In order to point out the great versatility of succession rules we present different research areas in which succession rules can be applied to.

The first line of research presented in this thesis is related to pattern avoidance, that is, strings over a given alphabet which can not contain a determined substring. The study of this subject is interesting in many scientific fields.

In the area of computer network security, the detection of intrusions, which are becoming increasingly frequent, is very important. Intrusion detection is primarily concerned with the detection of illegal activities and acquisitions of privileges that can not 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 pattern matching (see [1, 35, 48]).

In the area of computational biology it could be interesting to detect the occurrences of a particular pattern in a genomic sequence over the alphabet $\{A, G, C, T\}$, see for instance [63, 79].

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

If we consider the set of binary words without any restriction then the binary words avoiding a fixed pattern constitute a regular language and can be enumerated by using classical results obtaining rational generating functions (see, e.g., [40, 41, 45, 69]). When the restriction to words with no more 0's than 1's is valid, the language of words avoiding a pattern is more difficult to deal with. For each forbidden pattern an "ad hoc" grammar (from which the generating function can be obtained) should be defined. Consequently, for each pattern a different generating function enumerating the words avoiding it must to be computed.

Our first aim is to determine an algorithmic unified approach, based on the concept of succession rule, for the construction and the enumeration, according to the number of ones, of binary words in $\{0,1\}^*$ having the number of 1's greater than or equal to the number of 0's, and which can be applied to any forbidden pattern.

Moreover, we extend this approach to the class of binary words in $\{0,1\}^*$ having the number of 1's greater than or equal to the number of 0's, and avoiding a set of forbidden patterns. Surprisingly, the number of words avoiding a set of forbidden patterns does not depend on the shape of the avoided patterns themselves, but only on the number of ones in the patterns.

In the second line of research, succession rules are studied posing a specific relevance on their relationships with recurrence relations.

More recently, there have been some efforts in developing methods to pass from a recurrence relation defining an integer sequence to a succession rule defining the same sequence - in this case we say that the succession rule and the recurrence relation are *equivalent*.

It is worth mentioning that almost all studies realized until now on this topic have regarded linear recurrence relations with integer coefficients [19, 27]. Following Zeilberger [83], we will address to these as *C-finite recurrence relations*, and to the defined sequences as *C-finite sequences*. One of the open problems on this subject is the so called *Positivity Problem*, that is, given a C-finite sequence $\{f_n\}_{n\geq 0}$, establish whether all its terms are positive.

This problem was originally proposed as an open problem in [17], and then re-presented in [66] (Theorems 12.1-12.2, pages 73-74), but no general solution has been found yet.

The positivity problem can be solved for a large class of C-finite sequences, precisely those whose generating function is a $\mathbb{N}$-rational series. We also recall that the class of $\mathbb{N}$-rational series is precisely the class of the generating functions of regular languages, and that Soittola's Theorem [72] states that the problem of establishing whether a rational generating function is $\mathbb{N}$-rational is decidable.

Soittola's Theorem has recently been proved in different ways in [19, 62], using different approaches and some algorithms, to pass from an $\mathbb{N}$-rational series to a regular expression enumerated by such a series, have been proposed [9, 47]. However, none of these techniques provides a method to face C-finite recurrence relations which are not $\mathbb{N}$-rational.

Following the attempt of enlightening some questions on positive sequences, some researches have recently focused on determining sufficient conditions to establish the possible positivity of a given C-finite recurrence relation, as interestingly described in [38]. As a matter of fact, up to now, we only know that the positivity problem is decidable for C-finite recurrences of two [42] or three terms [49]. Another possible approach to tackle the positivity problem is to develop algorithms to test positivity of recursively defined sequences (and, in particular, C-finite sequences) by means computer algebra, as in [39].

Our work relates to the positivity problem, since we propose a sufficient condition for testing the positivity of a C-finite sequence.

If the recurrence relation has degree k with coefficients $a_1, \dots, a_k$, such a condition can be expressed in terms of a set of k inequalities which can be obtained from a set of quotients and remainders involving by the coefficients. To our knowledge, such a condition is completely new in literature coming from a completely different approach to tackle the problem.

A brief summary.

In **Chapter 1** we define the basic tools and objects used and studied in this work.

In **Chapter 2** we study particular strings over the alphabet $\{0, 1\}$ which avoid a single pattern. Firstly, we propose a general algorithm constructing such words and secondly we solve the enumeration problem by means of generating functions.

In **Chapter 3** we characterize a particular set of binary words having the property that no prefix of any word is a suffix of any other word. Working which such kind of sets is required to further develop the studies of pattern avoidance.

In **Chapter 4** we expand the results of Chapter 2 using the ones of Chapter 3 passing from a single to a set of forbidden patterns.

In **Chapter 5** we apply the tool of succession rule in a different context from the previous ones furnishing a sufficient condition to determine if an integer sequence, arise from a recurrence relation with integer coefficients, are all positive.

In **Chapter 6** we summarize the main results of this thesis and list some open problems in the context both of pattern avoidance and positivity problem.

Related publications by the author

Many of the results described in this thesis also appear in the following articles by the author:

1. S. Bilotta, D. Merlini, E. Pergola, R. Pinzani. Pattern $1^{j+1}0^j$ avoiding binary words. *Fundamenta Informaticae* 177 (1-4), 35–55, (2012).

2. S. Bilotta, E. Pergola, R. Pinzani. A new approach to cross-bifix-free sets. *IEEE Transactions on Information Theory* 58 (6), 4058–4063, (2012).

3. S. Bilotta, F. Disanto, R. Pinzani, S. Rinaldi. Catalan structures and Catalan pairs. *Theoretical Computer Science*, (2012). In Press.

4. S. Bilotta, E. Pergola, R. Pinzani. A construction for a class of binary words avoiding 1^j0^i. *Pure Mathematics and Applications*, (2012). In Press.

5. S. Bilotta, E. Grazzini, E. Pergola, R. Pinzani. Avoiding cross-bifix-free binary words. Submitted to *ACTA Informatica*, (2012).

6. S. Bilotta, E. Pergola, R. Pinzani, S. Rinaldi. Recurrence relations versus succession rules. Submitted to *International Journal of Foundations of Computer Science*, (2012).

7. S. Bilotta, E. Grazzini, E. Pergola. Counting binary words avoiding alternating patterns. Submitted to *Journal of Integer Sequences*, (2012).

8. S. Bilotta, E. Grazzini, E. Pergola. Primitive Dyck path as an invariant class for the enumeration. *Proceedings of the International Conference GASCom 2012*, Bordeaux, (2012).

9. S. Bilotta, E. Pergola, R. Pinzani. Pattern 1^j0^i avoiding binary words. *WORDS 2011, Electronic Proceedings in Theoretical Computer Science* 63, 53–64, (2011).

10. S. Bilotta, F. Disanto, R. Pinzani, S. Rinaldi. Catalan structures and Catalan pairs. *Proceedings of the International Conference GASCom 2010*, Montreal, (2010).

11. S. Bilotta, D. Merlini, E. Pergola, R. Pinzani. Binary words avoiding a pattern and marked succession rule. *Proceedings of the International Conference Lattice Path Combinatorics and Applications 2010*, 45–49, Siena, (2010).

1 Basic definitions

In this chapter we give some basic definitions and notations related to succession rules and pattern avoidance which are used in this thesis.

1.1 Succession rules

The concept of a succession rule was introduced in [22] by Chung, Graham, Hoggat and Kleiman to study reduced Baxter permutations. Later it was applied to the enumeration of permutations with forbidden subsequences [28, 29, 81]. Recently this technique has been successfully applied to other combinatorial objects [10, 12]. In all these cases there is a common approach to the examined enumeration problem: a *generating tree* is associated to certain combinatorial class, according to some enumerative parameters, in such a way that the number of nodes appearing on level n of the tree gives the number of n-sized objects in the class.

In [81] West discusses how generating trees can be used to some enumerate classes of permutations with forbidden subsequences and argues for a broader use of generating trees in combinatorial enumeration. Only later this has been recognized as an extremely useful tool for the ECO method [12].

In the next we give some basic definitions and notations related to the concept of succession rule.

1.1.1 Ordinary succession rules

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

$$\{(k) \overset{1}{\rightsquigarrow} (e_1(k))(e_2(k))\dots(e_k(k))\}_{k\in\Sigma} \;\; e_i(k) \in \Sigma,\ 1 \leq i \leq k.$$

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

$$\Omega : \begin{cases} (a) & \\ (k) & \overset{1}{\rightsquigarrow} (e_1(k))(e_2(k))\dots(e_k(k)) \end{cases} \tag{1.1}$$

In this context Σ is called the set of *labels*. The rule Ω can be represented by means of a *generating tree*, that is an infinite tree with the root labelled by (a). Each node labelled by (k) has k successors labelled $(e_1(k)), (e_2(k)), \dots, (e_k(k))$, respectively. By convention the root of the tree is at level 0, and a node lies at level n if its parent lies at level $n-1$. A tree is a generating tree for a class of combinatorial objects if there exists a bijection between the objects of size n and the nodes at level n in the tree, and in such a case a given object can be coded by the sequence of labels met from the root of the generating tree to the object itself.

For instance, the succession rule for *Catalan numbers* (that is: $1, 2, 5, 14, 132, 429, 1430, 4862, 16796, \dots$ sequence A000108 in the The On-Line Encyclopedia of Integer Sequences [71]) which describe the growth of several combinatorial objects, first of all the *well-parenthesized expressions* [73, 74], is:

$$\begin{cases} (2) & \\ (k) & \overset{1}{\rightsquigarrow} (2)(3)\dots(k)(k+1) \end{cases} \tag{1.2}$$

and some levels of its generating tree are shown in Figure 1.1. By applying the succession rule (1.2) to the well-parenthesized expressions we have the generating tree in Figure 1.2. For example, the word ()(())() is encoded by (2)(2)(3)(2). We refer to [2, 10, 12, 15] for further details and examples on succession rules.

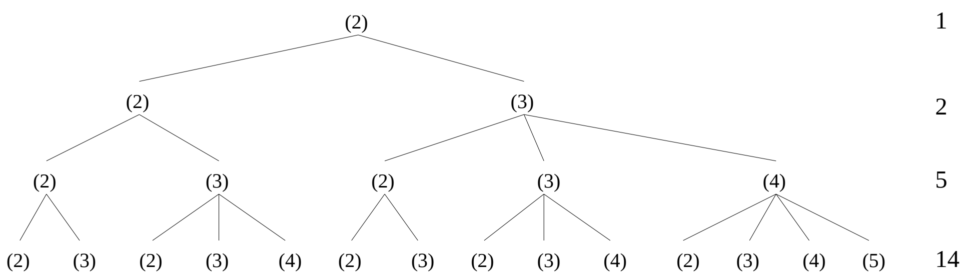


Figure 1.1: Some levels of the generating tree associated to the succession rule (1.2)

We remark that, from the above definition, a node labelled (k) has precisely k successors, this kind of rules are often called *consistent* succession rules. However, we can also consider succession rules, introduced in [25], in

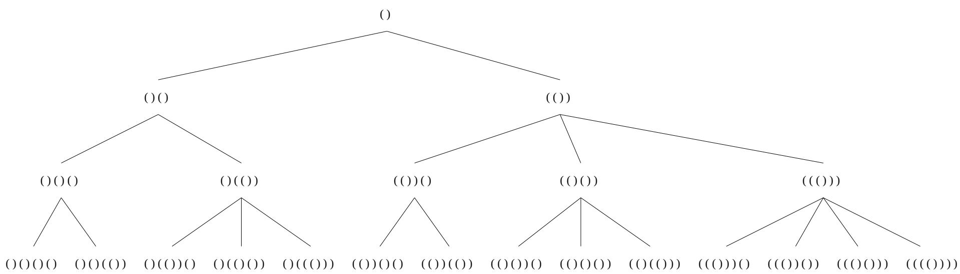


Figure 1.2: Some levels of the generating tree associated to the growth of the well-parenthesized expressions

which the value of a label does not necessarily represent the number of its successors. As an instance see the succession rule (1.3) for Catalan numbers whose generating tree is shown in Figure 1.3.

$$\left\{ \begin{array}{ll} (0) & \\ (k) & \stackrel{1}{\rightsquigarrow} (0)(1)\dots(k)(k+1) \end{array} \right. \tag{1.3}$$

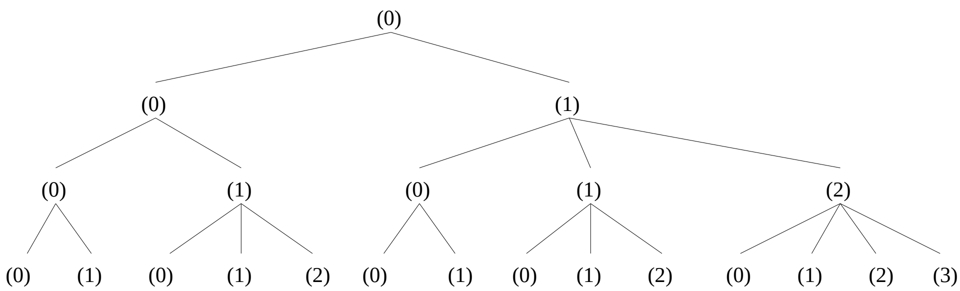


Figure 1.3: Some levels of the generating tree associated to the succession rule (1.3)

A succession rule Ω defines a sequence of positive integers $\{f_n\}_{n\geq 0}$ where f_n is the number of the nodes at level n in the generating tree defined by Ω. As the root is at level 0, so $f_0 = 1$. The function $f_\Omega(x) = \sum_{n\geq 0} f_n x^n$ is the *generating function* determined by Ω.

Two succession rules are *equivalent* if they have the same generating function. A succession rule is *finite* if it has a finite number of labels and productions.

For example, the two succession rules:

$$\left\{ \begin{array}{l} (2) \\ (2) \rightsquigarrow (2)(2) \end{array} \right. \qquad \left\{ \begin{array}{l} (2) \\ (k) \rightsquigarrow (1)^{k-1}(k+1) \end{array} \right.$$

are equivalent rules, and define the sequence $f_n = 2^n$. The one on the left is a finite rule, since it uses only the label (2), while the one on the right is a rule which generates an infinite number of labels.

1.1.2 Jumping succession rules

Succession rules such as (1.1) or (1.3) 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 productions acting on the objects of a class and producing successors at different levels.

To indicate a jumping succession rule Γ is used the following notation:

$$\Gamma : \begin{cases} (a) & \\ (k) & \overset{1}{\rightsquigarrow} (e_1(k))(e_2(k))\dots(e_k(k)) \\ (k) & \overset{j}{\rightsquigarrow} (d_1(k))(d_2(k))\dots(d_k(k)) \end{cases} \tag{1.4}$$

The generating tree associated to Γ has the property that each node labelled by (k) and lying at level n produces two sets of successors, the first set at level $n+1$ and having labels $(e_1(k)), (e_2(k)), \dots, (e_k(k))$, respectively, and the second one at level $n+j$, with $j > 1$, and having labels $(d_1(k)), (d_2(k)), \dots, (d_k(k))$, respectively.

For example, the jumping succession rule (1.5) counts the number of *2-generalized Motzkin paths* (see [76] for further details) and Figure 1.4 shows some levels of the associated generating tree. For more details about these topics, see [31].

$$\begin{cases} (1) & \\ (k) & \overset{1}{\rightsquigarrow} (1)(2)\cdots(k-1)(k+1) \\ (k) & \overset{2}{\rightsquigarrow} (k) \end{cases} \tag{1.5}$$

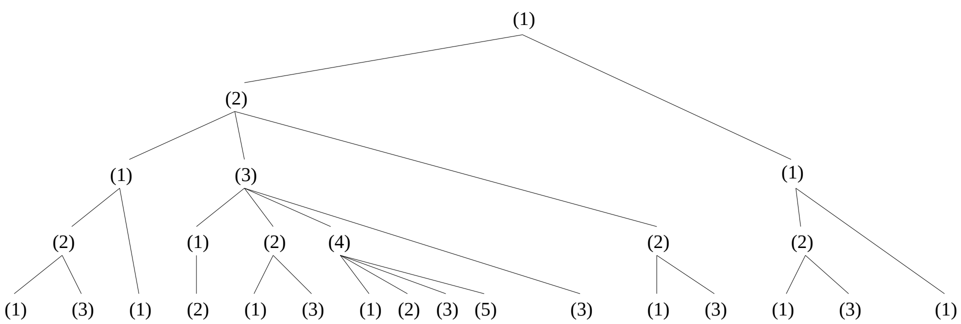


Figure 1.4: Some levels of the generating tree associated to the succession rule (1.5)

Of course the most general expression for a jumping succession rule can have more than one jump as in (1.6).

$$(1.6)\qquad \begin{cases} (a) & \\ (k) & \overset{j_1}{\rightsquigarrow} (e_{11}(k))(e_{12}(k))\dots(e_{1k}(k)) \\ (k) & \overset{j_2}{\rightsquigarrow} (e_{21}(k))(e_{22}(k))\dots(e_{2k}(k)) \\ \vdots & \\ (k) & \overset{j_m}{\rightsquigarrow} (e_{m1}(k))(e_{m2}(k))\dots(e_{mk}(k)) \end{cases}$$

Moreover, by leaving the notion of consistency for a succession rule we can have jumping succession rules of the form in (1.7).

$$(1.7)\qquad \begin{cases} (a) & \\ (k) & \overset{j_1}{\rightsquigarrow} (e_{11}(k))(e_{12}(k))\dots(e_{1k_1}(k)) \\ (k) & \overset{j_2}{\rightsquigarrow} (e_{21}(k))(e_{22}(k))\dots(e_{2k_2}(k)) \\ \vdots & \\ (k) & \overset{j_m}{\rightsquigarrow} (e_{m1}(k))(e_{m2}(k))\dots(e_{mk_m}(k)) \end{cases}$$

1.1.3 Marked succession rules

Another generalization is introduced in [55], where the authors deal with *marked succession rules*. In this case the labels appearing in a succession rule can be marked, and the *marked* labels are considered together with the unmarked labels.

A *marked generating tree* is a rooted labelled tree where there appear marked and unmarked labels according to the corresponding succession rule. The main property is that in the generating tree a marked label $(\overline{k})$ kills or annihilates the unmarked label (k) lying on the same level n. In particular, the enumeration of the combinatorial objects in a class is the difference between the number of unmarked and marked labels lying on a given level.

Note that, a compact notation (1.8) for a marked succession rule:

$$(1.8)\qquad \begin{cases} (a) & \\ (k) & \overset{1}{\rightsquigarrow} (e_1(k))(e_2(k))\dots(e_k(k)) \\ (k) & \overset{j}{\rightsquigarrow} (\overline{d_1(k)})(\overline{d_2(k)})\dots(\overline{d_k(k)}) \end{cases}$$

describes also the behavior for $(\overline{k})$, that is (1.9), and for any label (k), we have $(\overline{\overline{k}}) = (k)$.

$$(1.9)\qquad \begin{cases} (\overline{k}) & \overset{1}{\rightsquigarrow} (\overline{e_1(k)})(\overline{e_2(k)})\dots(\overline{e_k(k)}) \\ (\overline{k}) & \overset{j}{\rightsquigarrow} (d_1(k))(d_2(k))\dots(d_k(k)) \end{cases}$$

Of course a marked succession rule can also be not consistent. By the way, each succession rule (1.1) can be trivially rewritten as (1.10).

$$(1.10)\qquad \begin{cases} (a) & \\ (k) & \overset{1}{\rightsquigarrow} (e_1(k))(e_2(k))\dots(e_k(k))(k) \\ (k) & \overset{1}{\rightsquigarrow} (\overline{k}) \end{cases}$$

For example, the succession rule for Catalan numbers can be rewritten in the form (1.11) and Figure 1.5 shows some levels of the associated generating tree.

$$(1.11)\qquad \begin{cases} (2) & \\ (k) & \overset{1}{\rightsquigarrow} (2)(3)\dots(k)(k+1)(k) \\ (k) & \overset{1}{\rightsquigarrow} (\overline{k}) \end{cases}$$

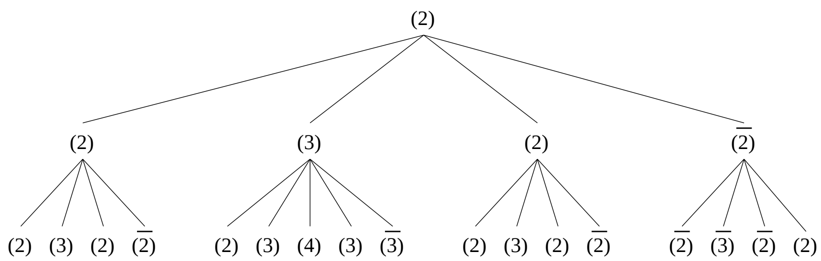


Figure 1.5: Three levels of the generating tree associated to the succession rule (1.11)

The various concepts previously described can be mixed together giving rise to *jumping and marked succession rules.*

1.2 Pattern avoidance

Let B be a finite, non-empty set called *alphabet*. The elements of B are called *letters*. A (finite) sequence of letters in B is called (finite) *word*. Let B^* denote the monoid of all finite words over B where ε denotes the *empty word* and $B^+ = B^* \backslash \varepsilon$. Let ω be a word in B^*, then $|\omega|$ indicates the length of ω and $|\omega|_a$ denotes the number of occurrences of a in ω, being $a \in B$. A word $\omega = \omega_0\omega_1\dots\omega_{n-1}$ of length $|\omega| = n$ contains a *pattern* $\mathfrak{p} = p_0p_1\dots p_{h-1} \in B^*$ of length $|\mathfrak{p}| = h$, with $h \leq n$, if there is an index i such that $\omega_i\omega_{i+1}\dots\omega_{i+h-1} = p_0p_1\dots p_{h-1}$. Otherwise, we say that ω *avoids* the pattern $\mathfrak{p}$, or that $\mathfrak{p}$ is a *forbidden* pattern for ω.

1.2.1 Pattern avoidance in binary words

Let $B = \{0,1\}^*$ be the set of binary words, the subclass $B^{[\mathfrak{p}]}$ of B denotes the set of binary words excluding a given pattern $\mathfrak{p} \in \{0,1\}^*$, i.e. each binary word $\omega \in B^{[\mathfrak{p}]}$ avoids $\mathfrak{p}$.

In particular, if $B^{[\mathfrak{p}]}_{n,k}$ denotes the number of words excluding the pattern $\mathfrak{p}$ and having n bits 1 and k bits 0, then by using the results in [3] we have:

$$B^{[\mathfrak{p}]}(x,y) = \sum_{n,k\geq 0} B^{[\mathfrak{p}]}_{n,k} x^n y^k = \frac{C^{[\mathfrak{p}]}(x,y)}{(1-x-y)C^{[\mathfrak{p}]}(x,y) + x^{|\mathfrak{p}|_1} y^{|\mathfrak{p}|_0}}, \tag{1.12}$$

where $|\mathfrak{p}|_1$ and $|\mathfrak{p}|_0$ correspond to the number of ones and zeroes in the pattern and $C^{[\mathfrak{p}]}(x,y)$ is the autocorrelation polynomial with coefficients given by the autocorrelation vector (see also [40, 41, 69]). For a given $\mathfrak{p}$, this vector of bits $c = (c_0, \ldots, c_{h-1})$ can be defined in terms of Iverson's bracket notation (for a predicate P, the expression $[\![P]\!]$ has value 1 if P is true and 0 otherwise) as follows: $c_i = [\![p_0 p_1 \cdots p_{h-1-i} = p_i p_{i+1} \cdots p_{h-1}]\!]$. In other words, the bit c_i is determined by shifting $\mathfrak{p}$ right by i positions and setting $c_i = 1$ if and only if the remaining letters match the original. For example, when $\mathfrak{p} = 101010$ the autocorrelation vector is $c = (1,0,1,0,1,0)$, as illustrated in Table 1.1, and $C^{[\mathfrak{p}]}(x,y) = 1 + xy + x^2y^2$, that is, we mark with $x^j y^i$ the tails of the pattern with j bits 1, i bits 0 and $c_{j+i} = 1$. Therefore, in this case we have:

$$B^{[\mathfrak{p}]}(x,y) = \frac{1 + xy + x^2y^2}{(1-x-y)(1+xy+x^2y^2) + x^3y^3}.$$

1	0	1	0	1	0	Tails					
1	0	1	0	1	0						1
	1	0	1	0	1	0					0
		1	0	1	0	1	0				1
			1	0	1	0	1	0			0
				1	0	1	0	1	0		1
					1	0	1	0	1	0	0

Table 1.1: The autocorrelation vector for $\mathfrak{p} = 101010$

As another example, when $\mathfrak{p} = 11100$ then $C^{[\mathfrak{p}]}(x,y) = 1$ and

$$B^{[\mathfrak{p}]}(x,y) = 1/(1-x-y+x^3y^2).$$

In this thesis, we are interested in studying the class $F^{[\mathfrak{p}]} \subset \{0,1\}^*$ of binary words such that $|\omega|_0 \leq |\omega|_1$, for all $\omega \in F^{[\mathfrak{p}]}$.

Note that, each binary word ω 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 ω, and a *fall step*, defined by $(1, -1)$ and denoted by $\overline{x}$, to each 0's in ω. From now on, we refer interchangeably to words or their graphical representations on the Cartesian plane, that is paths.

For example, the geometrical representation of $\mathfrak{p} = 11100$ is illustrate in Figure 1.6.

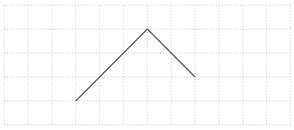


Figure 1.6: geometrical representation of $\mathfrak{p} = 11100$

2

Single pattern avoidance

In this chapter we study the enumeration and the construction of the class F of binary words in $\{0,1\}^*$ having the number of 1's greater than or equal to the number of 0's, and avoiding a *single* fixed pattern $\mathfrak{p}$. Firstly, we consider the forbidden pattern $\mathfrak{p}$ as the pattern $\mathfrak{p}(j) = 1^{j+1}0^j$, for any fixed $j \geq 1$. Secondly, 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 $\mathfrak{p}(j,i) = 1^j0^i$, $0 < i \leq j$.

2.1 A note on the class $F^{[\mathfrak{p}]}$

In this section we introduce some technics for the enumeration and the construction of the class $F^{[\mathfrak{p}]}$, according to the numbers of ones, with $\mathfrak{p} = \mathfrak{p}(j) = 1^{j+1}0^j$, for any fixed $j \geq 1$.

The first used tool is a context-free grammar for the class $F^{[\mathfrak{p}(j)]}$. We exemplify this approach to the class $F^{[\mathfrak{p}(1)]}$ where $\mathfrak{p}(1) = 110$, from which the generating function can be obtained.

We note that for each parameter j an "ad hoc" context-free grammar depending on the shape of the forbidden pattern must be defined.

A way for a more unified approach, which does not depend on the shape of the forbidden pattern $\mathfrak{p}(j)$, is initially obtained by means of the theory of Riordan arrays solving the enumeration problem and giving a *jumping and marked succession rule* which describes the growth of such words according to their number of ones.

2.1.1 A context-free grammar for $F^{[\mathfrak{p}]}$

A context-free grammar can be used to construct the class $F^{[\mathfrak{p}(j)]}$. In this case the shape of the forbidden pattern $\mathfrak{p}(j)$ must be taken into consideration, therefore we furnish the grammar for $F^{[\mathfrak{p}(1)]}$, with $\mathfrak{p}(j) = \mathfrak{p}(1) = 110$, that is $G = \{\{0,1\},\{F,A,B,C,D,E\},P,F\}$ where P is the set:

$$P : \begin{cases} F & \to BCA|BCDA \\ A & \to 1A|\varepsilon \\ B & \to 10B|\varepsilon \\ C & \to 01C|\varepsilon \\ D & \to 00E11 \\ E & \to B0E1|B \end{cases}$$

For example, the word $\omega = 001011 \in F^{[\mathfrak{p}(1)]}$ is generated by the following derivation:

$$F \to BCDA \to CDA \to DA \to 00E11A \to$$

$$\to 00B11A \to 0010B11A \to 001011A \to 001011.$$

This kind of grammar does not allow to the forbidden pattern 110 to be created and, at the same time, it does not allow to the number of 1's to be less than the number of 0's. In order to do that different subcases must be dealt which are managed by the variables appearing in the grammar.

By applying the Schützenberger's methodology [68] it is possible to translate the productions of the grammar into a system of functional equations whose solution is the generating function of the language $F^{[\mathfrak{p}(1)]}$ according to the number of ones:

$F(x) = B(x)C(x)A(x) + B(x)C(x)D(x)A(x) = B(x)C(x)A(x)(1 + D(x))$

$A(x) = 1 + xA(x)$ then $A(x) = \frac{1}{1-x}$.

$B(x) = 1 + xB(x)$ then $B(x) = \frac{1}{1-x}$.

$C(x) = 1 + xC(x)$ then $C(x) = \frac{1}{1-x}$.

$D(x) = x^2E(x)$.

$E(x) = xB(x)E(x) + B(x)$ then $E(x) = \frac{B(x)}{1-xB(x)}$.

Therefore, the generating function $F(x)$ according to the number of 1 is:

$$F(x) = \frac{1}{(1-x)^3}(1 + \frac{x^2}{1-2x}) = \frac{1}{(1-x)(1-2x)} \tag{2.1}$$

meaning that the number of words containing exactly n occurrences of 1's is $2^{n+1} - 1$.

We note that, the simplest class $F^{[\mathfrak{p}(1)]}$ with $\mathfrak{p}(j) = \mathfrak{p}(1) = 110$ is defined by a nontrivial context-free grammar. A general approach by means of context-free grammars is rather cumbersome as a different number of productions must be added strictly depending on the parameter j.

2.1.2 Riordan array and the class $F^{[\mathfrak{p}]}$

In this section, we establish necessary and sufficient conditions for the number of words counted according to the number of zeroes and ones to be related to proper Riordan arrays.

This problem is interesting in the context of the Riordan arrays theory because the matrices arising there are naturally defined by recurrence relations following the characterization given in [56] (see formula (2.3) below).

In order to study the binary words avoiding a pattern in terms of Riordan arrays, we consider the array $\mathcal{F}^{[\mathfrak{p}]} = (F^{[\mathfrak{p}]}_{n,k})$ given by the lower triangular part of the array $\mathcal{B}^{[\mathfrak{p}]} = (B^{[\mathfrak{p}]}_{n,k})$ by means of (1.12) in Section 1.2, that is, $F^{[\mathfrak{p}]}_{n,k} = B^{[\mathfrak{p}]}_{n,n-k}$ with $k \leq n$. More precisely, $F^{[\mathfrak{p}]}_{n,k}$ counts the number of words avoiding $\mathfrak{p}$ and having length $2n-k$, n bits one and $n-k$ bits zero. Given a pattern $\mathfrak{p} = p_0 \dots p_{h-1} \in \{0,1\}^h$, let $\bar{\mathfrak{p}} = \bar{p}_0 \dots \bar{p}_{h-1}$ be the pattern with $\bar{p}_i = 1 - p_i, \forall i = 0, \cdots, h-1$.

We obviously have $F^{[\bar{\mathfrak{p}}]}_{n,k} = B^{[\bar{\mathfrak{p}}]}_{n,n-k} = B^{[\mathfrak{p}]}_{n-k,n}$, therefore, the matrices $\mathcal{F}^{[\mathfrak{p}]}$ and $\mathcal{F}^{[\bar{\mathfrak{p}}]}$ represent the lower and upper triangular part of the array $\mathcal{B}^{[\mathfrak{p}]}$, respectively.

Moreover, we have $F^{[\mathfrak{p}]}_{n,0} = F^{[\bar{\mathfrak{p}}]}_{n,0} = B^{[\mathfrak{p}]}_{n,n}$, $\forall n \in \mathbb{N}$, that is, columns zero of $\mathcal{F}^{[\mathfrak{p}]}$ and $\mathcal{F}^{[\bar{\mathfrak{p}}]}$ correspond to the main diagonal of $\mathcal{B}^{[\mathfrak{p}]}$.

Tables 2.1, 2.2 and 2.3 illustrate some rows for the matrices $\mathcal{B}^{[\mathfrak{p}]}$, $\mathcal{F}^{[\mathfrak{p}]}$ and $\mathcal{F}^{[\bar{\mathfrak{p}}]}$ when $\mathfrak{p} = 11100$.

n/k	0	1	2	3	4	5	6	7
0	1	1	1	1	1	1	1	1
1	1	2	3	4	5	6	7	8
2	1	3	6	10	15	21	28	36
3	1	4	9	18	32	52	79	114
4	1	5	13	29	58	106	180	288
5	1	6	18	44	96	192	357	624
6	1	7	24	64	151	325	650	1222
7	1	8	31	90	228	524	1116	2232

Table 2.1: The matrix $\mathcal{B}^{[\mathfrak{p}]}$ for $\mathfrak{p} = 11100$

We briefly recall that a Riordan array is an infinite lower triangular array $(d_{n,k})_{n,k\in\mathbb{N}}$, defined by a pair of formal power series $(d(t), h(t))$, such

n/k	0	1	2	3	4	5	6	7
0	1							
1	2	1						
2	6	3	1					
3	18	9	4	1				
4	58	29	13	5	1			
5	192	96	44	18	6	1		
6	650	325	151	64	24	7	1	
7	2232	1116	524	228	90	31	8	1

Table 2.2: The triangle $\mathcal{F}^{[\mathfrak{p}]}$ for $\mathfrak{p} = 11100$

n/k	0	1	2	3	4	5	6	7
0	1							
1	2	1						
2	6	3	1					
3	18	10	4	1				
4	58	32	15	5	1			
5	192	106	52	21	6	1		
6	650	357	180	79	28	7	1	
7	2232	1222	624	288	114	36	8	1

Table 2.3: The triangle $\mathcal{F}^{[\bar{\mathfrak{p}}]}$ for $\bar{\mathfrak{p}} = 00011$

that $d(0) \neq 0, h(0) = 0, h'(0) \neq 0$ and the generic element $d_{n,k}$ is the n-th coefficient in the series $d(t)h(t)^k$, i.e.:

$$d_{n,k} = [t^n]d(t)h(t)^k, \qquad n, k \geq 0.$$

From this definition we have $d_{n,k} = 0$ for $k > n$. An alternative definition is in terms of the so-called A-sequence and Z-sequence, with generating functions $A(t)$ and $Z(t)$ satisfying the relations:

$$h(t) = tA(h(t)), \quad d(t) = \frac{d_0}{1 - tZ(h(t))} \quad \text{with} \quad d_0 = d(0).$$

In other words, Riordan arrays correspond to matrices where each element $d_{n,k}$ is described by a linear combination of the elements in the previous row, starting from the previous column, with coefficients in A:

$$d_{n+1,k+1} = a_0 d_{n,k} + a_1 d_{n,k+1} + a_2 d_{n,k+2} + \cdots \tag{2.2}$$

Another characterization (see [56]) states that a lower triangular array $(d_{n,k})_{n,k\in\mathbb{N}}$ is Riordan if and only if there exists another array $(\alpha_{i,j})_{i,j\in\mathbb{N}}$, with $\alpha_{0,0} \neq 0$, and a sequence $(\rho_j)_{j\in\mathbb{N}}$ such that:

$$d_{n+1,k+1} = \sum_{i\geq 0}\sum_{j\geq 0}\alpha_{i,j}d_{n-i,k+j} + \sum_{j\geq 0}\rho_j d_{n+1,k+j+2}. \tag{2.3}$$

Matrix $(\alpha_{i,j})_{i,j\in\mathbb{N}}$ is called the A-matrix of the Riordan array.
If $P^{[0]}(t), P^{[1]}(t), P^{[2]}(t), \ldots$ denote the generating functions of rows $0, 1, 2, \ldots$ in the A-matrix, i.e.:

$$P^{[i]}(t) = \alpha_{i,0} + \alpha_{i,1}t + \alpha_{i,2}t^2 + \alpha_{i,3}t^3 + \ldots$$

and $Q(t)$ is the generating function for the sequence $(\rho_j)_{j\in\mathbb{N}}$, then we have:

$$\frac{h(t)}{t} = \sum_{i\geq 0} t^i P^{[i]}(h(t)) + \frac{h(t)^2}{t}Q(h(t)), \tag{2.4}$$

$$A(t) = \sum_{i\geq 0} t^i A(t)^{-i}P^{[i]}(t) + tA(t)Q(t). \tag{2.5}$$

The theory of Riordan arrays and the proofs of their properties can be found in [56]. The Riordan arrays which arise in the context of pattern avoidance (see [3, 54]) have the nice property to be defined by a quite simple recurrence relation following the characterization (2.3), while the relation induced by the A-sequence is, in general, more complex. From a combinatorial point of view, this means that it is very challenging to find a construction allowing to build objects of size $n+1$ from objects of size n. Instead, the existence of a simple A-matrix corresponds to a possible construction from objects of different sizes less than $n+1$. On the other hand, as we will see in the sequel, the recurrence following characterization (2.3) contains negative coefficients and therefore gives rise to interesting non trivial combinatorial problems.

In this section we examine in particular the family of patterns $\mathfrak{p} = \mathfrak{p}(j) = 1^{j+1}0^j$ and show that the corresponding recurrence relation can be combinatorially interpreted. To this purpose, we translate the recurrence into a succession rule, as it is typically done from problems related to Riordan arrays (see, e.g., [57]), and give a construction for the class of binary words avoiding the pattern $\mathfrak{p}$.

The Riordan array for the pattern $\mathfrak{p}(j) = 1^{j+1}0^j$

Let us consider the family of patterns $\mathfrak{p} = \mathfrak{p}(j) = 1^{j+1}0^j$ and let $B^{[\mathfrak{p}(j)]}_{n,k}$ denote the number of words excluding the pattern and having n bits 1 and k bits 0; from (1.12) we have

$$B^{[\mathfrak{p}(j)]}(x,y) = \sum_{n,k\geq 0} B^{[\mathfrak{p}(j)]}_{n,k}x^n y^k = \frac{1}{1-x-y+x^{j+1}y^j}. \tag{2.6}$$

Now, let $F_{n,k}^{[\mathfrak{p}(j)]}$ count the number of words avoiding $\mathfrak{p}(j)$ and having n bits one and $n-k$ bits zero. Obviously we have $F_{n,k}^{[\mathfrak{p}(j)]} = B_{n,n-k}^{[\mathfrak{p}(j)]}$ with $k \leq n$. By extracting the coefficients from (2.6) we have:

$$[x^{n+1}y^{k+1}](1-x-y+x^{j+1}y^j)B^{[\mathfrak{p}(j)]}(x,y) =$$

$$= B_{n+1,k+1}^{[\mathfrak{p}(j)]} - B_{n,k+1}^{[\mathfrak{p}(j)]} - B_{n+1,k}^{[\mathfrak{p}(j)]} + B_{n-j,k+1-j}^{[\mathfrak{p}(j)]} = 0$$

and therefore:

$$F_{n+1,k+1}^{[\mathfrak{p}(j)]} = F_{n,k}^{[\mathfrak{p}(j)]} + F_{n+1,k+2}^{[\mathfrak{p}(j)]} - F_{n-j,k}^{[\mathfrak{p}(j)]}. \tag{2.7}$$

This is a recurrence relation of type (2.3) and therefore $\mathcal{F}^{[\mathfrak{p}(j)]} = (F_{n,k}^{[\mathfrak{p}(j)]})$ is a Riordan array. In particular, the coefficients of the relation correspond to $P^{[j]}(t) = -1$, $P^{[0]}(t) = 1$, and $Q(t) = 1$, therefore we have

$$\frac{h^{[\mathfrak{p}(j)]}(t)}{t} = \sum_{i\geq 0} t^i P^{[i]}(h^{[\mathfrak{p}(j)]}(t)) + \frac{h^{[\mathfrak{p}(j)]}(t)^2}{t} Q(h(t)) = 1 - t^j + \frac{h^{[\mathfrak{p}(j)]}(t)^2}{t}$$

that is,

$$h^{[\mathfrak{p}(j)]}(t)^2 - h^{[\mathfrak{p}(j)]}(t) + t - t^{j+1} = 0, \quad h^{[\mathfrak{p}(j)]}(t) = \frac{1-\sqrt{1-4t+4t^{j+1}}}{2}.$$

We explicitly observe that from formula (2.5) the generating function $A(t)$ of the A-sequence is the solution of a $j+1$ degree equation $(1-t)A(t)^{j+1} - A(t)^j + t^j = 0$. For example, when $\mathfrak{p}(j) = \mathfrak{p}(2) = 11100$ by developing into series we find:

$$A(t) = 1 + t + 2t^3 - t^4 + 7t^5 - 12t^6 + 38t^7 - 99t^8 + 281t^9 + O(t^{10})$$

and this result excludes that there might exist a simple dependence of the elements in row $n+1$ from the elements in row n. For what concerns $d^{[\mathfrak{p}(j)]}(t)$, we simply use the Cauchy formula for finding the main diagonal of matrix $\mathcal{B}^{[\mathfrak{p}(j)]}$ (see, e.g., [73, Cap. 6, p. 182]):

$$d^{[\mathfrak{p}(j)]}(t) = [x^0]B^{[\mathfrak{p}(j)]}(x, \frac{t}{x}) = \frac{1}{2\pi i}\oint B^{[\mathfrak{p}(j)]}(x, \frac{t}{x})\frac{dx}{x}.$$

We have:

$$\frac{1}{x}B^{[\mathfrak{p}(j)]}(x, \frac{t}{x}) = \frac{-1}{x^2(1-t^j) - x + t}$$

and in order to compute the integral, it is necessary to find the singularities $x(t)$ such that $x(t) \to 0$ with $t \to 0$ and apply the Residue theorem. In this case the right singularity is:

$$x(t) = \frac{1-\sqrt{1-4t(1-t^j)}}{2(1-t^j)}$$

and finally we have:

$$d^{[\mathfrak{p}(j)]}(t) = \lim_{x \to x(t)} \frac{-1}{x^2(1-t^j) - x + t}(x - x(t)) = \frac{1}{\sqrt{1-4t+4t^{j+1}}}.$$

Observe also that:

$$\frac{d^{[\mathfrak{p}(j)]}(t) - 1}{d^{[\mathfrak{p}(j)]}(t) h^{[\mathfrak{p}(j)]}(t)} = 2$$

and therefore $F^{[\mathfrak{p}(j)]}_{n+1,0} = 2F^{[\mathfrak{p}(j)]}_{n+1,1}$. Recurrence (2.7) is quite simple, however, the presence of negative coefficients leads to a possible non trivial combinatorial interpretation. In order to study this problem we proceed as follows. The dependence of $F^{[\mathfrak{p}(j)]}_{n+1,k+1}$ from the same row $n+1$ can be simply eliminated and we have:

$$F^{[\mathfrak{p}(j)]}_{n+1,k+1} = F^{[\mathfrak{p}(j)]}_{n,k} - F^{[\mathfrak{p}(j)]}_{n-j,k} + F^{[\mathfrak{p}(j)]}_{n+1,k+2} =$$

$$= F^{[\mathfrak{p}(j)]}_{n,k} - F^{[\mathfrak{p}(j)]}_{n-j,k} + F^{[\mathfrak{p}(j)]}_{n,k+1} - F^{[\mathfrak{p}(j)]}_{n-j,k+1} + F^{[\mathfrak{p}(j)]}_{n+1,k+3} = \cdots =$$

$$\begin{aligned} &= (F^{[\mathfrak{p}(j)]}_{n,k} + F^{[\mathfrak{p}(j)]}_{n,k+1} + F^{[\mathfrak{p}(j)]}_{n,k+2} + \cdots) - \\ &-(F^{[\mathfrak{p}(j)]}_{n-j,k} + F^{[\mathfrak{p}(j)]}_{n-j,k+1} + F^{[\mathfrak{p}(j)]}_{n-j,k+2} + \cdots) \end{aligned} \tag{2.8}$$

Similarly we have:

$$\begin{aligned} F^{[\mathfrak{p}(j)]}_{n+1,0} &= 2(F^{[\mathfrak{p}(j)]}_{n,0} + F^{[\mathfrak{p}(j)]}_{n,1} + F^{[\mathfrak{p}(j)]}_{n,2} + \cdots) - \\ &-2(F^{[\mathfrak{p}(j)]}_{n-j,0} + F^{[\mathfrak{p}(j)]}_{n-j,1} + F^{[\mathfrak{p}(j)]}_{n-j,2} + \cdots) \end{aligned} \tag{2.9}$$

Finally, by using the results in [3], recurrences (2.8) and (2.9) translate into the following succession rule, for each $k \geq 0$:

$$\begin{cases} (0) & \\ (k) & \overset{1}{\rightsquigarrow} (k+1)(k)\cdots(1)(0_2)(0_1) \\ (k) & \overset{j+1}{\rightsquigarrow} (\overline{k+1})(\overline{k})\cdots(\overline{1})(\overline{0_2})(\overline{0_1}) \end{cases} \tag{2.10}$$

This rule can be represented as a tree having its root labelled (0) and where each node with label (k) at a given level n has $k+3$ sons at level $n+1$ labelled $(k+1), (k), \cdots, (1), (0_2), (0_1)$ and $k+3$ sons at level $n+j+1$ with labels $(\overline{k+1})(\overline{k})\cdots(\overline{1})(\overline{0_2})(\overline{0_1})$. If we denote by $d_{n,k}$ the number of nodes having label k at level n in the tree and count as negative the marked nodes then we obtain matrix $\mathcal{F}^{[\mathfrak{p}(j)]} = (F^{[\mathfrak{p}(j)]}_{n,k})_{n,k\in\mathbb{N}}$, that is, $\mathcal{F}^{[\mathfrak{p}(j)]}$ corresponds to the matrix associated to the rule (2.10). The relations between Riordan arrays and succession rules has been widely studied and we refer the reader to [55, 57] for more details.

2.2 A generating algorithm for $F^{[\mathfrak{p}(j)]}$

In this section we present a combinatorial interpretation of the jumping and marked succession rule (2.10) for the class $F^{[\mathfrak{p}(j)]}$, for any fixed $j \geq 1$. In particular, the problem of associating a word to a path in the generating tree obtained by the succession rule (2.10) is solved by introducing an algorithm which constructs all binary words in F and then kills those containing the forbidden pattern $\mathfrak{p}(j) = 1^{j+1}0^j$, for any fixed $j \geq 1$. Then, a generating function is given by using the ECO-method for the enumeration of combinatorial objects which admit recursive descriptions in terms of generating trees.

2.2.1 A construction for $F^{[\mathfrak{p}(j)]}$

In this section we define an algorithm which associates a lattice path in $F^{[\mathfrak{p}(j)]}$, where $\mathfrak{p}(j) = x^{j+1}\overline{x}^j = 1^{j+1}0^j$, to a sequence of labels obtained by means of the succession rule (2.10), where the subscripts of labels (0) are simply used in order to distinguish the two labels one from each other, since they are obtained in two different ways in the generating process. Note that the labels (0_1) and (0_2) have the same set of successors regardless their subscripts. This give a construction for the set $F^{[\mathfrak{p}(j)]}$ according to the number of rise steps or equivalently the number of ones.

The axiom (0) is associated to the empty path ε.
A path $\omega \in F$, with n rise steps and such that its endpoint has ordinate k, provides $k+3$ paths with $n+1$ rise steps, according to the first production of (2.10) having $k+1, k, \ldots, 1, 0, 0$ as endpoint ordinate, respectively. The first $k+2$ labels are obtained by adding to ω a sequence of steps consisting of one rise step followed by $k+1-h$, $0 \leq h \leq k+1$, fall steps (see Figure 2.1). Each 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 and having a rise step as last step are never obtained. These paths are bound to the label (0_1) of the first production in (2.10) and the way to obtain them will be described later in this section.

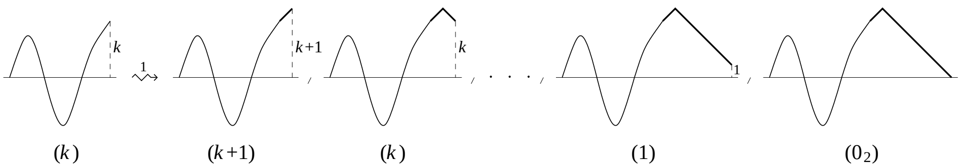


Figure 2.1: The mapping associated to $(k) \overset{1}{\rightsquigarrow} (k+1)(k)\ldots(1)(0_2)$ of (2.10)

We define a *marked forbidden pattern* $\mathfrak{p}(j)$ as a pattern $\mathfrak{p}(j) = x^{j+1}\overline{x}^j$

whose steps cannot be split, that is, they must always be contained all together in that defined sequence. We say that a point is strictly contained in a given marked forbidden pattern $\mathfrak{p}(j)$ if it is in $\mathfrak{p}(j)$ and it is different from both its initial point and its last point. We denote a marked forbidden pattern by marking its peak.

A cut operation, i.e the procedure which splits a given path into two subpaths, is not possible within a marked forbidden pattern $\mathfrak{p}(j)$. After a cut operation, it is not allowed to switch any rise step with a fall one, and viceversa, inside a marked forbidden pattern, but it can be translated.

A path $\omega \in F$, with n rise steps and such that its endpoint has ordinate k, provides $k+3$ paths, with $n+j+1$ rise steps, according to the second production of (2.10) having $k+1, k, \dots, 1, 0, 0$ as endpoint ordinate, respectively. The first $k+2$ labels are obtained by adding to ω a sequence of steps consisting of the marked forbidden pattern $\mathfrak{p}(j) = x^{j+1}\overline{x}^j$ followed by $k+1-h$, $0 \le h \le k+1$, fall steps (see Figure 2.2). Each 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. At this point the label (0_1) due to the first and the second production of (2.10) yield paths which either do not contain marked forbidden patterns in its rightmost suffix and end on the x-axis by a rise step or having the rightmost marked point with ordinate less than or equal to j.

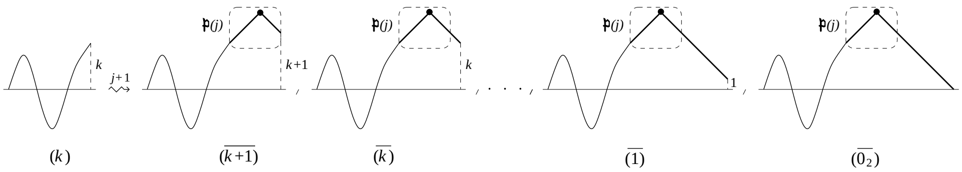


Figure 2.2: The mapping associated to $(k) \overset{j+1}{\rightsquigarrow} (\overline{k+1})(\overline{k})\dots(\overline{1})(\overline{0_2})$ of (2.10)

In order to obtain the label (0_1) according to the first production of (2.10), we consider the path ω' obtained from ω by adding a sequence of steps consisting of one rise step followed by k fall steps, while in order to obtain the label (0_1) according to the second production of (2.10), we consider the path ω' obtained from ω by adding a sequence of steps consisting of the marked forbidden pattern $\mathfrak{p}(j) = x^{j+1}\overline{x}^j$ 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_1) 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 2.3).

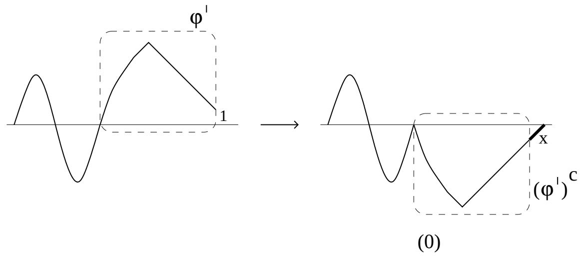


Figure 2.3: A graphical representation of the actions giving the label (0_1) in case of no marked points in φ'

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_1) 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, called α, and the path running from the initial point in φ' to z, called β (see Figure 2.4 and 2.5).

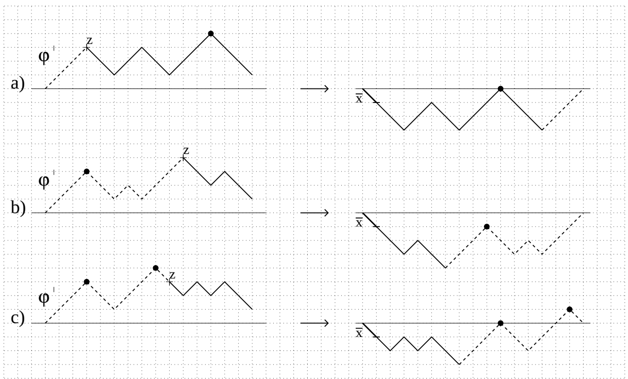


Figure 2.4: Some examples of the actions giving the label (0_1) in the case of marked points in φ', $\mathfrak{p}(j) = \mathfrak{p}(1) = x^2\overline{x}$

This last mapping can be inverted as follows. Let d be the rightmost fall step in a path ω^* labelled (0_1) beginning from the x-axis and such that each marked point, on its right, has ordinate less than or equal to j. Let us $\omega^* = v\varphi^*$, where φ^* is the rightmost suffix in ω^* beginning with d and let P 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 P to the endpoint of the path and α is the path φ^* running from the endpoint of d to P (see Figure 2.6). Figure 2.7 shows the cut and paste actions related to the inverted mapping with the pattern $\mathfrak{p}(j) = \mathfrak{p}(1) = x^2\overline{x}$.

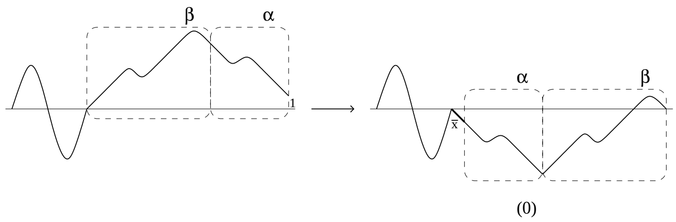


Figure 2.5: A graphical representation of the cut and paste actions giving the label (0_1) in case of marked points in φ'

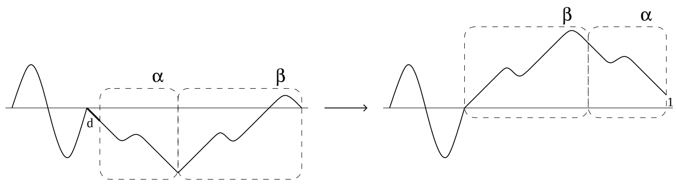


Figure 2.6: A graphical representation of the path obtained by means of the inverted mapping related to the label (0_1) in case of marked points in φ'

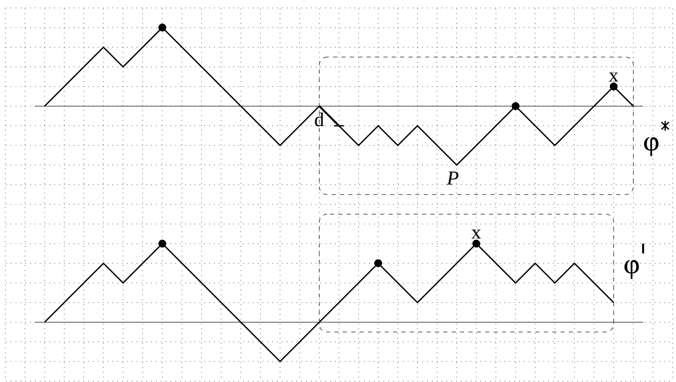


Figure 2.7: The inverted mapping related to the label (0_1) in case of marked points in φ'

At this point, we can describe the complete mapping defined by the succession rule (2.10). In particular Figure 2.8 shows this complete mapping with the pattern $\mathfrak{p}(j) = \mathfrak{p}(1) = x^2\overline{x}$ and Figure 2.9 sketches some levels of the generating tree for the paths in $F^{[\mathfrak{p}(j)]}$ enumerated according to the number of the rise steps.

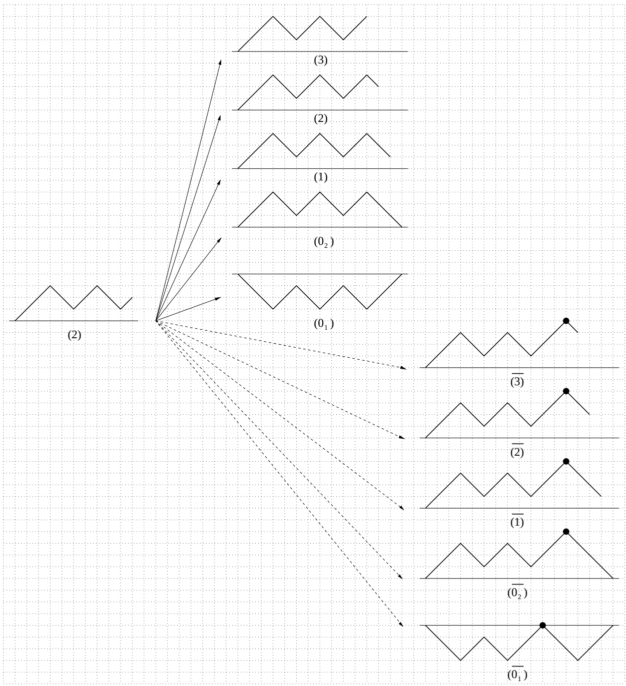


Figure 2.8: The set of lattice paths obtained from a given (k), by means of the succession rule (2.10)

2.2.2 Proving the construction

This construction generates 2^C copies of each path having C forbidden patterns such that 2^{C-1} instances are coded by a sequence of labels ending by a marked label, say $(\overline{k})$, and contain an odd number of marked forbidden patterns, and 2^{C-1} instances are coded by a sequence of labels ending by a non-marked label, say (k), and contain an even number of marked forbidden patterns. For example, Figure 2.10 shows the 4 copies of a given path having 2 forbidden patterns $\mathfrak{p}(j) = \mathfrak{p}(1) = x^2\overline{x}$, where the sequences of labels show the derivation of each path in the generating tree.

This observation is due to the fact that when a path is obtained according to the first production of (2.10) then no marked forbidden pattern is added. Moreover, when a path is obtained according to the second production of (2.10) exactly one marked forbidden pattern is added. In any case, the actions performed to obtain the label (0_1) do not change the number of marked forbidden patterns in the path.

Theorem 1 *The generating tree of the paths in $F^{[\mathfrak{p}(j)]}$, where $\mathfrak{p}(j) = x^{j+1}\overline{x}^j$, $j \geq 1$, according to the number of rise steps, is isomorphic to the tree having its root labelled (0) and recursively defined by the succession rule (2.10).*

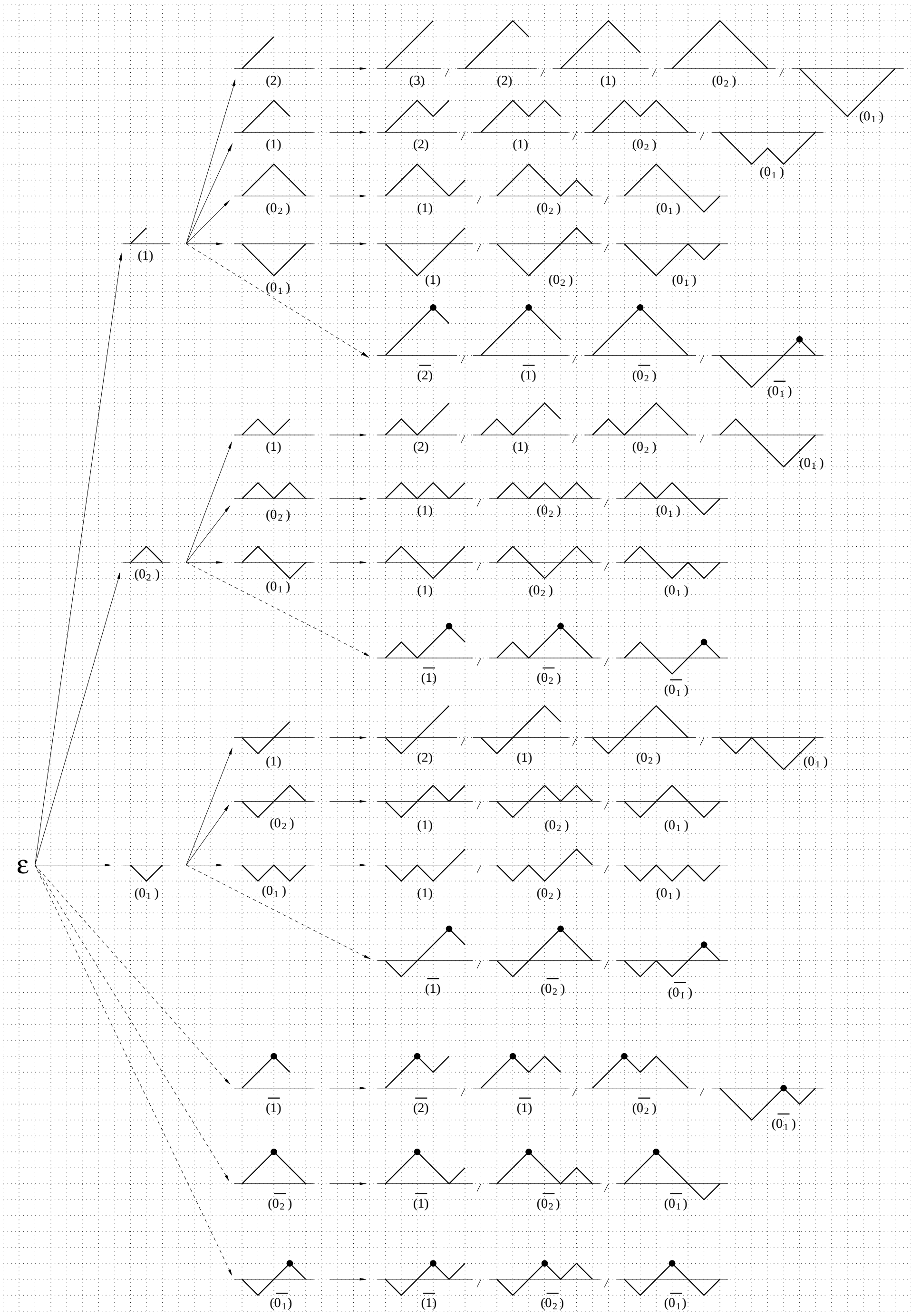


Figure 2.9: Some levels of the generating tree associated to the succession rule (2.10) for the path in $F^{[\mathfrak{p}(j)]}$, being $\mathfrak{p}(j) = \mathfrak{p}(1) = x^2\overline{x}$

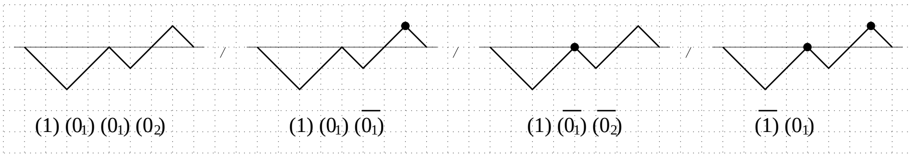


Figure 2.10: The 4 copies of a given path having 2 forbidden patterns, $\mathfrak{p}(j) = \mathfrak{p}(1) = x^2\overline{x}$

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

The first assertion is an immediate consequence of the construction according to the first production of (2.10).

In order to prove the second assertion we have to distinguish two cases (which in their turn are subdivided in 5 and 3 subitems respectively) depending on whether the last forbidden pattern is marked or not. For sake of completeness we report the entire proof, which is indeed rather cumbersome. Anyhow, the interested reader could skip all the subitems, except the first ones. In fact, all the others are obtained from these by means of slight modifications.

We denote by h be 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: $h > j$, $h = j$, $0 < h < j$, $h = 0$ and $h < 0$.

$h > j$: Each path ω in $F \backslash F^{[\mathfrak{p}(j)]}$ can be written as $\omega = \mu x^{j+1}\overline{x}^f\nu$, where $\mu \in F$, $\nu \in F^{[\mathfrak{p}(j)]}$ and $j \leq f \leq d + j + 1$ where $d \geq 0$ is the ordinate of the endpoint of μ (see Figure 2.11).

The path $\overline{\omega}$ which kills ω is obtained by performing on the path μ the following: add the path x^j by applying j times the mapping associated to $(k) \overset{1}{\rightsquigarrow} (k+1)$ of the first production of (2.10), add the path $x\overline{x}^f$ by applying the mapping associated to $(k) \overset{1}{\rightsquigarrow} (d + j + 1 - f)$ of the first production of (2.10). The path ν in $\overline{\omega}$ is obtained as in ω.

$h = j$: Each path ω in $F \backslash F^{[\mathfrak{p}(j)]}$ can be written as $\omega = \mu\overline{x}\gamma x^{j+1}\overline{x}^j\nu$, where $\mu, \gamma \in F$ and $\nu \in F^{[\mathfrak{p}(j)]}$ (see Figure 2.12). We observe that the path γ can contain marked points, with ordinate $b < j$, or not. If the path γ contains no marked point, then it remains strictly under the x-axis,

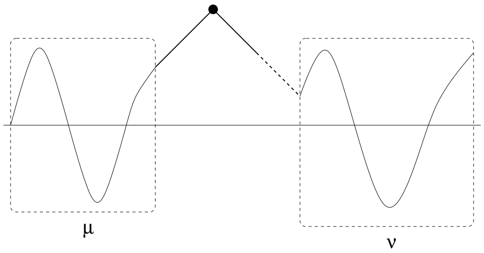


Figure 2.11: A graphical representation of the path ω in the case $h > j$

otherwise the marked forbidden patterns intersect the x-axis when $0 \leq b < j$. In the following cases we consider a path γ having the same property.

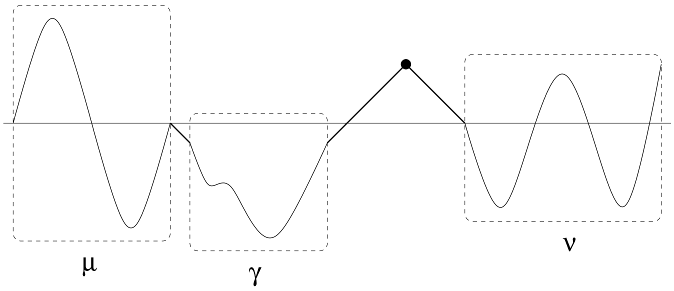


Figure 2.12: A graphical representation of the path ω in the case $h = j$

The path ϖ which kills ω is obtained by performing on $\mu\overline{x}\gamma x$ 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 (2.10), add the path $x\overline{x}^j$ by applying the mapping associated to $(k) \overset{1}{\rightsquigarrow} (0_2)$ of the first production of (2.10). The path ν in ϖ is obtained as in ω.

$0 < h < j$: Each path ω in $F \backslash F^{[\mathfrak{p}(j)]}$ can be written as $\omega = \mu\overline{x}\gamma x^{j+1}\overline{x}^j\eta x\nu$, where $\mu, \gamma \in F$ and $\eta, \nu \in F^{[\mathfrak{p}(j)]}$ (see Figure 2.13). We observe that the path η remains strictly under the x-axis. In the following cases we consider a path η having the same property.

The path ϖ which kills ω is obtained by performing on $\mu\overline{x}\gamma x^{j+1-h}$ 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 (2.10), add the path $x\overline{x}^h$ by applying the mapping associated to $(k) \overset{1}{\rightsquigarrow} (0_2)$ of the first production of (2.10), add the path $\overline{x}^{j-h}\eta x$ by applying consecutive and appropriate mappings of the first production of (2.10) and these

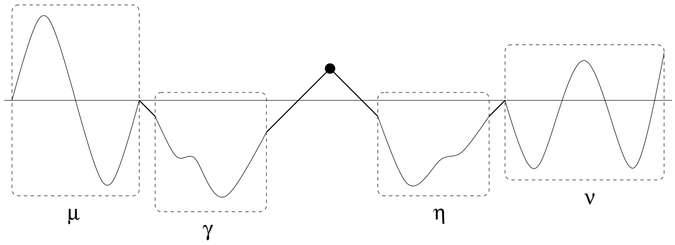


Figure 2.13: A graphical representation of the path ω in the case $0 < h < j$

mappings must be completed by performing the actions giving the label (0_1) in case of no marked points. The path ν in $\overline{\omega}$ is obtained as in ω.

$h = 0$: Each path ω in $F \backslash F^{[\mathfrak{p}(j)]}$ can be written as $\omega = \mu\overline{x}\gamma x^{j+1}\overline{x}^j\eta x\nu$, where $\mu, \gamma \in F$ and $\eta, \nu \in F^{[\mathfrak{p}(j)]}$ (see Figure 2.14).

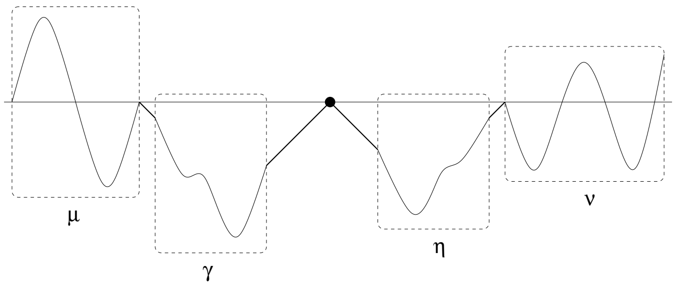


Figure 2.14: A graphical representation of the path ω in the case $h = 0$

The path $\overline{\omega}$ which kills ω is obtained by performing on $\mu\overline{x}\gamma x^{j+1}$ the following: add the path $\overline{x}^j\eta x$ by applying consecutive and appropriate mappings of the first production of (2.10), apply the actions giving the label (0_1) in case of no marked points. The path ν in $\overline{\omega}$ is obtained as in ω.

$h < 0$: Each path ω in $F \backslash F^{[\mathfrak{p}(j)]}$ can be written as $\omega = \mu\overline{x}\gamma x^{j+1}\overline{x}^j\eta x\nu$, where $\mu, \gamma \in F$ and $\eta, \nu \in F^{[\mathfrak{p}(j)]}$ (see Figure 2.15).

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 $\overline{\omega}$ which kills ω is obtained by performing on μ the following: add the path $\overline{x}\gamma x^{j+1}\overline{x}^j\eta x$ by applying consecutive and appropriate mappings of the first production of (2.10), apply the

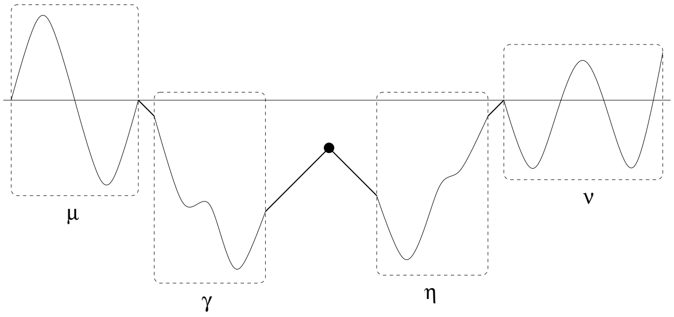


Figure 2.15: A graphical representation of the path ω in the case $h < 0$

actions giving the label (0_1) in case of no marked points. The path ν in $\overline{\omega}$ is obtained as in ω.

In the second subcase, we consider the rightmost point P of the path $\overline{x}\gamma x^{j+1}\overline{x}^j\eta x$ with lowest ordinate. The path $\overline{\omega}$ which kills ω is obtained by performing on μ the following: add the path in $\gamma x^{j+1}\overline{x}^j\eta x$ running from P to the endpoint of the path by applying consecutive and appropriate mappings of the first and second production of (2.10), add the path in $\gamma x^{j+1}\overline{x}^j\eta x$ running from its initial point to P by applying consecutive and appropriate mappings of the first and second production of (2.10), apply the cut and paste actions giving the label (0_1) 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 (2.10). 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: $h > j$, $h = j$ and $h < j$.

$h > j$: Each path ω in $F \backslash F^{[\mathfrak{p}(j)]}$ can be written as $\omega = \mu x^{j+1}\overline{x}^f\nu$, where $\mu \in F$, $\nu \in F^{[\mathfrak{p}(j)]}$ and $j \leq f \leq d + j + 1$ where $d \geq 0$ is the ordinate of the endpoint of μ (Figure 2.16).

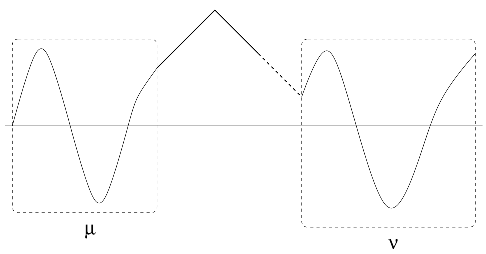


Figure 2.16: A graphical representation of the path ω in the case $h > j$

The path $\overline{\omega}$ which kills ω is obtained by performing on the path μ the following: add the path $x^{j+1}\overline{x}^f$ by applying the mapping associated to $(k) \overset{j+1}{\rightsquigarrow} (\overline{d+j+1-f})$ of the second production of (2.10). The path ν in $\overline{\omega}$ is obtained as in ω.

$h = j$: Each path ω in $F \backslash F^{[\mathfrak{p}(j)]}$ can be written as $\omega = \mu\overline{x}\gamma x^{j+1}\overline{x}^j\nu$, where $\mu, \gamma \in F$ and $\nu \in F^{[\mathfrak{p}(j)]}$ (see Figure 2.17). We observe that the path γ can contains marked points, with ordinate $b < j$, 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 $0 \leq b < j$. In the following case we consider a path γ having the same property.

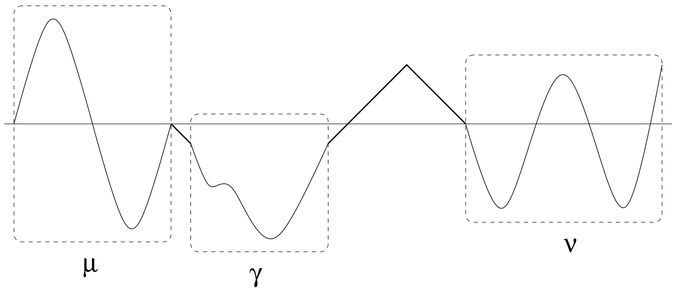


Figure 2.17: A graphical representation of the path ω in the case $h = j$

Let P be the rightmost point of the path $\overline{x}\gamma x^{j+1}\overline{x}^j$ with lowest ordinate. The path $\overline{\omega}$ which kills ω is obtained by performing on μ the following: add the path in $\gamma x^{j+1}\overline{x}^j$ running from P to the endpoint of the path by applying consecutive and appropriate mappings of the first and second production of (2.10), add the path in $\gamma x^{j+1}\overline{x}^j$ running from its initial point to P by applying consecutive and appropriate mappings of the first and second production of (2.10), apply the cut and paste actions giving the label (0_1) 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 (2.10). The path ν in $\overline{\omega}$ is obtained as in ω.

$h < j$: Each path ω in $F \backslash F^{[\mathfrak{p}(j)]}$ can be written as $\omega = \mu\overline{x}\gamma x^{j+1}\overline{x}^j\eta x\nu$, where $\mu, \gamma \in F$ and $\eta, \nu \in F^{[\mathfrak{p}(j)]}$ (see Figure 2.18). We observe that the path η remains strictly under the x-axis.

Let P be the rightmost point of the path $\overline{x}\gamma x^{j+1}\overline{x}^j\eta x$ with lowest ordinate. The path $\overline{\omega}$ which kills ω is obtained by performing on μ the following: add the path in $\gamma x^{j+1}\overline{x}^j\eta x$ running from P to the endpoint of the path by applying consecutive and appropriate mappings

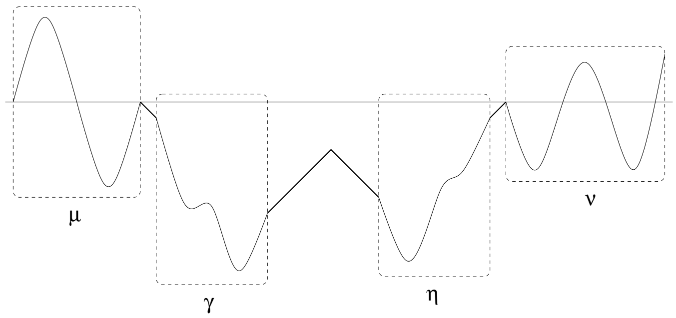


Figure 2.18: A graphical representation of the path ω in the case $h < j$

of the first and second production of (2.10), add the path in $\gamma x^{j+1}\overline{x}^j \eta x$ running from its initial point to P by applying consecutive and appropriate mappings of the first and second production of (2.10), apply the cut and paste actions giving the label (0_1) 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 (2.10). The path ν in ϖ is obtained as in ω.

We observe that for each path ω in $F \backslash F^{[\mathfrak{p}(j)]}$ with n rise steps, C forbidden patterns and last label (k), there exists one and only one path $\overline{\omega}$ in $F \backslash F^{[\mathfrak{p}(j)]}$ with n rise steps, C forbidden patterns and last label $(\overline{k})$ having the same form as ω but such that the last forbidden pattern is marked if it is not in ω and vice-versa.

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

2.2.3 Enumeration of $F^{[\mathfrak{p}(j)]}$

In order to obtain the enumeration of the class $F^{[\mathfrak{p}(j)]}$ according to the number of rise steps, for any fixed $j \geq 1$, we use a standard method, called ECO-method, for the enumeration of combinatorial objects which admit recursive descriptions in terms of generating trees, see [8, 32].

Let N be the set of paths generated by the algorithm described in Section 2.2.1 whose instances are coded by a sequence of labels in the generating tree ending by a non-marked label and M be the set of instances coded by a sequence of labels ending by a marked label. Then $F^{[\mathfrak{p}(j)]} = N \backslash M$.

The paths in N with n rise steps are obtained from the paths in N with $n-1$ rise steps by means of the first production of (2.10) and from those in M with $n-j-1$ rise steps by means of the second production of (2.10).

The paths in M with n rise steps are obtained from the paths in M with $n-1$ rise steps by means of the first production of (2.10) and from those in N with $n-j-1$ rise steps by means of the second production of (2.10).

So, given a path $\omega \in F$ with $n(\omega)$ rise steps and ending point at ordinate $h(\omega)$, from the succession rule (2.10) we have:

$$\begin{aligned} N(x,y) &= 1+\sum_{\omega\in N}\left(\sum_{i=0}^{h(\omega)+1} x^{n(\omega)+1}y^i + x^{n(\omega)+1}y^0\right)+ \\ &\quad +\sum_{\omega\in M}\left(\sum_{i=0}^{h(\omega)+1} x^{n(\omega)+j+1}y^i + x^{n(\omega)+j+1}y^0\right) \end{aligned}$$

$$\begin{aligned} M(x,y) &= \sum_{\omega\in M}\left(\sum_{i=0}^{h(\omega)+1} x^{n(\omega)+1}y^i + x^{n(\omega)+1}y^0\right)+ \\ &\quad +\sum_{\omega\in N}\left(\sum_{i=0}^{h(\omega)+1} x^{n(\omega)+j+1}y^i + x^{n(\omega)+j+1}y^0\right) \end{aligned}$$

Since $\sum_{\omega\in N}\left(\sum_{i=0}^{h(\omega)+1} x^{n(\omega)+1}y^i + x^{n(\omega)+1}y^0\right) =$

$= \sum_{\omega\in N} x^{n(\omega)+1}\left(\frac{y^{h(\omega)+2}-1}{y-1}+1\right) = \frac{xy^2}{y-1}N(x,y) - \frac{x}{y-1}N(x,1) + x\,N(x,1)$,

going on in the same way with the other terms we obtain:

$$\begin{aligned} N(x,y) &= 1+\frac{xy^2}{y-1}N(x,y) - \frac{x}{y-1}N(x,1) + x\,N(x,1)+ \\ &\quad +\frac{x^{j+1}y^2}{y-1}M(x,y) - \frac{x^{j+1}}{y-1}M(x,1) + x^{j+1}M(x,1) \end{aligned}$$

$$\begin{aligned} M(x,y) &= \frac{xy^2}{y-1}M(x,y) - \frac{x}{y-1}M(x,1) + x\,M(x,1)+ \\ &\quad +\frac{x^{j+1}y^2}{y-1}N(x,y) - \frac{x^{j+1}}{y-1}N(x,1) + x^{j+1}N(x,1) \end{aligned}$$

Since $F_j(x,y) = N(x,y) - M(x,y)$ then

$$\begin{aligned} F_j(x,y) &= 1+\frac{xy^2}{y-1}F_j(x,y) - \frac{x}{y-1}F_j(x,1)+ \\ &\quad +x\,F_j(x,1) - \frac{x^{j+1}y^2}{y-1}F_j(x,y)+ \\ &\quad +\frac{x^{j+1}}{y-1}F_j(x,1) - x^{j+1}F_j(x,1) \end{aligned}$$

and

$$
\begin{aligned}
(y-1-x(1-x^j)y^2)F_j(x,y) &= y-1-xF_j(x,1)+x(y-1)\,F_j(x,1)+\\
&\quad +x^{j+1}F_j(x,1)-x^{j+1}(y-1)F_j(x,1)
\end{aligned}
$$

Going on and using the kernel method [8] we obtain the generating function $F_j(x)$ for the words $\omega \in F^{[\mathfrak{p}(j)]}$, for any fixed $j \geq 1$, according to the number of ones:

$$F_j(x) = F_j(x,1) = \frac{1-y_0(x)}{x(y_0(x)-2)(1-x^j)}$$

where

$$y_0(x) = \frac{1-\sqrt{1-4x(1-x^j)}}{2x(1-x^j)}.$$

Note that, for $\mathfrak{p}(j) = \mathfrak{p}(1) = 110$, the generating function $F_j(x)$ coincides with (2.1) in Section 2.1 and the first number of the sequence enumerating the binary words in $F^{[\mathfrak{p}(1)]}$, according to the number of ones, are:

$$1, 3, 7, 15, 31, 63, 127, 255, 511, 1023, \ldots$$

.

2.3 A generating algorithm for $F^{[\mathfrak{p}(j,j)]}$

In this section, we study the construction and the enumeration of the class $F^{[\mathfrak{p}(j,j)]}$, where $\mathfrak{p}(j,j) = 1^j0^j$, for any fixed $j \geq 1$, which is a slight modification of study in Section 2.2.1.

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

In this section, we define an algorithm to construct the set $F^{[\mathfrak{p}(j,j)]}$, where $\mathfrak{p}(j,j) = x^j\overline{x}^j = 1^j0^j$, for any fixed $j \geq 1$. The growth of the set, according to the number of rise steps or equivalently the number of ones, can be described by the following jumping and marked succession rule:

$$
\begin{cases}
(0) & \\
(k) \overset{1}{\rightsquigarrow} (k+1)(k)\cdots(1)(0_2)(0_1) & k \geq 0\\
(0) \overset{j}{\rightsquigarrow} (\overline{0_2}) & \\
(k) \overset{j}{\rightsquigarrow} (\overline{k})(\overline{k-1})\cdots(\overline{1})(\overline{0_2})(\overline{0_1}) & k \geq 1
\end{cases}
\tag{2.11}
$$

This rule can be represented as a tree having its root labelled (0) and where each node with label (k) at level n gives $k+3$ sons at level $n+1$ labelled $(k+1), \cdots, (1), (0_2), (0_1)$ and $k+2$ sons at level $n+j$ with labels

$(\overline{k}), \cdots, (\overline{1}), (\overline{0_2}), (\overline{0_1})$, if $k \geq 1$, or only one son with label $(\overline{0_2})$ at level $n+j$ if $k = 0$. The generating algorithm associates a lattice path in $F^{[\mathfrak{p}(j,j)]}$ to a sequence of labels obtained by means of the succession rule (2.11). This give a construction for the set $F^{[\mathfrak{p}(j,j)]}$ according to the number of rise steps or equivalently the number of ones.

The axiom (0) is associated to the empty path ε.
A lattice 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 (2.11) having $k+1, \ldots, 1, 0, 0$ as endpoint ordinate, respectively.

As in Section 2.2.1, the first $k+2$ paths are obtained by adding to ω a sequence of steps consisting of one rise step followed by $k+1-h$, $0 \leq h \leq k+1$, fall steps (see Figure 2.1 in Section 2.2.1).

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 and having a rise step as last step are never obtained. These paths have the label (0_1) of the first production in (2.11).

We define a *marked forbidden pattern* $\mathfrak{p}(j,j)$ as a pattern $\mathfrak{p}(j,j) = x^j\overline{x}^j$ whose steps cannot be split, that is they must always be contained all together in that defined sequence. We say that a point is strictly contained in a given marked forbidden pattern $\mathfrak{p}(j,j)$ if it is in $\mathfrak{p}(j,j)$ and it is different from both its initial point and its last point. We denote a marked forbidden pattern by marking its peak.

A cut operation, i.e the procedure which splits a given path into two subpaths, is not possible within a marked forbidden pattern $\mathfrak{p}(j,j)$. After a cut operation, it is not allowed to switch any rise step with a fall one, and viceversa, inside a marked forbidden pattern, but it can be translated.

A lattice path $\omega \in F$, with n rise steps and such that its endpoint has ordinate 0, provides one lattice path, with $n+j$ rise steps, according to the second production of (2.11), having 0 as endpoint ordinate, obtained by adding to ω a sequence of steps consisting of the marked forbidden pattern $\mathfrak{p}(j,j)$ (see Figure 2.19).

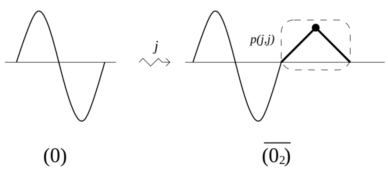


Figure 2.19: The mapping associated to $(0) \stackrel{j}{\rightsquigarrow} (\overline{0_2})$ of (2.11)

A lattice path $\omega \in F$, with n rise steps and such that its endpoint has ordinate $k \geq 1$, provides $k+2$ lattice paths, with $n+j$ rise steps, according to the last production of (2.11), having $k, \ldots, 1, 0, 0$ as endpoint ordinate, respectively. The first $k+1$ labels are obtained by adding to ω a sequence of steps consisting of the marked forbidden pattern $\mathfrak{p}(j, j)$ followed by $k-h$, $0 \leq h \leq k$, fall steps (see Figure 2.20). 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.

At this point the label (0_1) due to the productions of (2.11) yields lattice paths which either do not contain marked forbidden patterns in its rightmost suffix and end on the x-axis by a rise step or having the rightmost marked point with ordinate less than j.

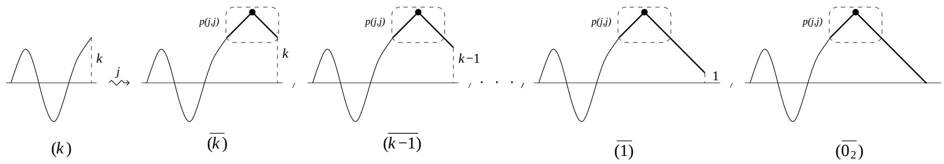


Figure 2.20: The mapping associated to $(k) \overset{j}{\rightsquigarrow} (\overline{k}) \ldots (\overline{1})(\overline{0_2})$ of (2.11)

In order to obtain the label (0_1) according to the first production of (2.11), we consider the path ω' obtained from ω by adding a sequence of steps consisting of one rise step followed by k fall steps, while in order to obtain the label (0_1) according to the second production of (2.11), we consider the path ω' obtained from ω by adding a sequence of steps consisting of the marked forbidden pattern $\mathfrak{p}(j, j) = x^j \overline{x}^j$ followed by $k-1$ 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.

As in Section 2.2.1, 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_1) 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 2.3 is Section 2.2.1).

If the suffix φ' contains marked points, then the desired label (0_1) is associated to the path obtained by applying cut and paste actions described in Section 2.2.1.

At this point, we have the complete mapping defined by the succession rule (2.11).

2.3.2 Proving the construction

The above construction generates 2^C copies of each path having C forbidden patterns such that 2^{C-1} instances are coded by a sequence of labels ending by a marked label, say $(\overline{k})$, and contain an odd number of marked forbidden patterns, and 2^{C-1} instances are coded by a sequence of labels ending by a non-marked label, say (k), and contain an even number of marked forbidden patterns. This is due to the fact that when a path is obtained according to the first production of (2.11) then no marked forbidden pattern is added. Moreover, when a path is obtained according to the other productions of (2.11) exactly one marked forbidden pattern is added. In any case, the actions performed to obtain the label (0_1) do not change the number of marked forbidden patterns in the path itself.

Theorem 2 *The generating tree of the paths in $F^{[\mathfrak{p}(j,j)]}$, where $\mathfrak{p}(j,j) = x^j\overline{x}^j$, $j \geq 1$, according to the number of rise steps, is isomorphic to the tree having its root labelled* (0) *and recursively defined by the succession rule* (2.11).

The proof of the Theorem 2 is analogous to the proof of Theorem 1 and it is omitted for brevity.

2.3.3 Enumeration of $F^{[\mathfrak{p}(j,j)]}$

As in Section 2.2.1, in order to obtain the enumeration of the class $F^{[\mathfrak{p}(j,j)]}$ according to the number of rise steps, for any fixed $j \geq 1$, we use a standard method, called ECO-method, for the enumeration of combinatorial objects which admit recursive descriptions in terms of generating trees, see [8, 32].

let Z be the set of paths whose instances are coded by a sequence of labels in the generating tree ending by a non-marked zero, S be the set of paths whose instances are coded by a sequence of labels ending by a marked zero, N be the set of paths whose instances are coded by a sequence of labels ending by a non-marked $k \geq 1$ and M be the set of paths whose instances are coded by a sequence of labels ending by a marked $k \geq 1$. Then $F^{[\mathfrak{p}(j,j)]} = (Z \backslash S) \cup (N \backslash M)$.

The succession rule (2.11) can be written as:

$$\left\{\begin{array}{ll} (0) & \\ (0) \overset{1}{\rightsquigarrow} (1)(0_2)(0_1) & \\ (k) \overset{1}{\rightsquigarrow} (k+1)(k)\cdots(1)(0_2)(0_1) & k \geq 1 \\ (0) \overset{j}{\rightsquigarrow} (\overline{0_2}) & \\ (k) \overset{j}{\rightsquigarrow} (\overline{k})(\overline{k-1})\cdots(\overline{1})(\overline{0_2})(\overline{0_1}) & k \geq 1 \end{array}\right. \tag{2.12}$$

Let us denote by $n(\omega)$ the number of rise steps of a path $\omega \in F$ and by $h(\omega)$ the last point's ordinate of ω itself. From the succession rule (2.12) we have:

$$Z(x,1) = 1 + 2xZ(x,1) + 2xN(x,1) + x^jS(x,1) + 2x^jM(x,1), \tag{2.13}$$

$$S(x,1) = 2xS(x,1) + 2xM(x,1) + x^j Z(x,1) + 2x^j N(x,1), \tag{2.14}$$

$$N(x,y) = xyZ(x,1) + \sum_{\omega \in N} \sum_{i=1}^{h(\omega)+1} x^{n(\omega)+1} y^i + \sum_{\omega \in M} \sum_{i=1}^{h(\omega)} x^{n(\omega)+j} y^i, \tag{2.15}$$

$$M(x,y) = xyS(x,1) + \sum_{\omega \in M} \sum_{i=1}^{h(\omega)+1} x^{n(\omega)+1} y^i + \sum_{\omega \in N} \sum_{i=1}^{h(\omega)} x^{n(\omega)+j} y^i. \tag{2.16}$$

Since $\sum_{\omega \in N} \sum_{i=1}^{h(\omega)+1} x^{n(\omega)+1} y^i = \sum_{\omega \in N} x^{n(\omega)+1} \left(\frac{y^{h(\omega)+2} - y}{y-1} \right) =$

$= \frac{xy^2}{y-1} N(x,y) - \frac{xy}{y-1} N(x,1)$ going on in the same way with the other terms, then (2.15) and (2.16) can rewritten as:

$$N(x,y) = xyZ(x,1) + \frac{xy^2}{y-1} N(x,y) - \frac{xy}{y-1} N(x,1) + \frac{x^j y}{y-1} M(x,y) - \frac{x^j y}{y-1} M(x,1),$$

$$M(x,y) = xyS(x,1) + \frac{xy^2}{y-1} M(x,y) - \frac{xy}{y-1} M(x,1) + \frac{x^j y}{y-1} N(x,y) - \frac{x^j y}{y-1} N(x,1).$$

Since $T(x,y) = N(x,y) - M(x,y)$ then:

$$T(x,y) = xy(Z(x,1) - S(x,1)) + \frac{(xy^2 - x^j y)}{y-1} T(x,y) - \frac{(xy - x^j y)}{y-1} T(x,1)$$

that is

$$T(x,y)(y - 1 - xy^2 + x^j y) = xy(y-1)(Z(x,1) - S(x,1)) - (xy - x^j y) T(x,1).$$

Let $y_0(x) = \frac{x^j + 1 - \sqrt{(x^j+1)^2 - 4x}}{2x}$ be a solution of $xy^2 - (x^j + 1)y + 1 = 0$. For any $j > 1$, we have:

$$T(x,1) = \frac{y_0(x) - 1}{1 - x^{j-1}} (Z(x,1) - S(x,1)) \tag{2.17}$$

so obtaining the desired equation according to the number of ones, only. From (2.13) and (2.14) we obtain:

$$W(x,1) = Z(x,1) - S(x,1) = \frac{1 + 2x(1 - x^{j-1}) T(x,1)}{1 - 2x + x^j}. \tag{2.18}$$

By solving (2.17) and (2.18) we have:

$$T(x,1) = \frac{y_0(x) - 1}{(1 - x^{j-1})(1 + x^j - 2xy_0(x))},$$

$$W(x,1) = \frac{1}{(1 + x^j - 2xy_0(x))}.$$

Therefore the generating function $F_{jj}(x)$, $j > 1$, for the words $\omega \in F^{[\mathfrak{p}(j,j)]}$ according to the number of ones is:

$$F_{jj}(x) = T(x,1) + W(x,1) = \frac{y_0(x) - x^{j-1}}{(1 - x^{j-1})(1 + x^j - 2xy_0(x))}.$$

Let us remark that the generating function $F_{jj}(x)$ depends only on the number of ones in the forbidden pattern.

Example 1 *Let $\mathfrak{p}(j,j) = \mathfrak{p}(4,4)$, the first numbers of the sequence enumerating the binary words in $F^{[\mathfrak{p}(4,4)]}$, according to the number of ones, are:* $1, 3, 10, 35, 125, 454, 1671, 6211, 23261, \cdots$ *being*

$$F_{44}(x) \quad = \quad \frac{1 - x^4 - \sqrt{x^8 + 2x^4 + 1 - 4x}}{2x(1 - x^3)\sqrt{x^8 + 2x^4 + 1 - 4x}}$$

the associated generating function.

In the case $j = 1$, from (2.13) and (2.14) we have $W(x,1) = 1 + xW(x,1)$, and from (2.15) and (2.16) we obtain $T(x,y) = xyW(x,1) + xyT(x,y)$ that is

$$W(x,1) = \frac{1}{1-x}, \qquad T(x,1) = \frac{x}{(1-x)^2}.$$

Therefore the generating function $F_{11}(x)$ is:

$$F_{11}(x) = T(x,1) + W(x,1) = \frac{1}{(1-x)^2}$$

that is $F_{11}(x)$ is the generating function of the succession $x_n = n+1$, $n \geq 1$.

Indeed, when $j = 1$, the set $F^{[\mathfrak{p}(1,1)]}$ of binary words with n ones and avoiding the forbidden pattern $\mathfrak{p} = 10$ is $F^{[\mathfrak{p}(1,1)]} = \{0^m 1^n \,|\, 0 \leq m \leq n\}$.

2.4 A generating algorithm for $F^{[\mathfrak{p}(j,i)]}$

In this section, 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 $\mathfrak{p}(j,i) = 1^j 0^i$, $0 < i < j$. 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 $\mathfrak{p}(j,i) = 1^j 0^i$, $0 < i < j$.

In this case the theory of Riordan arrays is not applicable, neither the algorithmic approach allows to obtain their generating functions. Anyway it gives us a way to construct all the objects in this class.

2.4.1 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 split, that is, they must always be contained all together in that defined sequence. We say that a point is strictly contained in a given marked forbidden pattern $\mathfrak{p}(j,i)$ if it is in $\mathfrak{p}(j,i)$ and it is different from both its initial point and its last point. We denote a marked forbidden pattern by marking its peak.

A cut operation, i.e the procedure which splits a given path into two subpaths, is not possible within a marked forbidden pattern $\mathfrak{p}(j,i)$. After a cut operation, it is not allowed to switch any rise step with a fall one, and viceversa, inside a marked forbidden pattern, but it can be translated.

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 1 *A path ω in F is a* Δ-path *if:*

- *it ends on the x-axis (see Figure 2.21.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 2.21.b)) or not (see Figure 2.21.c)). If φ contains marked forbidden patterns $\mathfrak{p}(j,i)$, then their marked points have ordinate $b \geq j$.*

Definition 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 2.21.d)).*

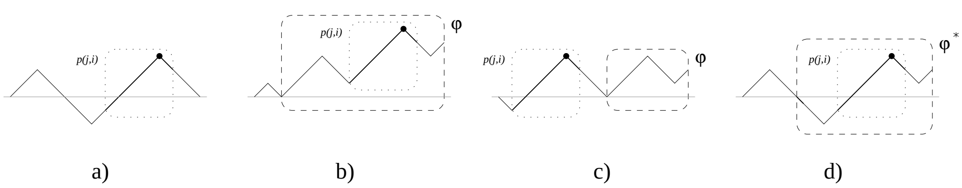


Figure 2.21: Some examples of paths in F

Δ-paths in F

For each Δ-path ω in F having k as ordinate of its endpoint, we apply the following succession rule (2.19), for each $k \geq 0$:

$$(2.19) \qquad \begin{cases} (0) \\ (k) & \overset{1}{\rightsquigarrow} (k+1)(k)\cdots(2)(1)(0)^2 \\ (k) & \overset{j}{\rightsquigarrow} (\overline{k+j-i})\ldots(\overline{s+1})(\overline{s})(\overline{s-1})^2\ldots(\overline{1})^s(\overline{0})^{s+1} \end{cases}$$

In the second production of (2.19), 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 (2.19).

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 (2.19) having $k+1, k, \ldots, 1, 0, 0$ as endpoint ordinate, respectively.

In the similar way described in Section 2.2.1, the first $k+2$ 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 2.1 in Section 2.2.1.

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 last label (0) of the first production in (2.19).

In order to obtain the last label (0) according to the first production of (2.19), we consider the 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.

In the similar way described in Section 2.2.1, 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 2.3 in Section 2.2.1.

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 - described in Section 2.2.1 - 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, called α, and the path running from the initial point in φ' to z, called β, see Figure 2.5 in Section 2.2.1.
This last mapping can be inverted as in Section 2.2.1. In particular, 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 P 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 P to the endpoint of the path and α is the path φ^* running from the endpoint of d to P, see Figure 2.6 in Section 2.2.1.

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 (2.19). The first $1+k+j-i$ lattice paths have $k+j-i, k+j-i-1, \dots, s, \dots, 2, 1, 0$ 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 2.22).

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 (2.19), must give lattice paths having the rightmost marked point with ordinate less than j.

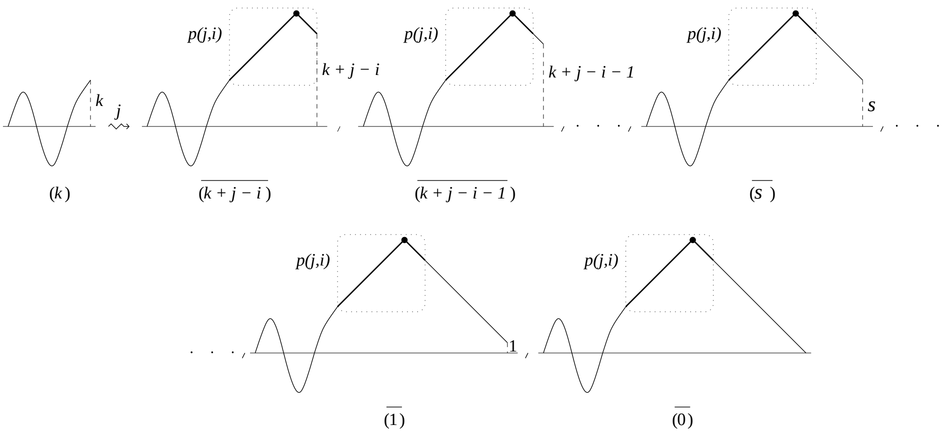


Figure 2.22: The mapping associated to $(k) \overset{j}{\rightsquigarrow} (\overline{k+j-i})(\overline{k+j-i-1})\dots$
$\dots(\overline{s})\dots(\overline{1})(\overline{0})$ of (2.19)

In order to obtain the $\sum_{m=1}^{s} m$ marked labels according to the second production of (2.19), 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 2.22.

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 ω'' beginning from the x-axis and strictly remaining above the x-axis (see Figure 2.23).

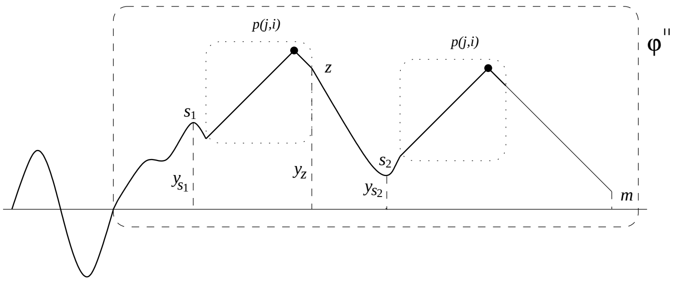


Figure 2.23: A representation of $\omega'' = \omega\mathfrak{p}(j,i)\overline{x}^{k+j-i-m} = v\varphi''$, $1 \leq m \leq s$

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 (2.19) 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'', $1 \leq m \leq s$, then s_2 does not exist and so $s = y_z - y_{s_1}$.

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 P (see Figure 2.24).

Remind that, from the reverse of the cut and paste actions, the point P 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 P 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 P 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

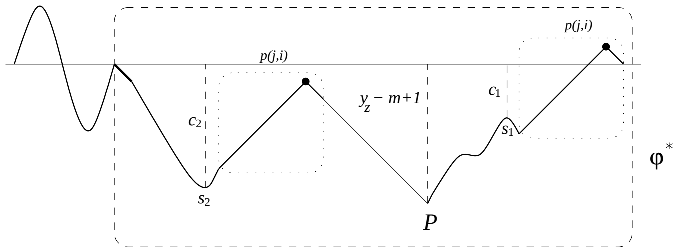


Figure 2.24: A graphical representation of the path $v\varphi^*$ obtained by applying the cut and paste actions to the path ω''

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, $1 \le m \le s$. By adding g fall steps for each g, $0 < g \le m-1$, to each of such paths (see Figure 2.25), we obtain the complete mapping associated to the second production of (2.19).

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', $0 \le m' < k+j-i-s$, we have already obtained paths.

Figure 2.26 shows the complete mapping according to the rule (2.19) on an example with the pattern $\mathfrak{p}(j,i) = x^5\overline{x}^2$.

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 $k \ge 1$:

$$(2.20) \qquad \begin{cases} (k) \overset{1}{\rightsquigarrow} (k+1)(k)\cdots(2)(1)(0) \\ (k) \overset{j}{\rightsquigarrow} (\overline{k+j-i})(\overline{k+j-i-1})\cdots(\overline{2})(\overline{1})(\overline{0}) \end{cases}$$

A Γ-path $\omega \in F$, with n rise steps and such that its endpoint has ordinate k, provides $k+2$ lattice paths, with $n+1$ rise steps, according to the first production of (2.20) having $k+1, k, \ldots, 2, 1, 0$ 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, $0 \le h \le k+1$.

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 (2.20) having $k+j-i, k+j-i-1, \ldots, 2, 1, 0$ as endpoint ordinate, respectively. These labels 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, $0 \le h \le k+j-i$.

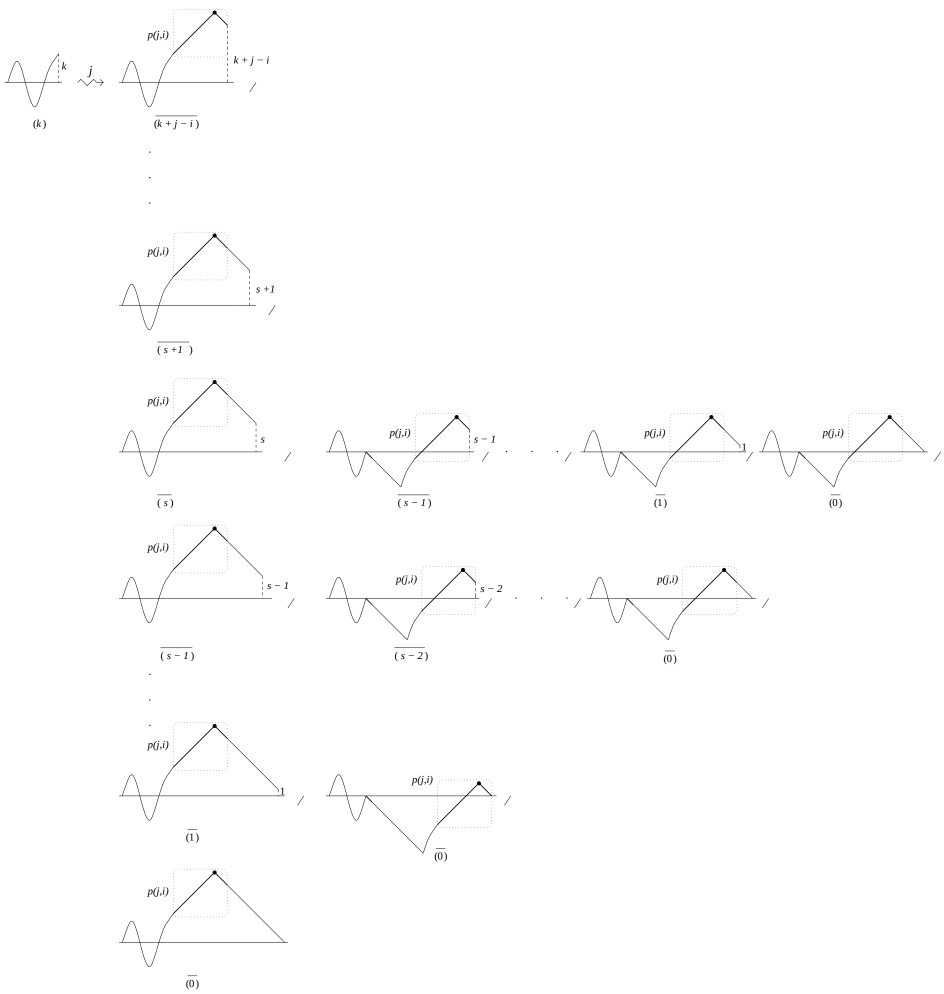


Figure 2.25:
The mapping associated to $(k) \overset{j}{\rightsquigarrow} (\overline{k+j-i})\dots(\overline{s})(\overline{s-1})^2\dots(\overline{1})^s(\overline{0})^{s+1}$ of (2.19)

Figure 2.27 shows the complete mapping according to the rule (2.20) on an example with the pattern $\mathfrak{p}(j,i) = \mathfrak{p}(5,2) = x^5\overline{x}^2$.

2.4.2 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 2.28 shows

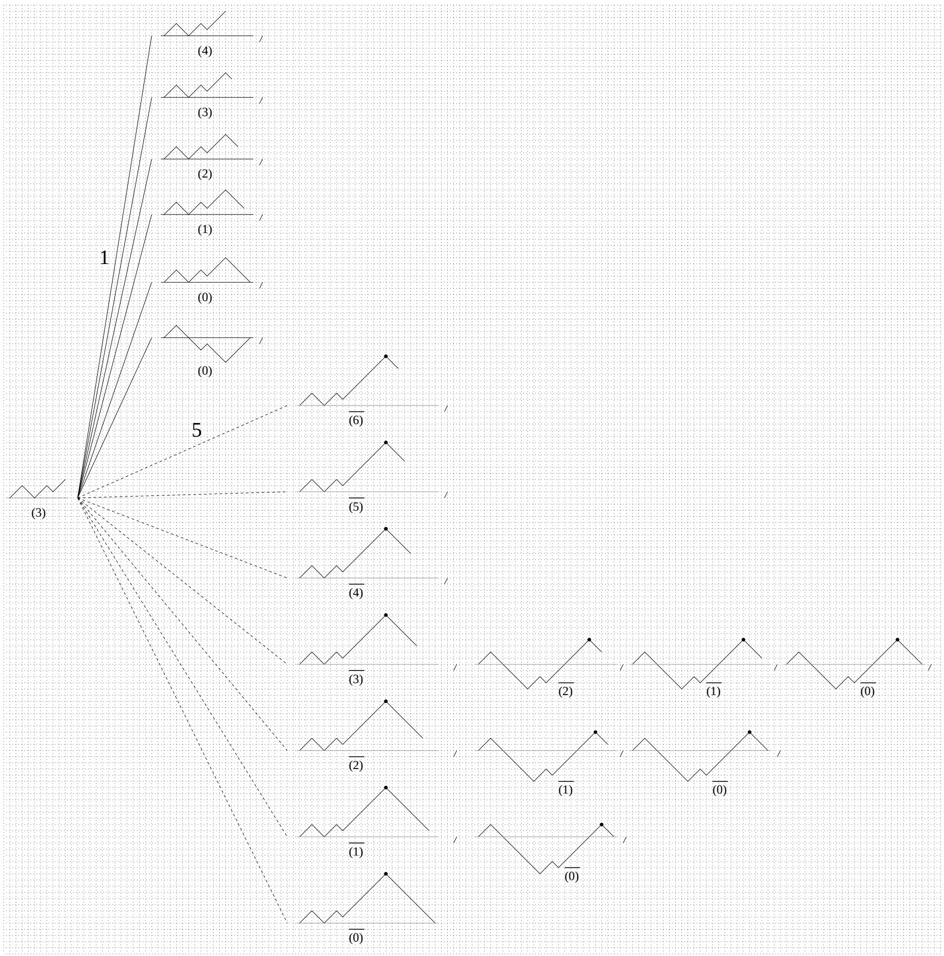


Figure 2.26: The mapping associated to the succession rule (2.19), being $\mathfrak{p}(j,i) = x^5\overline{x}^2$

the 4 copies of a given path having two forbidden patterns $\mathfrak{p}(j,i) = \mathfrak{p}(5,2) = x^5\overline{x}^2$. 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. This observation is due to the fact that when a path is obtained either according to the first production of (2.19) or according to the first production of (2.20) then no marked forbidden pattern is added. Moreover when a path is obtained either according to the second production of (2.19) or according to the second production of (2.20) 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 (2.19) or the $\sum_{m=1}^{s} m$ marked labels, according to the second production of (2.19), do not change the number of marked forbidden patterns in the path.

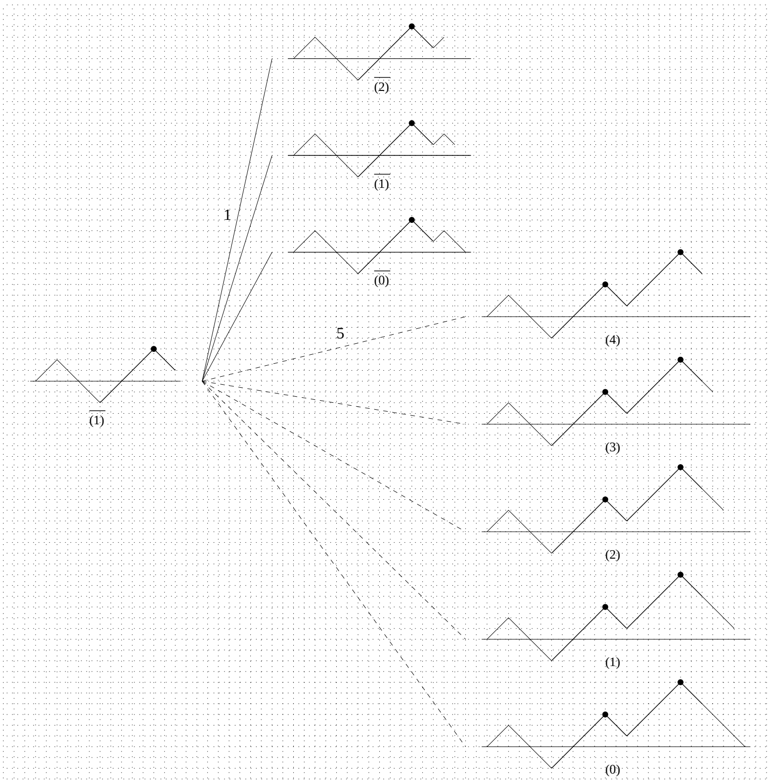


Figure 2.27: The mapping associated to the succession rule (2.20), being $\mathfrak{p}(j,i) = \mathfrak{p}(5,2) = x^5\overline{x}^2$

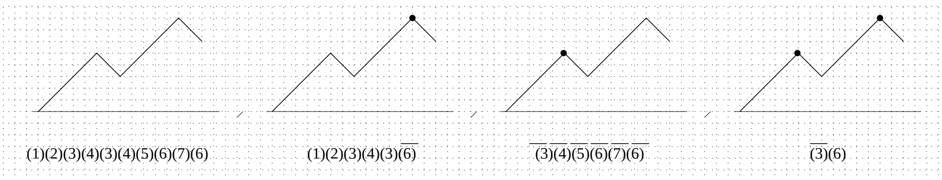


Figure 2.28: The 4 copies of a given path having 2 forbidden patterns, $\mathfrak{p}(j,i) = \mathfrak{p}(5,2) = x^5\overline{x}^2$

Theorem 3 *The generating tree of the 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 rules (2.19) and (2.20), related to the shape of the path $\omega \in F$.*

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 ordinate 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 ordinate 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 (2.19).

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 $h < 0$.

$h \geq j$: 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 2.29).

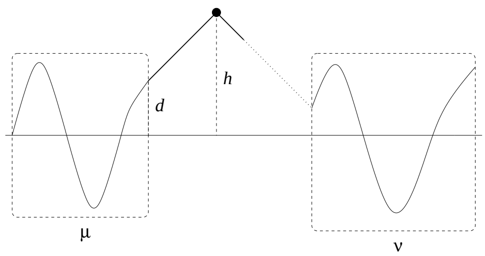


Figure 2.29: A representation of the path ω in the case $h \geq j$

If μ is a Δ-path then 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 $(k) \stackrel{1}{\rightsquigarrow} (k+1)$ of the first production of (2.19); add $x\overline{x}^f$ by applying the mapping associated to $(k) \stackrel{1}{\rightsquigarrow} (d+j-f)$ of the first production of (2.19). Otherwise, if μ is a Γ-path then 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 $(k) \stackrel{1}{\rightsquigarrow} (k+1)$ of the first production of (2.20); add the path $x\overline{x}^f$ by applying the mapping associated to $(k) \stackrel{1}{\rightsquigarrow} (d+j-f)$ of the first production of (2.20). In each case the path ν in ϖ is obtained as in ω.

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

$h > i$: 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 2.30). 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 $0 \leq b < i$.

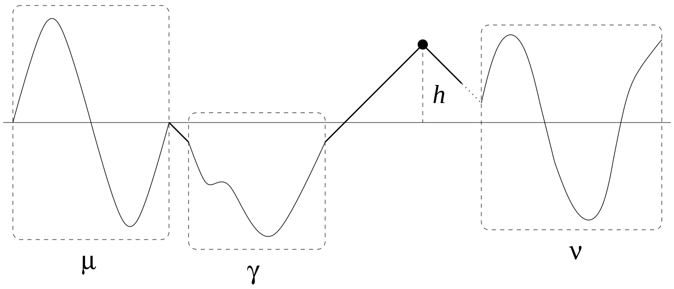


Figure 2.30: A representation of the path ω in the case $0 < h < j$ with $h > i$

The path ϖ 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 (2.19); add the path $x\overline{x}^f$ by applying the mapping associated to $(k) \overset{1}{\rightsquigarrow} (h-f)$ of the first production of (2.19). The path ν in ϖ is obtained as in ω.

$h = i$: 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 2.31).

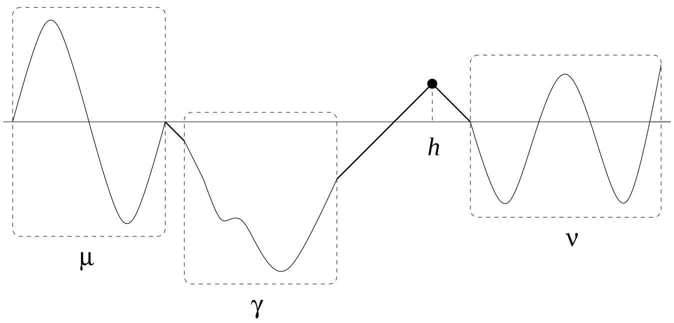


Figure 2.31: A representation of the path ω in the case $0 < h < j$ with $h = i$

The path ϖ 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 (2.19); add the path $x\overline{x}^{i}$ by applying the mapping associated to $(k) \overset{1}{\rightsquigarrow} (0)$ of the first production of (2.19) for the second label (0). The path ν in $\overline{\omega}$ is obtained as in ω.

$h < i$: 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 2.32). We observe that the path η remains strictly under the x-axis.

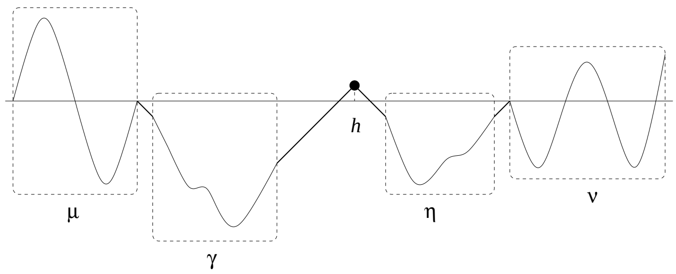


Figure 2.32: A representation of the path ω in the case $0 < h < j$ with $h < i$

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 (2.19); add the path $x\overline{x}^{h}$ by applying the mapping associated to $(k) \overset{1}{\rightsquigarrow} (0)$ of the first production of (2.19) for the second label (0); add the path $\overline{x}^{i-h}\eta x$ by applying consecutive and appropriate mappings associated to $(k) \overset{1}{\rightsquigarrow} (k+1)\dots(2)(1)$ of the first production of (2.19) 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 $\overline{\omega}$ is obtained as in ω.

$h = 0$: 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 2.33).

The path $\overline{\omega}$ which kills ω is obtained by performing on $\mu\overline{x}\gamma x^{j}$ the following: add the path $\overline{x}^{i}\eta x$ by applying consecutive and appropriate mappings associated to $(k) \overset{1}{\rightsquigarrow} (k+1)\dots(1)$ of the first production of (2.19) 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 $\overline{\omega}$ is obtained as in ω.

$h < 0$: 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 2.34).

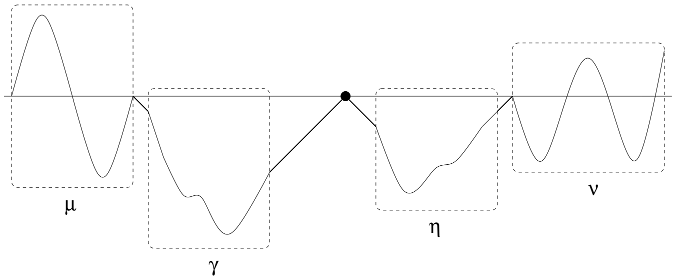


Figure 2.33: A representation of the path ω in the case $h = 0$

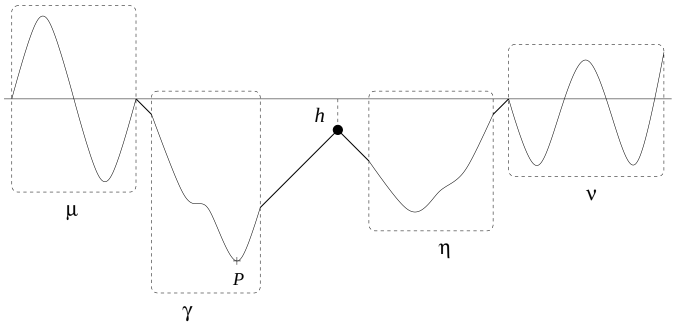


Figure 2.34: A representation of the path ω in the case $h < 0$

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 $\overline{\omega}$ 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 $(k) \overset{1}{\rightsquigarrow} (k+1)\dots(2)(1)$ of the first production of (2.19) 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 $\overline{\omega}$ is obtained as in ω.

In the second subcase, we consider the rightmost point P 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 $\gamma x^{j+1} \overline{x}^j \eta x$ running from P to the endpoint of the path by applying consecutive and appropriate mappings associated to $(k) \overset{1}{\rightsquigarrow} (k+1)\dots(2)(1)$ of the first production of (2.19) and by applying consecutive and appropriate mappings associated to $(k) \overset{j}{\rightsquigarrow} (\overline{k+j-i})\dots(\overline{2})(\overline{1})$ of the second production of (2.19); add the path in $\gamma x^j \overline{x}^i \eta x$ running from its initial point to P by applying consecutive and appropriate mappings associated to $(k) \overset{1}{\rightsquigarrow} (k+1)\dots(2)(1)$ of the first production of (2.19) and by applying consecutive and appropriate mappings associated to $(k) \overset{j}{\rightsquigarrow} (\overline{k+j-i})\dots(\overline{2})(\overline{1})$ of the second production of (2.19); 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 (2.19). The path ν in ϖ 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 $h < j$.

$h \geq j$: 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 2.35).

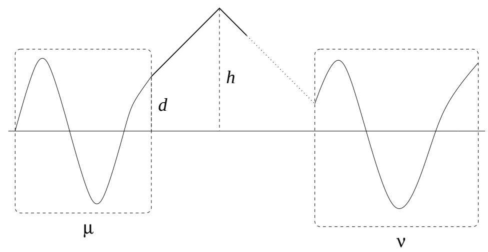


Figure 2.35: A representation of the path ω in the case $h \geq j$

If μ is a Δ-path then the path ϖ 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 (2.19). Otherwise, if μ is a Γ-path then the path ϖ 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 (2.20). In each case, the path ν in ϖ is obtained as in ω.

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

$h > i$: 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 2.36). 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 P be the rightmost point of the path $\overline{x}\gamma x^j \overline{x}^i$ with lowest ordinate. The path ϖ which kills ω is obtained by performing on μ the following: add the path in $\gamma x^j \overline{x}^i$ running from P to the

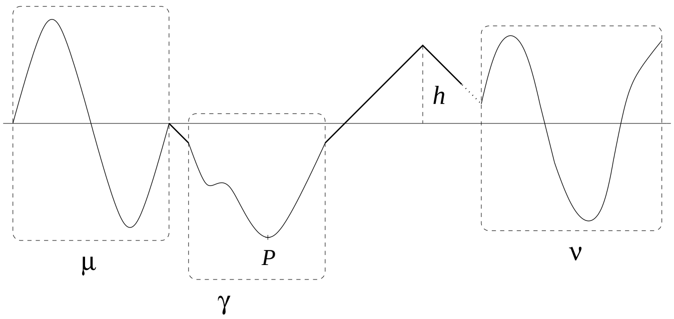


Figure 2.36: A representation of the path ω in the case $h < j$ with $h > i$

endpoint of the path by applying consecutive and appropriate mappings associated to $(k) \overset{1}{\rightsquigarrow} (k+1)\dots(2)(1)$ of the first production of (2.19) and by applying consecutive and appropriate mappings associated to $(k) \overset{j}{\rightsquigarrow} (\overline{k+j-i})\dots(\overline{2})(\overline{1})$ of the second production of (2.19); add the path in $\gamma x^j \overline{x}^i$ running from its initial point to P by applying consecutive and appropriate mappings associated to $(k) \overset{1}{\rightsquigarrow} (k+1)\dots(2)(1)$ of the first production of (2.19) and by applying consecutive and appropriate mappings associated to $(k) \overset{j}{\rightsquigarrow} (\overline{k+j-i})\dots(\overline{2})(\overline{1})$ of the second production of (2.19); 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 (2.19). Obviously the last forbidden pattern in the path must be generated by applying the mapping of the second production of (2.19). The path ν in ϖ is obtained as in ω. Note that, in case of $i \leq f < h$, any prefix of ν in ϖ 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 (2.20).

$h = i$: 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 2.37).

Let P be the rightmost point of the path $\overline{x}\gamma x^j \overline{x}^i$ with lowest ordinate. The path ϖ which kills ω is obtained by performing on μ the following: add the path in $\gamma x^j \overline{x}^i$ running from P to the endpoint of the path by applying consecutive and appropriate mappings associated to $(k) \overset{1}{\rightsquigarrow} (k+1)\dots(2)(1)$ of the first production of (2.19) and by applying consecutive and appropriate mappings associated to $(k) \overset{j}{\rightsquigarrow} (\overline{k+j-i})\dots(\overline{2})(\overline{1})$ of the second production of (2.19); add the path in $\gamma x^j \overline{x}^i$ running from its initial point to P by applying consecutive and appropriate mappings associated to $(k) \overset{1}{\rightsquigarrow} (k+1)\dots(2)(1)$ of the first production

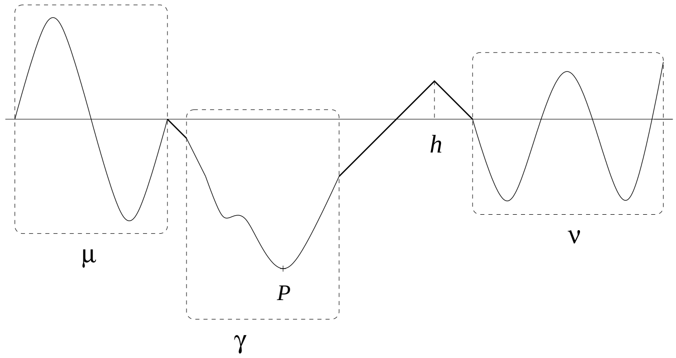


Figure 2.37: A representation of the path ω in the case $h < j$ with $h = i$

of (2.19) and by applying consecutive and appropriate mappings associated to $(k) \overset{j}{\rightsquigarrow} (\overline{k+j-i}) \dots (\overline{2})(\overline{1})$ of the second production of (2.19); 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 (2.19). The path ν in $\overline{\omega}$ is obtained as in ω.

$h < i$: 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 2.38).

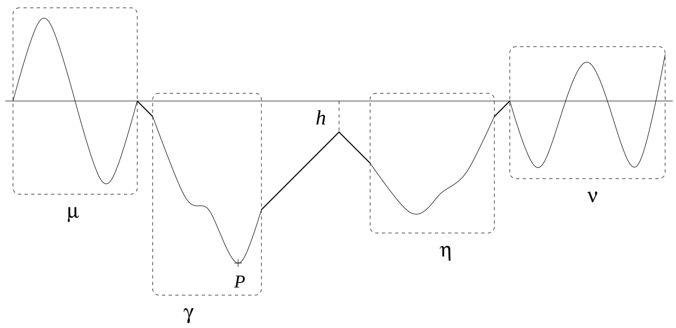


Figure 2.38: A representation of the path ω in the case $h < j$ with $h < i$

Let P 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 P to the endpoint of the path by applying consecutive and appropriate mappings associated to $(k) \overset{1}{\rightsquigarrow} (k+1) \dots (2)(1)$ of the first production of (2.19) and by applying consecutive and appropriate mappings associated to $(k) \overset{j}{\rightsquigarrow} (\overline{k+j-i}) \dots (\overline{2})(\overline{1})$ of the second production of (2.19); add the path in $\gamma x^j \overline{x}^i \eta x$ running from its initial point to P by applying consecutive and appropriate mappings associated to $(k) \overset{1}{\rightsquigarrow} (k+1)\dots(2)(1)$ of the first production of (2.19) and by applying consecutive and appropriate mappings associated to $(k) \overset{j}{\rightsquigarrow} (\overline{k+j-i})\dots(\overline{2})(\overline{1})$ of the second production of (2.19); 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 (2.19). 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. ∎

3
Cross-bifix-free sets

Cross-bifix-free sets are sets of words such that no prefix of any word is a suffix of any other word. In this chapter, we introduce a general constructive method for the sets of cross-bifix-free binary words of fixed length. It enables us to determine a cross-bifix-free words subset which has the property to be non-expandable and whose cardinality is greater than the ones known in the literature till now [4].

This kind of sets play an important role in the field of Telecommunications but in the present work will be used to study strings avoiding set of different patterns.

3.1 Different areas involving

In digital communication systems, synchronization is an essential requirement to establish and maintain a connection between a transmitter and a receiver.

Analytical approaches to the synchronization acquisition process and methods for the construction of sequences with the best aperiodic autocorrelation properties [13, 50, 51, 67] have been the subject of numerous analyses in the digital transmission.

The historical engineering approach started with the introduction of bifix, a name proposed by J. L. Massey as acknowledged in [59]. It denotes a subsequence that is both a prefix and suffix of a longer observed sequence. Rather than to the bifix, much attention has been devoted to a bifix-indicator, an indicator function implying the existence of the bifix [59]. Such indicators were shown to be without equal in performing various statistical analysis, mainly concerning the search process [7, 59], whose goal is to find a fixed sequence in random data.

However, an analytical study of simultaneous search for a set of sequences urged the invention of cross-bifix indicators [4, 5] and, accordingly, turned attention to the so called cross-bifix-free sets, that is, the sets of words such that no prefix of any word is a suffix of any other word.

In [4], the author analyzes some properties of binary words that form a cross-bifix-free set, giving a constructive method.

This approach leads to sets $S(n)$ of cross-bifix-free binary words, of fixed length n, having cardinality $1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233$ for $n = 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15$, respectively.

This sequence forms a Fibonacci progression and satisfies the recurrence relation $|S(n)| = |S(n-1)| + |S(n-2)|$ with $|S(3)| = 1$ and $|S(4)| = 1$.

The problem of determining cross-bifix-free sets is also related to several other scientific applications, for instance in frame synchronization for multiaccess systems [6, 50, 70], pattern matching [26] and automata theory [18].

3.2 Cross-bifix-free words

Let A be a finite, non-empty alphabet and ψ be a word in A^*.

Let $\psi = vwu$ then v is called *prefix* of ψ and u is called *suffix* of ψ. A *bifix* of ψ is a subsequence of ψ that is both its prefix and suffix.

A word ψ of A^+ is said to be *bifix-free* or *unbordered* [43, 60] if and only if no strict prefix of ψ is also a suffix of ψ. Therefore, ψ is bifix-free if and only if $\psi \neq uwu$, being u any necessarily non-empty word and w any word. Obviously, a necessary condition for ψ to be bifix-free is that the first and the last letters of ψ must be different.

Example 2 *In the monoid* $\{0,1\}^*$*, the word* 111010100 *of length* $n = 9$ *is bifix-free, while the word* 101001010 *contains two bifixes,* 10 *and* 1010*.*

Let $BF_q(n)$ denote the set of all bifix-free words of length n over an alphabet of fixed size q. The following formula for the cardinality of $BF_q(n)$, denoted by $|BF_q(n)|$, is well-known [60].

$$\begin{cases} |BF_q(1)| = q \\ \\ |BF_q(2n+1)| = q|BF_q(2n)| \\ \\ |BF_q(2n)| = q|BF_q(2n-1)| - |BF_q(n)| \end{cases} \tag{3.1}$$

The number sequences related to this recurrence can be found in Sloane's database of integer sequences [71]: sequences A003000 ($q = 2$), A019308 ($q = 3$) and A019309 ($q = 4$).

Table 3.1 lists the set $BF_2(n)$, $2 \leq n \leq 6$, the last row reports the cardinality of each set.

Let $q > 1$ and $n > 1$ be fixed. Two distinct words $\psi, \psi' \in BF_q(n)$ are said to be *cross-bifix-free* [7] if and only if no strict prefix of ψ is also a suffix of ψ' and vice-versa.

Table 3.1: The set $BF_2(n)$, $2 \leq n \leq 6$

n=2	**n=3**	**n=4**	**n=5**	**n=6**
10 01	100 001	1000 0001	10000 00001	100000 000001
	110 011	1100 0011	10100 00101	101000 000101
		1110 0111	11000 00011	101100 001101
			11100 00111	110000 000011
			11010 01011	110100 001011
			11110 01111	111000 000111
				111100 001111
				110010 010011
				111010 010111
				111110 011111
2	4	6	12	20

Example 3 *The binary words* 111010100 *and* 110101010 *in* $BF_2(9)$ *are cross-bifix-free, while the binary words* 111001100 *and* 110011010 *in* $BF_2(9)$ *have the cross-bifix* 1100*.*

A subset of $BF_q(n)$ is said to be *cross-bifix-free set* if and only if for each ψ, ψ', with $\psi \neq \psi'$, in this set, ψ and ψ' are cross-bifix-free. This set is said to be *non-expandable* on $BF_q(n)$ if and only if the set obtained by adding any other word is not a cross-bifix-free set. A non-expandable cross-bifix-free set on $BF_q(n)$ having maximal cardinality is called *maximal cross-bifix-free set* on $BF_q(n)$.

Each word $\psi \in BF_2(n)$ 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 $BF_2(n)$, and a *fall step*, defined by $(1, -1)$ and denoted by $\overline{x}$, to each 0's in $BF_2(n)$, running from $(0, 0)$ to (n, h), $-n < h < n$.

From now on, we will refer interchangeably to words or their graphical representations on the Cartesian plane, that is paths.

The definition of bifix-free and cross-bifix-free can be easily extended to paths. Figure 3.1 shows the two paths corresponding to the cross-bifix-free words of Example 3.

Figure 3.1: Two paths in $BF_2(9)$ which are cross-bifix-free

A lattice path on the Cartesian plane using the steps $(1,1)$ and $(1,-1)$ and running from $(0,0)$ to $(2m,0)$, with $m \geq 0$, is said to be *Grand-Dyck* or *Binomial* path (see [23] for further details). A *Dyck* path is a sequence of rise steps and fall steps running from $(0,0)$ to $(2m,0)$ and remaining weakly above the x-axis (see Figure 3.2). The number of $2m$-length Dyck paths is the mth Catalan number $C_m = 1/(m+1)\binom{2m}{m}$, see [73] for further details.

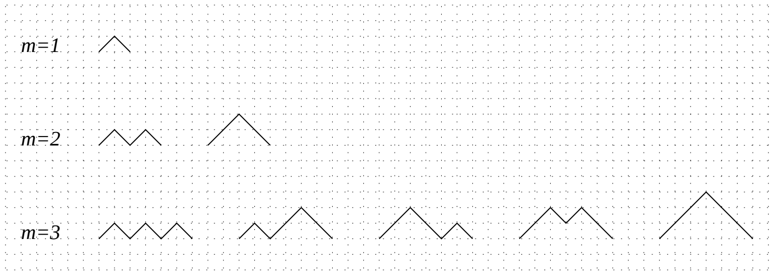


Figure 3.2: The $2m$-length Dyck paths, $1 \leq m \leq 3$

In the next section of the present chapter we are interested in determining one among all the possible non-expandable cross-bifix-free sets of words of fixed length $n > 1$ on the monoid $\{0,1\}^*$. We denote this set by $CBFS_2(n)$.

In order to do so, we focus on the set $\hat{BF}_2(n)$ of bifix-free binary words of fixed length n having 1 as the first letter and 0 as the last letter or equivalently the set of bifix-free lattice paths on the Cartesian plane using the steps $(1,1)$ and $(1,-1)$, running from $(0,0)$ to (n,h), $-n < h < n$, beginning with a rise step and ending with a fall step. Of course $\check{BF}_2(n) = BF_2(n) \backslash \hat{BF}_2(n)$ is obtained by switching rise and fall steps.

Let $\hat{BF}_2^h(n)$ denote the set of the paths in $\hat{BF}_2(n)$ having h as the ordinate of their endpoint, $-n < h < n$.

3.3 On the non-expandability of $CBFS_2(n)$

In order to prove that $CBFS_2(n)$ is a non-expandable cross-bifix-free set on $BF_2(n)$ we have to distinguish the following two cases depending on the parity of n.

3.3.1 Non-expandable $CBFS_2(2m+1)$

Let $CBFS_2(2m+1) = \{x\alpha : \alpha \in D_{2m}\}$ that is the set of paths beginning with a rise step linked to a $2m$-length Dyck path (see Figure 3.3). Note that $CBFS_2(2m+1)$ is a subset of $\hat{BF}_2^1(2m+1)$, $m \geq 1$.

Of course $|CBFS_2(2m+1)| = C_m$, being C_m the mth Catalan number, $m \geq 1$.

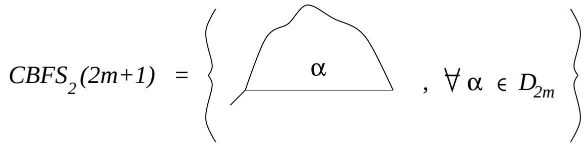


Figure 3.3: A representation of $CBFS_2(2m+1)$, with $m \geq 1$

Figure 3.4 shows the set $CBFS_2(7)$, with $|CBFS_2(7)| = C_3 = 5$.

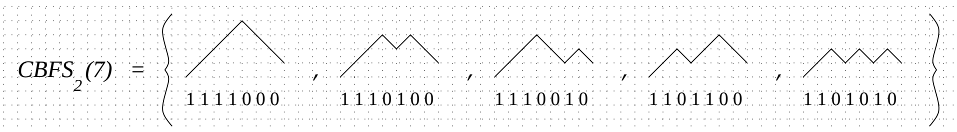


Figure 3.4: A graphical representation of $CBFS_2(7)$

Proposition 1 *$CBFS_2(2m+1)$ is a cross-bifix-free set on $BF_2(2m+1)$, $m \geq 1$.*

Proof. The proof consists of two distinguished steps. The first one proves that each $\psi \in CBFS_2(2m+1)$ is bifix-free and the second one proves that $CBFS_2(2m+1)$ is a cross-bifix-free set. Each $\psi \in CBFS_2(2m+1)$ can be written as $\psi = vwu$, being v, u any necessarily non-empty word while w can be an empty word also. For each prefix v of ψ we have $|v|_1 > |v|_0$ and for each suffix u of ψ we have $|u|_1 \leq |u|_0$. Therefore $v \neq u$, $\forall v, u \in \psi$ so ψ is bifix-free.

The proof that, for each $\psi, \psi' \in CBFS_2(2m+1)$ then ψ and ψ' are cross-bifix-free, is quite analogous to the one just illustrated, being $\psi = vwu$ and $\psi' = v'w'u'$ and comparing the prefix v of ψ and the suffix u' of ψ'. ■

Proposition 2 *$CBFS_2(2m+1)$ is a non-expandable cross-bifix-free set on $BF_2(2m+1)$, $m \geq 1$.*

Proof. It is sufficient to prove that the set $CBFS_2(2m+1)$ is a non-expandable cross-bifix-free set on $\hat{BF}_2(2m+1)$, as each $\psi \in CBFS_2(2m+1)$ and $\varphi \in \check{BF}_2(2m+1)$ match on the last letter of ψ and the first one of φ at least.

Let $m \geq 1$ be fixed, we can prove that $CBFS_2(2m+1)$ is a non-expandable cross-bifix-free set on $\hat{BF}^h_2(2m+1)$ by distinguishing $h > 0$ from $h < 0$.

$h > 0:$ a path γ in $\hat{BF}_2^h(2m+1)\backslash CBFS_2(2m+1)$ can be written as $\gamma = \phi x\alpha_1 x\alpha_2 x \ldots x\alpha_r$ (see Figure 3.5, where $n = 2m+1$), being ϕ a Grand-Dyck path beginning with a rise step, x a rise step, α_l Dyck paths, for each $l : 1 \leq l \leq r-1$, and α_r a necessarily non-empty Dyck path. Therefore, we can find paths in $CBFS_2(2m+1)$ having a prefix which matches with a suffix of γ. The path $\psi = x\alpha_r\delta$ in $CBFS_2(2m+1)$, being δ a Dyck path of appropriate length, has the prefix $x\alpha_r$ which matches with the suffix $x\alpha_r$ of γ.

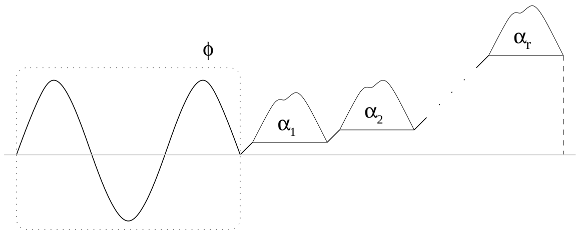


Figure 3.5: A graphical representation of a path γ in $\hat{BF}_2^h(n), h > 0$

$h < 0:$ a path γ in $\hat{BF}_2^h(2m+1)$ can be written as $\gamma = \alpha_r\overline{x}\alpha_{r-1}\overline{x}\ldots\overline{x}\alpha_1\overline{x}\phi$ (see Figure 3.6, where $n = 2m+1$), being α_r a necessarily non-empty Dyck path, $\overline{x}$ a fall step, α_l Dyck paths, for each $l : 1 \leq l \leq r-1$, and ϕ a Grand-Dyck path. Therefore, we can find paths in $CBFS_2(2m+1)$ having a suffix which matches with a prefix of γ. The path $\psi = x\delta\alpha_r$ in $CBFS_2(2m+1)$, being δ a Dyck path of appropriate length, has the suffix α_r which matches with the prefix α_r of γ.

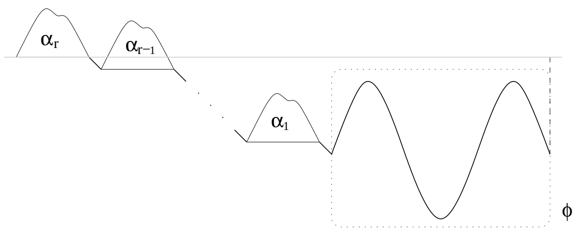


Figure 3.6: A graphical representation of a path γ in $\hat{BF}_2^h(n), h < 0$

■

3.3.2 Non-expandable $CBFS_2(2m+2)$

In this case we have to distinguish two further subcases depending on the parity of $m > 0$.

If m is an even number then $CBFS_2(2m+2) = \{\alpha x \beta \overline{x} : \alpha \in D_{2i}, \beta \in D_{2(m-i)}, 0 \leq i \leq \frac{m}{2}\}$, that is the set of paths consisting of the following consecutive sub-paths: a $2i$-length Dyck path, a rise step, a $2(m-i)$-length Dyck path, $0 \leq i \leq \frac{m}{2}$, a fall step (see Figure 3.7). Note that $CBFS_2(2m+2)$ is a subset of $\hat{BF}_2^0(2m+2)$, for any even number $m > 1$.

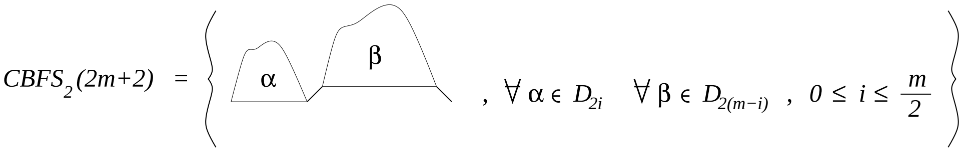


Figure 3.7: A representation of $CBFS_2(2m+2)$, for any even number $m > 1$

Of course $|CBFS_2(2m+2)| = \sum_{i=0}^{m/2} C_i C_{m-i}$, C_m is the mth Catalan number, for any even number $m > 1$. Figure 3.8 shows the set $CBFS_2(10)$, with $|CBFS_2(10)| = C_4 + C_1C_3 + C_2C_2 = 23$.

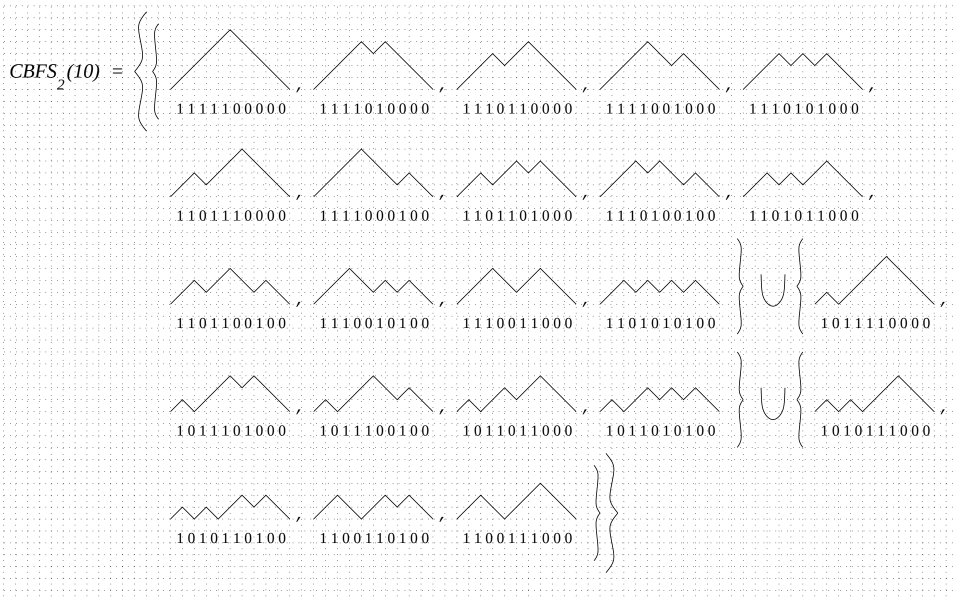


Figure 3.8: A graphical representation of $CBFS_2(10)$ giving evidence to the elements corresponding to C_4, C_1C_3 and C_2C_2 respectively

Proposition 3 *$CBFS_2(2m+2)$ is a cross-bifix-free set on $BF_2(2m+2)$, for any even number $m > 1$.*

Proof. The proof consists of two distinguished steps. The first one proves that each $\psi \in CBFS_2(2m+2)$ is bifix-free and the second one proves that

$CBFS_2(2m+2)$ is a cross-bifix-free set. Each $\psi \in CBFS_2(2m+2)$ can be written as $\psi = vwu$, being v, u any necessarily non-empty word while w can be an empty word also. Referring to Figure 3.7, let $m > 1$ be fixed, we have to take into consideration two different cases: in the first one $i = 0$ and in the second one $0 < i \leq \frac{m}{2}$.

If $i = 0$ then $\psi \in \{x\beta\overline{x} : \beta \in D_{2m}\}$, and for each prefix v of ψ we have $|v|_1 > |v|_0$ and for each suffix u of ψ we have $|u|_1 < |u|_0$. Therefore $v \neq u$, $\forall v, u \in \psi$, so ψ is bifix-free.

Otherwise, $\psi \in \{\alpha x\beta\overline{x} : \alpha \in D_{2i}, \beta \in D_{2(m-i)}, 0 < i \leq \frac{m}{2}\}$, then for each prefix v of ψ we have $|v|_1 \geq |v|_0$ and for each suffix u of ψ we have $|u|_1 \leq |u|_0$. If $|v|_1 > |v|_0$ then $v \neq u$, $\forall v, u \in \psi$ and therefore ψ is bifix-free. Let i be fixed, if $|v|_1 = |v|_0$ then the path v is a $2k$-length Dyck path, $1 \leq k \leq i$. In this case $\alpha = vv'$, so either $u = \mu\overline{x}$, where μ is any suffix of β, or $u = \mu' x\beta\overline{x}$, where μ' is any suffix of v'. If $u = \mu\overline{x}$ then $|u|_1 < |u|_0$, therefore $v \neq u$, $\forall v, u \in \psi$ and therefore ψ is bifix-free. If $u = \mu' x\beta\overline{x}$ then v does not match with $x\beta\overline{x}$, therefore $v \neq u$, $\forall v, u \in \psi$, so ψ is bifix-free.

The proof that, for each $\psi, \psi' \in CBFS_2(2m+2)$ then ψ and ψ' are cross-bifix-free, is quite analogous to the one just illustrated, being $\psi = vwu$ and $\psi' = v'w'u'$ and comparing the prefix v of ψ and the suffix u' of ψ'. ■

Proposition 4 *$CBFS_2(2m+2)$ is a non-expandable cross-bifix-free set on $BF_2(2m+2)$, for any even number $m > 1$.*

Proof. It is sufficient to prove that the set $CBFS_2(2m+2)$ is a non-expandable cross-bifix-free set on $\hat{BF}_2(2m+2)$, as each $\psi \in CBFS_2(2m+2)$ and $\varphi \in \check{BF}_2(2m+2)$ match on the last letter of ψ and the first one of φ at least.

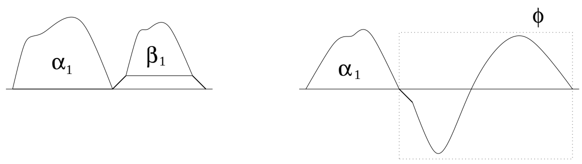
Let $m > 1$ be fixed, we have to take into consideration three different cases: in the first one we prove that $CBFS_2(2m+2)$ is a non-expandable cross-bifix-free set on $\hat{BF}^h_2(2m+2)$, $h > 0$, in the second one we prove that $CBFS_2(2m+2)$ is a non-expandable cross-bifix-free set on $\hat{BF}^h_2(2m+2)$, $h < 0$, and in the last one we prove that $CBFS_2(2m+2)$ is a non-expandable cross-bifix-free set on $\hat{BF}^0_2(2m+2)$.

$h > 0$: a path γ in $\hat{BF}^h_2(2m+2)$ can be written as $\gamma = \phi x\alpha_1 x\alpha_2 x \ldots x\alpha_r$ (see Figure 3.5, where $n = 2m+2$), being ϕ a Grand-Dyck path beginning with a rise step, x a rise step, α_l Dyck paths, for each $l : 1 \leq l \leq r-1$, and α_r a necessarily non-empty Dyck path. Therefore, we can find paths in $CBFS_2(2m+2)$ having a prefix which matches with a suffix of γ. The path $\psi = x\alpha_r\delta\overline{x}$ in $CBFS_2(2m+2)$, being δ a Dyck path of appropriate length, has the prefix $x\alpha_r$ which matches with the suffix $x\alpha_r$ of γ.

$h < 0$: a path γ in $\hat{BF}_2^h(2m+2)$ can be written as $\gamma = \alpha_r \overline{x} \alpha_{r-1} \overline{x} \dots \overline{x} \alpha_1 \overline{x} \phi$ (see Figure 3.6, where $n = 2m+2$), being α_r a necessarily non-empty Dyck path, $\overline{x}$ a fall step, α_l Dyck paths, for each $l : 1 \leq l \leq r-1$, and ϕ a Grand-Dyck path. Therefore, we can find paths in $CBFS_2(2m+2)$ having a suffix which matches with a prefix of γ. The path $\psi = x\delta\alpha_r\overline{x}$ in $CBFS_2(2m+2)$, being δ a Dyck path of appropriate length, has the suffix $\alpha_r\overline{x}$ which matches with the prefix $\alpha_r\overline{x}$ of γ.

$h = 0$: a path γ in $\hat{BF}_2^0(2m+2) \backslash CBFS_2(2m+2)$ either never falls below the x-axis or crosses the x-axis. In the first case, it can be written as $\gamma = \alpha_1 x \beta_1 \overline{x}$, where α_1 is a necessarily non-empty $2k$-length Dyck path and β_1 is a $2(m-k)$-length Dyck path, with $\frac{m}{2} + 1 \leq k \leq m$, see Figure 3.9 a). Therefore, we can find paths in $CBFS_2(2m+2)$ having a prefix which matches with a suffix of γ. The path $\psi = x\beta_1\overline{x}x\beta\overline{x}$ in $CBFS_2(2m+2)$, since $x\beta_1\overline{x} \in D_{2i}$ being $i = m-k+1$, has the prefix $x\beta_1\overline{x}$ which matches with the suffix $x\beta_1\overline{x}$ of γ.

If a path γ in $\hat{BF}_2^0(2m+2) \backslash CBFS_2(2m+2)$ crosses the x-axis then it can be written as $\gamma = \alpha_1 \phi$ where α_1 is a necessarily non-empty $2k$-length Dyck path, $1 \leq k \leq m$, and ϕ is a necessarily non-empty Grand-Dyck path beginning with a fall step, see Figure 3.9 b). Therefore, we can find paths in $CBFS_2(2m+2)$ having a suffix which matches with a prefix of γ. The path $\psi = x\delta\alpha_1\overline{x}$ in $CBFS_2(2m+2)$, being δ a Dyck path of appropriate length, has the suffix $\alpha_1\overline{x}$ which matches with the prefix $\alpha_1\overline{x}$ of γ.



a) b)

Figure 3.9: The two cases for a path γ in $\hat{BF}_2^0(2m+2) \backslash CBFS_2(2m+2)$, for any $m > 1$

■

If m is an odd number then $CBFS_2(2m+2) = \{\alpha x \beta \overline{x} : \alpha \in D_{2i}, \beta \in D_{2(m-i)}, 0 \leq i \leq \frac{m+1}{2}\} \backslash \{x\alpha'\overline{x}x\beta'\overline{x} : \alpha', \beta' \in D_{m-1}\}$, that is the set of paths consisting of the following consecutive sub-paths: a $2i$-length Dyck path, a rise step, a $2(m-i)$-length Dych path, $0 \leq i \leq \frac{m+1}{2}$, a fall step, and excluding those consisting of the following consecutive sub-paths: a rise step, a $(m-1)$-length Dyck path, a fall step followed by a rise step, a $(m-1)$-length Dyck path, a fall step (see Figure 3.10). In other words, the paths which result from the concatenation of two elevated Dyck paths of the same length must be excluded.

In particular, if $\alpha' = \beta'$ then the excluded paths are not bifix-free, otherwise if $\alpha' \neq \beta'$ then the excluded paths match with the paths $\{\alpha x \beta \overline{x} : \alpha \in D_{m+1}, \beta \in D_{m-1}\}$ in $CBFS_2(2m+2)$. Note that $CBFS_2(2m+2)$ is a subset of $\hat{BF}_2^0(2m+2)$, for any odd number $m \geq 1$.

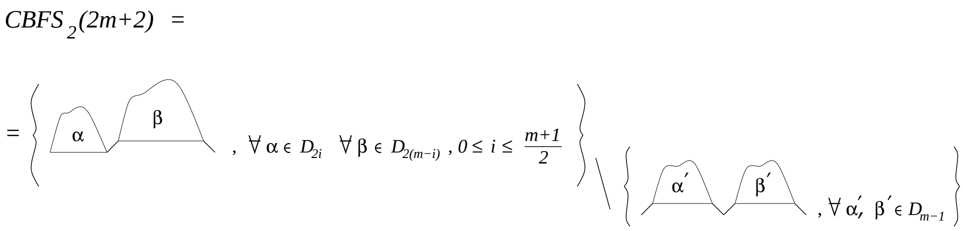


Figure 3.10: A representation of $CBFS_2(2m+2)$, for any odd number $m \geq 1$

Of course $|CBFS_2(2m+2)| = (\sum_{i=0}^{\frac{m+1}{2}} C_i C_{m-i}) - (C_{\frac{m-1}{2}})^2$, C_m is the mth Catalan number, for any odd number $m \geq 1$. Figure 3.11 shows the set $CBFS_2(8)$, with $|CBFS_2(8)| = (C_3 + C_1C_2 + C_2C_1) - (C_1)^2 = 8$.

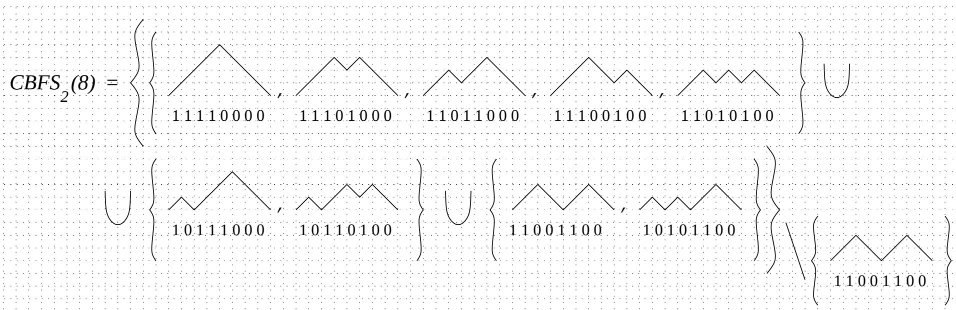


Figure 3.11: A graphical representation of the set $CBFS_2(8)$ giving evidence to the elements corresponding to C_3, C_1C_2 and C_2C_1 respectively

Proposition 5 *$CBFS_2(2m+2)$ is a cross-bifix-free set on $BF_2(2m+2)$, for any odd number $m \geq 1$.*

Proof. The proof consists of two distinguished steps. The first one proves that each $\psi \in CBFS_2(2m+2)$ is bifix-free and the second one proves that $CBFS_2(2m+2)$ is a cross-bifix-free set. Each $\psi \in CBFS_2(2m+2)$ can be written as $\psi = vwu$, being v, u any necessarily non-empty word while w can be an empty word also. Referring to Figure 3.10, let $m \geq 1$ be fixed, we have

to take into consideration three different cases: in the first one $i = 0$, in the second one $0 < i < \frac{m+1}{2}$ and in the last one $i = \frac{m+1}{2}$.
If $i = 0$ then $\psi \in \{x\beta\overline{x} : \beta \in D_{2m}\}$ and the proof that ψ is bifix-free is equal to the corresponding case of Proposition 3.
If $0 < i < \frac{m+1}{2}$ then $\psi \in \{\alpha x\beta\overline{x} : \alpha \in D_{2i}, \beta \in D_{2(m-i)}, 0 < i < \frac{m+1}{2}\}$ and the proof that ψ is bifix-free is equal to the case $0 < i \leq \frac{m}{2}$ of Proposition 3.
If $i = \frac{m+1}{2}$ then $\psi \in \{\alpha x\beta\overline{x} : \alpha \in D_{m+1}, \beta \in D_{m-1}\} \backslash \{x\alpha'\overline{x}x\beta'\overline{x} : \alpha', \beta' \in D_{m-1}\}$. For each prefix v of ψ we have $|v|_1 \geq |v|_0$ and for each suffix u of ψ we have $|u|_1 \leq |u|_0$. If $|v|_1 > |v|_0$ then $v \neq u$, $\forall v, u \in \psi$ and therefore ψ is bifix-free. If $|v|_1 = |v|_0$ then the path v is a $2k$-length Dyck path, for each $k : 1 \leq k \leq \frac{m+1}{2}$. If $1 \leq k \leq \frac{m-1}{2}$ the proof that ψ is bifix-free is equal to the case $0 < i \leq \frac{m}{2}$ of Proposition 3. Therefore, we have to solve only the case for $i = \frac{m+1}{2}$ and $k = \frac{m+1}{2}$. Since $\psi \notin \{x\alpha'\overline{x}x\beta'\overline{x} : \alpha', \beta' \in D_{m-1}\}$, necessarily $\alpha = v = v'v''$, where v' and v'' are necessary non-empty Dyck paths. Consequently v does not match with any suffix u of $x\beta\overline{x}$ and so ψ is bifix-free.

The proof that, for each $\psi, \psi' \in CBFS_2(2m + 2)$ then ψ and ψ' are cross-bifix-free, is quite analogous to the one just illustrated, being $\psi = vwu$ and $\psi' = v'w'u'$ and comparing the prefix v of ψ and the suffix u' of ψ'. ■

Proposition 6 *$CBFS_2(2m+2)$ is a non-expandable cross-bifix-free set on $BF_2(2m+2)$, for any odd number $m \geq 1$.*

Proof. It is sufficient to prove that the set $CBFS_2(2m + 2)$ is a non-expandable cross-bifix-free set on $\hat{BF}_2(2m+2)$, as each $\psi \in CBFS_2(2m+2)$ and $\varphi \in \check{BF}_2(2m + 2)$ match on the last letter of ψ and the first one of φ at least.

Let $m \geq 1$ be fixed, we have to take into consideration three different cases: in the first one we prove that $CBFS_2(2m + 2)$ is a non-expandable cross-bifix-free set on $\hat{BF}^h_2(2m + 2)$, $h > 0$, in the second one we prove that $CBFS_2(2m + 2)$ is a non-expandable cross-bifix-free set on $\hat{BF}^h_2(2m + 2)$, $h < 0$, and in the last one we prove that $CBFS_2(2m+2)$ is a non-expandable cross-bifix-free set on $\hat{BF}^0_2(2m + 2)$.

Both for $h > 0$ and $h < 0$ the proof that $CBFS_2(2m + 2)$ is a non-expandable cross-bifix-free set on $\hat{BF}^h_2(2m+2)$ is equal to the corresponding cases proved in Proposition 4.

For $h = 0$, a path γ in $\hat{BF}^0_2(2m + 2)\backslash CBFS_2(2m + 2)$ either never falls below the x-axis or crosses the x-axis.
In the first case, either $\gamma \in \{x\alpha'\overline{x}x\beta'\overline{x} : \alpha', \beta' \in D_{m-1}\}$ or $\gamma = \alpha_1 x\beta_1\overline{x}$, where the path α_1 is a necessarily non-empty $2k$-length Dyck path and β_1 is a $2(m - k)$-length Dyck path, with $\frac{m+1}{2} + 1 \leq k \leq m$, see Figure 3.9 a). If $\gamma \in \{x\alpha'\overline{x}x\beta'\overline{x} : \alpha', \beta' \in D_{m-1}\}$ then exist paths in $CBFS_2(2m + 2)$ having

a suffix which matches with a prefix of γ. The path $\delta x\alpha'\overline{x}$ in $CBFS_2(2m+2)$, where δ is a necessarily no elevated $(m+1)$-length Dyck path, has the suffix $x\alpha'\overline{x}$ which matches with the prefix $x\alpha'\overline{x}$ of γ. If $\gamma = \alpha_1 x\beta_1\overline{x}$, the path $\psi = x\beta_1\overline{x}x\beta\overline{x}$ in $CBFS_2(2m+2)$ has the prefix $x\beta_1\overline{x} \in D_{2i}$, $i = m-k+1$, which matches with the suffix of γ.

If a path γ in $\hat{BF}_2^0(2m+2)\backslash CBFS_2(2m+2)$ crosses the x-axis then it can be written as $\gamma = \alpha_1\phi$ where α_1 is a necessarily non-empty $2k$-length Dyck path, $1 \leq k \leq m$, and ϕ is a necessarily non-empty Grand-Dyck path beginning with a fall step, see Figure 3.9 b). Therefore, we can find paths in $CBFS_2(2m+2)$ having a suffix which matches with a prefix of γ. The path $\psi = x\delta\alpha_1\overline{x}$ in $CBFS_2(2m+2)$, being δ a Dyck path of appropriate length, has the suffix $\alpha_1\overline{x}$ which matches with the prefix $\alpha_1\overline{x}$ of γ. ∎

The presented constructive method gives sets $CBFS_2(n)$ of cross-bifix-free binary words, of fixed length n, having cardinality $1, 1, 2, 3, 5, 8, 14, 23, 42, 72, 132, 227, 429$ for $n =$ $3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15$, respectively.

4

Set of patterns avoidance

This chapter generalizes the study of Chapter 2. We are interested in studying the subclass $F^{[\mathcal{P}]}$ of F of binary words excluding the patterns $\mathfrak{p}_l = p_{l,0}p_{l,1}\dots p_{l,h_l-1} \in \{0,1\}^{h_l}$, where $p_{l,k}$ is the k-th letter of $\mathfrak{p}_l$, that is the words ω in $F^{[\mathcal{P}]}$ that do not admit a sequence of consecutive indices $j, j+1, \dots, j+h_l-1$ such that $\omega_j\omega_{j+1}\dots\omega_{j+h_l-1} = p_{l,0}p_{l,1}\dots p_{l,h_l-1}$, for any l such that $1 \leq l \leq m$. In order to do that the concept of cross-bifix-free set tackled in Chapter 3 is needed.

4.1 A generating algorithm for $F^{[\mathcal{P}_j]}$

In this section we define an algorithm for generating and enumerating the paths in the class $F^{[\mathcal{P}_j]}$ where $\mathcal{P}_j = \{\mathfrak{p}_1(j_1), \mathfrak{p}_2(j_2), \dots, \mathfrak{p}_m(j_m)\}$ is a cross-bifix-free set of patterns (see Chapter 3), none include in any other, such that each $\mathfrak{p}_l(j_l) = \mathfrak{p}_l$ begins with a rise step and $|\mathfrak{p}_l|_1 = |\mathfrak{p}_l|_0 + 1 = j_l + 1$, $1 \leq l \leq m$.

Let us observe that, since the patterns $\mathfrak{p}_l$ are cross-bifix-free and begins with a rise step, they necessarily end with a fall step. Moreover, we remark that it is not required to the patterns $\mathfrak{p}_l$ to have all the same length.

4.1.1 A construction for the class $F^{[\mathcal{P}_j]}$

The succession rule (2.10) given in Chapter 2 can be extended in order to obtain the generating algorithm for the class $F^{[\mathcal{P}_j]}$. The class $F^{[\mathcal{P}_j]}$ can be described by the following succession rule:

$$
(4.1)\qquad
\begin{cases}
(0) & \\
(k) & \overset{1}{\rightsquigarrow} (k+1)\cdots(1)(0_2)(0_1) \\
(k) & \overset{j_1+1}{\rightsquigarrow} (\overline{k+1})\cdots(\overline{1})(\overline{0_2})(\overline{0_1}) \\
(k) & \overset{j_2+1}{\rightsquigarrow} (\overline{k+1})\cdots(\overline{1})(\overline{0_2})(\overline{0_1}) \\
\vdots & \\
(k) & \overset{j_m+1}{\rightsquigarrow} (\overline{k+1})\cdots(\overline{1})(\overline{0_2})(\overline{0_1})
\end{cases}
$$

where the subscripts of labels (0) are simply used in order to distinguish the two labels one from each other, since they are obtained in two different ways in the generating process. Note that the labels (0_1) and (0_2) have the same set of successors regardless their subscripts.
The rule (4.1) can be represented as a tree having its root labelled (0) and where each node with label (k) at a given level n has $k+3$ sons at level $n+1$ labelled $(k+1), \cdots, (1), (0_2), (0_1)$ respectively, and $k+3$ sons at level $n+j_l+1$, $1 \leq l \leq m$, with labels $(\overline{k+1}), \cdots, (\overline{1}), (\overline{0_2}), (\overline{0_1})$ respectively. The generating algorithm associates a lattice path in $F^{[\mathcal{P}_j]}$ with a sequence of labels obtained by means of the succession rule (4.1). This gives a construction for the set $F^{[\mathcal{P}_j]}$ according to the number of rise steps or equivalently the number of ones.

The axiom (0) is associated to the empty path ε.
A path $\omega \in F$, with n rise steps and such that its endpoint has ordinate k, generates $k+3$ paths with $n+1$ rise steps, according to the first production of (4.1) having $k+1, \ldots, 1, 0, 0$ as endpoint ordinates, respectively.

In the similar way described in Section 2.2.1, the first $k+2$ paths are obtained by adding to ω a sequence of steps consisting of one rise step followed by $k+1-h$, $0 \leq h \leq k+1$, fall steps (see Figure 2.1 in Section 2.2.1). Each 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 and having a rise step as last step are never obtained. These paths have the label (0_1) of the first production in (4.1) and the way to obtain them will be described later in this section.

Let us denote by $L_l = (x_{L_l}, y_{L_l})$ and $R_l = (x_{R_l}, y_{R_l})$ the initial and last point of a pattern $\mathfrak{p}_l$, respectively, (see Figure 4.1). We define a *marked forbidden pattern* $\mathfrak{p}_l$ as a pattern $\mathfrak{p}_l = x\rho_l\overline{x} \in \mathcal{P}$, where $\rho_l \in F$ and $|\rho_l|_1 = |\rho_l|_0 + 1$, whose steps cannot be split, that is they must always be contained all together in that defined sequence. We say that a point is strictly contained in a given marked forbidden pattern $\mathfrak{p}_l$ if it is in $\mathfrak{p}_l$ and it is different from both L_l and R_l.

We denote a marked forbidden pattern $\mathfrak{p}_l$ by drawing its minimal bounding rectangle B_l. A rectangle B_l is like a *black box*, in the sense that it masks the included pattern $\mathfrak{p}_l$. A cut operation, i.e the procedure which splits a given path into two subpaths, is not possible within a marked forbidden pattern $\mathfrak{p}_l$. After a cut operation, it is not allowed to switch any rise step with a fall one, and viceversa, inside a marked forbidden pattern, but it can be translated.

A path $\omega \in F$, with n rise steps and such that its endpoint has ordinate k, generates $k+3$ paths, with $n+j_l+1$ rise steps, according to the production $(k) \overset{j_l+1}{\rightsquigarrow} (\overline{k+1}) \cdots (\overline{1})(\overline{0_2})(\overline{0_1})$ of (4.1), $1 \leq l \leq m$, having $k+1, \ldots, 1, 0, 0$ as endpoint ordinates, respectively. The first $k+2$ labels are obtained by

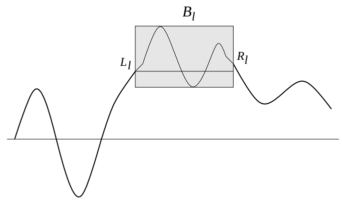


Figure 4.1: A marked forbidden pattern

adding to ω a sequence of steps consisting of the marked forbidden pattern $\mathfrak{p}_l$ followed by $k+1-h$, $0 \leq h \leq k+1$, fall steps (see Figure 4.2). Each path so obtained has the property that its rightmost suffix beginning from the x-axis, except the points in the rectangle B_l, either remains strictly above the x-axis itself or ends on the x-axis by a fall step. Let us note that, even if the path in the rectangle B_l intersects the x-axis, the initial point of the marked forbidden pattern, that is the point L_l, is always at ordinate $h \geq 0$.

At this point the label (0_1) due to the productions of (4.1) is associated to the paths which either do not contain marked forbidden patterns in its rightmost suffix and end on the x-axis by a rise step or having the initial point L_l in the rightmost marked forbidden pattern at ordinate $h < 0$.

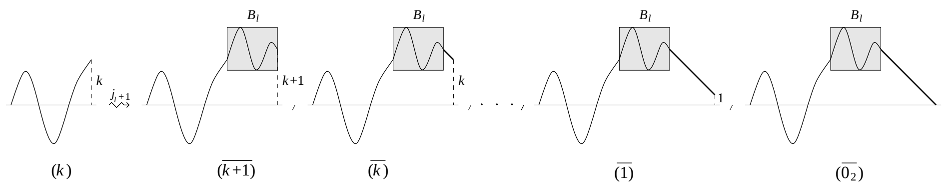


Figure 4.2: The mapping associated to $(k) \overset{j_l+1}{\rightsquigarrow} (\overline{k+1}) \ldots (\overline{1})(\overline{0_2})$ of (4.1)

In order to obtain the path labelled by (0_1) according to the first production of (4.1), we consider the path ω' obtained from ω by adding a sequence of steps consisting of one rise step followed by k fall steps. In order to obtain the path labelled by (0_1) according to every one of the other productions of (4.1), we consider the paths ω' obtained from ω by adding a sequence of steps consisting of the marked forbidden pattern $\mathfrak{p}_l$ 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 (see Figure 4.3). Clearly, in order to determine the suffix φ' of ω' we ignore the possible points on or below the x-axis which are within the black boxes.

In the similar way described in Section 2.2.1, if the suffix φ' does not contain any marked forbidden pattern, then the desired label (0_1) is asso-

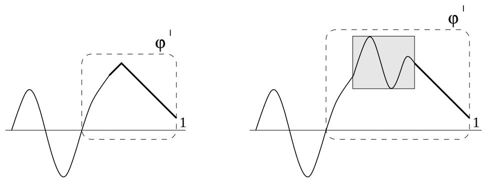


Figure 4.3: A graphical representation of the suffix φ' in ω'

ciated with the path $v(\varphi')^c x$, where $(\varphi')^c$ is the path obtained from φ' by switching rise and fall steps (see Figure 2.3 in Section 2.2.1).

If the suffix φ' contains marked forbidden patterns, let z be the leftmost point in φ' having highest ordinate and not strictly contained in a marked forbidden pattern. The desired label (0_1) is associated to the path obtained by concatenating to v a fall step $\overline{x}$ and then the path in φ' running from z to the endpoint of the path and the path running from the initial point in φ' to z. (see Figure 2.5 in Section 2.2.1).

This last mapping can be inverted as follows. Let d be the rightmost fall step in a path ω^* labelled (0_1) such that it begins from the x-axis and each point L_l of the marked forbidden patterns, on its right, has ordinate less than 0. Let $\omega^* = vd\varphi^*$ and P be the rightmost point in φ^* with lowest ordinate. The inverted lattice path of ω^* is given by $v\beta\alpha$, where β is the path in φ^* running from P to the endpoint of the path and α is the path running from the initial point in φ^* to P (see Figure 2.6 in Section 2.2.1).

Figure 4.4 shows the cut and paste actions related to the inverted mapping with the patterns $\mathfrak{p}_1 = x^2\overline{x}x\overline{x}$ and $\mathfrak{p}_2 = x^4\overline{x}^2x\overline{x}^2$ and Figure 4.5 shows the complete mapping defined by the succession rule (4.1) with the cross-bifix-free set $\mathcal{P}_j = \{\mathfrak{p}_1, \mathfrak{p}_2\}$.

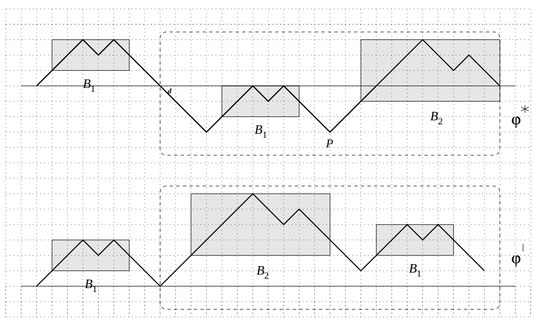


Figure 4.4: The inverted mapping related to the label (0_1) in case of marked forbidden patterns in φ'

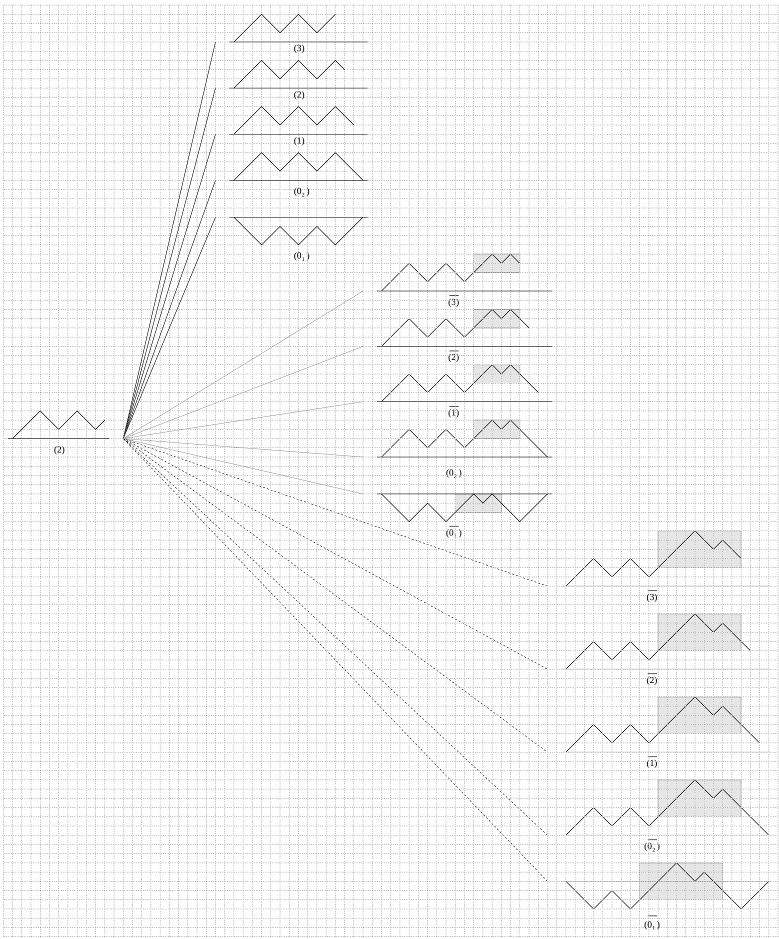


Figure 4.5: The set of lattice paths obtained from a given (k), by means of the succession rule (4.1)

4.1.2 Proving the construction

The above described construction generates 2^C copies of each path having C forbidden patterns such that 2^{C-1} instances are coded by a sequence of labels ending by a marked label, say $(\overline{k})$, and contain an odd number of marked forbidden patterns, and 2^{C-1} instances are coded by a sequence of labels ending by a non-marked label, say (k), and contain an even number of marked forbidden patterns. This is due to the fact that when a path is obtained according to the first production of (4.1) then no marked forbidden pattern is added. Moreover, when a path is obtained according to the other productions of (4.1) exactly one marked forbidden pattern is added. In any case, the actions performed to obtain the label (0_1) do not change the number of marked forbidden patterns in the path.

Theorem 4 *Let $\mathcal{P}_j = \{\mathfrak{p}_1(j_1), \ldots, \mathfrak{p}_m(j_m)\}$ be a cross-bifix-free set of forbidden patterns, none include in any other, such that each $\mathfrak{p}_l(j_l) = \mathfrak{p}_l$, $1 \leq l \leq m$, starts with a rise step and $|\mathfrak{p}_l|_1 = |\mathfrak{p}_l|_0 + 1 = j_l + 1$, $1 \leq l \leq m$. The generating tree of the paths in $F^{[\mathcal{P}_j]}$ according to the number of rise steps, is isomorphic to the tree having its root labelled by (0) and recursively defined by the succession rule (4.1).*

Proof. In order to prove the theorem we have to show that the algorithm described in the previous pages is a construction for the set $F^{[\mathcal{P}_j]}$ according to the number of rise steps. This means that all the paths in F with n rise steps are obtained once. Moreover, for each obtained path ω in $F \backslash F^{[\mathcal{P}_j]}$ with n rise steps, C forbidden patterns, and (k) as last label of the associated code, a path $\overline{\omega}$ in $F \backslash F^{[\mathcal{P}_j]}$ with n rise steps, C forbidden patterns and $(\overline{k})$ as last label of the associated code is also generated having the same form as ω but such that the last forbidden pattern is marked if it is not in ω and vice-versa.

The first assertion is a consequence of the construction according to the first production of (4.1).

In order to prove the second assertion we have to distinguish two cases depending on whether the last forbidden pattern $\mathfrak{p}_\ell$ is marked or not.

Let y_{L_ℓ} and y_{R_ℓ} be the ordinate of the initial point and ordinate the last point of $\mathfrak{p}_\ell$, respectively, and h_ℓ be the largest ordinate in $\mathfrak{p}_\ell$.

First case: the last forbidden pattern $\mathfrak{p}_\ell = x\rho_\ell\overline{x}$ in ω is marked. In the following we represent the marked forbidden pattern $\mathfrak{p}_\ell$ by its minimal bounding rectangle B_ℓ.

We consider the following subcases:

$y_{L_\ell} \geq 0$: The path ω in $F \backslash F^{[\mathcal{P}_j]}$ can be written as $\omega = \mu \mathfrak{p}_\ell \nu$, where $\mu \in F$, $\nu \in F^{[\mathcal{P}_j]}$ (see Figure 4.6).

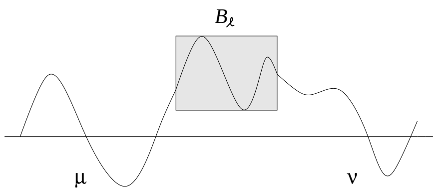


Figure 4.6: A representation of the path ω in the case $y_{L_\ell} \geq 0$

The path $\overline{\omega}$ which kills ω is obtained by adding on μ the path $x\rho_\ell\overline{x}\nu$ by applying consecutive and appropriate mappings of the first production of (4.1).

$y_{L_\ell} < 0$: In this case we distinguish the following two subcases: $h_\ell \geq 0$ and $h_\ell < 0$

$h_\ell \geq 0$: The path ω in $F \backslash F^{[\mathcal{P}_j]}$ can be written as $\omega = \mu\overline{x}\gamma\mathfrak{p}_\ell\,\nu$, where $\mu, \gamma \in F$, $\nu \in F^{[\mathcal{P}_j]}$ (see Figure 4.7).

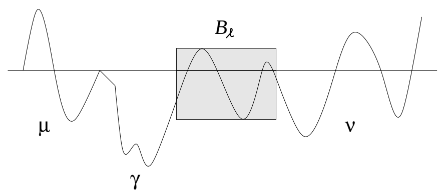


Figure 4.7: A representation of the path ω in the case $y_{L_\ell} < 0$ with $h_\ell \geq 0$

In this case exists at least one point intersecting the x-axis which is contained in $\mathfrak{p}_\ell$.

Let ς be the prefix of the forbidden (but no marked) pattern $\mathfrak{p}_\ell$ running from point L_ℓ to the leftmost point in which $\mathfrak{p}_\ell$ meets with the x-axis. The path ϖ which kills ω is obtained by adding to $\mu\overline{x}\gamma\varsigma$ the path $\varsigma'\nu$ by applying consecutive and appropriate mappings of the first production of (4.1), where ς' is the suffix of $\mathfrak{p}_\ell$ running from the endpoint of ς to the endpoint of $\mathfrak{p}_\ell$.

$h_\ell < 0$: The path ω in $F \backslash F^{[\mathcal{P}]}$ can be written as $\omega = \mu\overline{x}\gamma\mathfrak{p}_\ell\,\eta x\nu$, where $\mu, \gamma \in F$ and $\eta, \nu \in F^{[\mathcal{P}]}$ (see Figure 4.8).

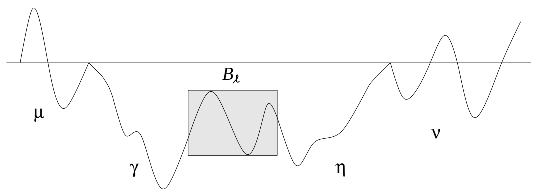


Figure 4.8: A representation of the path ω in the case $y_{L_\ell} < 0$ with $h_\ell < 0$

We observe that the path γ can contain marked forbidden patterns, with endpoints at ordinate less than 0. If the path γ contains no marked forbidden patterns, then it remains strictly under the x-axis, otherwise each marked forbidden pattern in γ intersects the x-axis when its largest ordinate is greater than 0. Moreover, the path η remains strictly under the x-axis. We distinguish two subcases.

In the first one the path γ contains no marked forbidden patterns and remains strictly under the x-axis. The path ϖ which kills ω is obtained by performing on μ the following actions: add the path $\overline{x}\gamma x\rho_\ell\overline{x}\eta x$ by applying consecutive and appropriate mappings of the first production of (4.1), apply the actions giving the label

(0_1) in case of no marked forbidden patterns. The path ν in $\overline{\omega}$ is obtained as in ω.

In the latter subcase the path γ contains at least a marked forbidden pattern. We consider the rightmost point P of the path $\overline{x}\gamma\mathfrak{p}_\ell\,\eta x$ with lowest ordinate. The path $\overline{\omega}$ which kills ω is obtained by performing on μ the following actions: add the path in $\gamma x\rho_\ell\overline{x}\eta x$ running from P to its endpoint by applying consecutive and appropriate mappings of the productions of (4.1), add the path in $\gamma x\rho_\ell\overline{x}\eta x$ running from its initial point to P by applying consecutive and appropriate mappings of the productions of (4.1), apply the cut and paste actions giving the label (0_1) in case of marked forbidden patterns. Obviously, the last forbidden pattern in the path must be generated by applying consecutive and appropriate mappings of the first production of (4.1). The path ν in $\overline{\omega}$ is obtained as in ω.

Second case: the last forbidden pattern $\mathfrak{p}_\ell$ in ω is not a marked forbidden pattern. We consider the subcases: $y_{L_\ell} \geq 0$ and $y_{L_\ell} < 0$.

$y_{L_\ell} \geq 0$: The path ω in $F\backslash F^{[\mathcal{P}_j]}$ can be written as $\omega = \mu\mathfrak{p}_\ell\,\nu$, where $\mu \in F$, $\nu \in F^{[\mathcal{P}_j]}$ (see Figure 4.9).

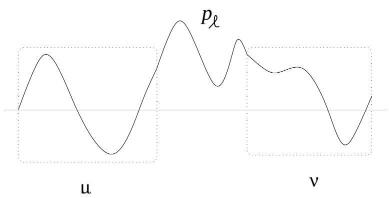


Figure 4.9: A representation of the path ω in the case $y_{L_\ell} \geq 0$

The path $\overline{\omega}$ which kills ω is obtained by adding on μ the path $\mathfrak{p}_\ell\nu$ by applying an appropriate mapping of $(k) \overset{j_\ell+1}{\rightsquigarrow} (\overline{k+1})\cdots(\overline{1})(\overline{0_2})$, that is the production generating the last marked forbidden pattern $\mathfrak{p}_\ell$ in $\overline{\omega}$, and consecutive and appropriate mappings of the first production of (4.1).

$y_{L_\ell} < 0$: The path ω in $F\backslash F^{[\mathcal{P}_j]}$ can be written as $\omega = \mu\overline{x}\gamma\mathfrak{p}_\ell\,\eta x\nu$, where $\mu, \gamma \in F$ and $\eta, \nu \in F^{[\mathcal{P}_j]}$ (see Figure 4.10).

We consider the rightmost point P of the path $\overline{x}\gamma\mathfrak{p}_\ell\,\eta x$ with lowest ordinate. The path $\overline{\omega}$ which kills ω is obtained by performing on μ the following actions: add the path in $\gamma\mathfrak{p}_\ell\eta x$ running from P to its endpoint by applying consecutive and appropriate mappings of the productions of (4.1), add the path in $\gamma\mathfrak{p}_\ell\eta x$ running from its initial

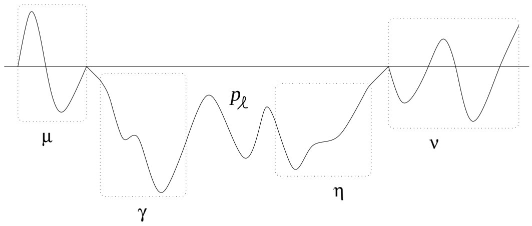


Figure 4.10: A representation of the path ω in the case $y_{L_\ell} < 0$

point to P by applying consecutive and appropriate mappings of the productions of (4.1), apply the cut and paste actions giving the label (0_1) in case of marked forbidden patterns. Obviously, the last marked forbidden pattern $\mathfrak{p}_\ell$ in $\overline{\omega}$ is generated by an appropriate mapping of the production $(k) \overset{j_\ell+1}{\rightsquigarrow} (\overline{k+1}) \cdots (\overline{1})(\overline{0_2})$. The path ν in $\overline{\omega}$ is obtained as in ω.

Note that, if the last point R_ℓ of $\mathfrak{p}_\ell$ in ω is on the x-axis then the path ω in $F \backslash F^{[\mathcal{P}_j]}$ can be written as $\omega = \mu\overline{x}\gamma\mathfrak{p}_\ell\nu$ and operations above described are applied on the path $\gamma\mathfrak{p}_\ell$.

We observe that for each path ω in $F \backslash F^{[\mathcal{P}_j]}$ with n rise steps, C forbidden patterns and last label (k), there exists one and only one path $\overline{\omega}$ in $F \backslash F^{[\mathcal{P}_j]}$ with n rise steps, C forbidden patterns and last label $(\overline{k})$ having the same form as ω but such that the last forbidden pattern is marked if it is not in ω and vice-versa.

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

4.1.3 Enumeration of $F^{[\mathcal{P}_j]}$

In order to obtain the enumeration of the class $F^{[\mathcal{P}_j]}$ according to the number of rise steps, we use a standard method, called ECO-method, for the enumeration of combinatorial objects which admit recursive descriptions in terms of generating trees, see [8, 32].

Let N be the set of paths generated by the algorithm described in Section 4.1 whose instances are coded by a sequence of labels in the generating tree ending by a non-marked one and M be the set of instances coded by a sequence of labels ending by a marked one. Then $F^{[\mathcal{P}_j]} = N \backslash M$.

The paths in N with n rise steps are obtained from the paths in N with $n-1$ rise steps by means of the first production of (4.1) and from those in

M with $n-j_l-1$, $1\leq l\leq m$, rise steps by means of the other productions of (4.1).

The paths in M with n rise steps are obtained from the paths in M with $n-1$ rise steps by means of the first production of (4.1) and from those in N with $n-j_l-1$, $1\leq l\leq m$, rise steps by means of the other productions of (4.1).

So, given a path $\omega\in F$ with $n(\omega)$ rise steps and ending point at ordinate $h(\omega)$, from the succession rule (4.1) we have:

$$\begin{aligned}
N(x,y) &= 1+\sum_{\omega\in N}\left(\sum_{i=0}^{h(\omega)+1}x^{n(\omega)+1}y^i+x^{n(\omega)+1}y^0\right)+\\
&\quad+\sum_{\omega\in M}\left(\sum_{i=0}^{h(\omega)+1}x^{n(\omega)+j_1+1}y^i+x^{n(\omega)+j_1+1}y^0\right)+\\
&\quad\vdots\\
&\quad+\sum_{\omega\in M}\left(\sum_{i=0}^{h(\omega)+1}x^{n(\omega)+j_m+1}y^i+x^{n(\omega)+j_m+1}y^0\right)
\end{aligned}$$

$$\begin{aligned}
M(x,y) &= \sum_{\omega\in M}\left(\sum_{i=0}^{h(\omega)+1}x^{n(\omega)+1}y^i+x^{n(\omega)+1}y^0\right)+\\
&\quad+\sum_{\omega\in N}\left(\sum_{i=0}^{h(\omega)+1}x^{n(\omega)+j_1+1}y^i+x^{n(\omega)+j_1+1}y^0\right)+\\
&\quad\vdots\\
&\quad+\sum_{\omega\in N}\left(\sum_{i=0}^{h(\omega)+1}x^{n(\omega)+j_m+1}y^i+x^{n(\omega)+j_m+1}y^0\right)
\end{aligned}$$

Since $\sum_{\omega\in N}\left(\sum_{i=0}^{h(\omega)+1}x^{n(\omega)+1}y^i+x^{n(\omega)+1}y^0\right)=$

$=\sum_{\omega\in N}x^{n(\omega)+1}\left(\frac{y^{h(\omega)+2}-1}{y-1}+1\right)=\frac{xy^2}{y-1}N(x,y)-\frac{x}{y-1}N(x,1)+x\,N(x,1),$

going on in the same way with the other terms we obtain:

$$
\begin{aligned}
N(x,y) &= 1+\frac{xy^2}{y-1}N(x,y)-\frac{x}{y-1}N(x,1)+x\,N(x,1)+\\
&+\frac{x^{j_1+1}y^2}{y-1}M(x,y)-\frac{x^{j_1+1}}{y-1}M(x,1)+x^{j_1+1}M(x,1)+\\
&\vdots\\
&+\frac{x^{j_m+1}y^2}{y-1}M(x,y)-\frac{x^{j_m+1}}{y-1}M(x,1)+x^{j_m+1}M(x,1)
\end{aligned}
$$

$$
\begin{aligned}
M(x,y) &= \frac{xy^2}{y-1}M(x,y)-\frac{x}{y-1}M(x,1)+x\,N(x,1)+\\
&+\frac{x^{j_1+1}y^2}{y-1}N(x,y)-\frac{x^{j_1+1}}{y-1}N(x,1)+x^{j_1+1}N(x,1)+\\
&\vdots\\
&+\frac{x^{j_m+1}y^2}{y-1}N(x,y)-\frac{x^{j_m+1}}{y-1}N(x,1)+x^{j_m+1}N(x,1)
\end{aligned}
$$

Since $F_{j_1,\dots,j_m}(x,y)=N(x,y)-M(x,y)$ then

$$
\begin{aligned}
F_{j_1,\dots,j_m}(x,y) &= 1+\frac{xy^2}{y-1}F_{j_1,\dots,j_m}(x,y)-\frac{x}{y-1}F_{j_1,\dots,j_m}(x,1)+\\
&+x\,F_{j_1,\dots,j_m}(x,1)-\\
&-(x^{j_1+1}+\ldots+x^{j_m+1})\frac{y^2}{y-1}F_{j_1,\dots,j_m}(x,y)+\\
&+(x^{j_1+1}+\ldots+x^{j_m+1})\frac{1}{y-1}F_{j_1,\dots,j_m}(x,1)-\\
&-(x^{j_1+1}+\ldots+x^{j_m+1})F_{j_1,\dots,j_m}(x,1)
\end{aligned}
$$

and

$$
\begin{aligned}
(y-1-x(1-x^{j_1}-\ldots-x^{j_m})y^2)F_{j_1,\dots,j_m}(x,y) &= \\
= y-1-xF_{j_1,\dots,j_m}(x,1)+x(y-1)\,F_{j_1,\dots,j_m}(x,1)+&\\
+(x^{j_1+1}+\ldots+x^{j_m+1})F_{j_1,\dots,j_m}(x,1)-&\\
-(x^{j_1+1}+\ldots+x^{j_m+1})(y-1)F_{j_1,\dots,j_m}(x,1)&
\end{aligned}
$$

Going on and using the kernel method [8] we obtain the generating function $F_{j_1,\dots,j_m}(x)$ for the words $\omega\in F^{[\mathcal{P}_j]}$ according to the number of ones:

$$F_{j_1,\dots,j_m}(x)=F_{j_1,\dots,j_m}(x,1)=\frac{1-y_0(x)}{x(y_0(x)-2)(1-x^{j_1}-\ldots-x^{j_m})}$$

where

$$y_0(x) = \frac{1-\sqrt{1-4x(1-x^{j_1}-\ldots-x^{j_m})}}{2x(1-x^{j_1}-\ldots-x^{j_m})}.$$

Let us remark that the generating function $F_{j_1,\ldots,j_m}(x)$ depends only on the number of ones in each forbidden pattern, so it enumerates all the sets of binary words in F avoiding cross-bifix-free forbidden patterns with $j_1,\ldots,j_m$ ones independently from their shapes.

Example 4 *For any set $\mathcal{P}_j = \{\mathfrak{p}_1(3),\mathfrak{p}_2(5)\}$, the first numbers of the sequence enumerating the binary words in $F^{[\mathcal{P}_j]}$, according to the number of ones, are:* $1, 3, 10, 35, 123, 442, 1608, 5911, 21905, \ldots$

being

$$\begin{aligned} F_{3,5}(x) &= \frac{1-\frac{1-\sqrt{1-4x+4x^4+4x^6}}{2x(1-x^3-x^5)}}{x\left(\frac{1-\sqrt{1-4x+4x^4+4x^6}}{2x(1-x^3-x^5)}-2\right)(1-x^3-x^5)} = \\ &= \frac{2x(1-x^3-x^5)-(1-\sqrt{1-4x+4x^4+4x^6})}{x(1-x^3-x^5)(1-\sqrt{1-4x+4x^4+4x^6}-4x(1-x^3-x^5))} \end{aligned}$$

the associated generating function.

4.2 A generating algorithm for $F^{[\mathcal{P}_{j,j}]}$

In this section, we study the construction and the enumeration of the class $F^{[\mathcal{P}_{j,j}]}$ where $\mathcal{P}_{j,j} = \{\mathfrak{p}_1(j_1,j_1),\mathfrak{p}_2(j_2,j_2),\ldots,\mathfrak{p}_m(j_m,j_m)\}$ is a cross-bifix-free set of patterns (see Chapter 3), none include in any other, such that each $\mathfrak{p}_l(j_l,j_l) = \mathfrak{p}_l$ begins with a rise step and $|\mathfrak{p}_l|_1 = |\mathfrak{p}_l|_0 = j_l$, $1 \le l \le m$, which is a slight modification of study in Section 4.1.

4.2.1 A construction for the class $F^{[\mathcal{P}_{j,j}]}$

In this section, we define an algorithm to construct the set $F^{[\mathcal{P}_{j,j}]}$ where $\mathcal{P}_{j,j} = \{\mathfrak{p}_1(j_1,j_1),\mathfrak{p}_2(j_2,j_2),\ldots,\mathfrak{p}_m(j_m,j_m)\}$ is a cross-bifix-free set of patterns, none include in any other, such that each $\mathfrak{p}_l(j_l,j_l) = \mathfrak{p}_l$ begins with a rise step and $|\mathfrak{p}_l|_1 = |\mathfrak{p}_l|_0 = j_l$, $1 \le l \le m$. Let us observe that, since the patterns $\mathfrak{p}_l$ are cross-bifix-free and begins with a rise step, they necessarily end with a fall step. Moreover, we remark that it is not required to the patterns $\mathfrak{p}_l$ to have all the same length.

The growth of the set $F^{[\mathcal{P}_{j,j}]}$, according to the number of rise steps or equivalently the number of ones, can be described by the following jumping and marked succession rule:

$$(4.2) \qquad \begin{cases} (0) \\ (k) \overset{1}{\rightsquigarrow} (k+1)(k)\cdots(1)(0_2)(0_1) & k \geq 0 \\ (0) \overset{j_1}{\rightsquigarrow} (\overline{0_2}) \\ (k) \overset{j_1}{\rightsquigarrow} (\overline{k})(\overline{k-1})\cdots(\overline{1})(\overline{0_2})(\overline{0_1}) & k \geq 1 \\ (0) \overset{j_2}{\rightsquigarrow} (\overline{0_2}) \\ (k) \overset{j_2}{\rightsquigarrow} (\overline{k})(\overline{k-1})\cdots(\overline{1})(\overline{0_2})(\overline{0_1}) & k \geq 1 \\ \vdots \\ (0) \overset{j_m}{\rightsquigarrow} (\overline{0_2}) \\ (k) \overset{j_m}{\rightsquigarrow} (\overline{k})(\overline{k-1})\cdots(\overline{1})(\overline{0_2})(\overline{0_1}) & k \geq 1 \end{cases}$$

This rule can be represented as a tree having its root labelled (0) and where each node with label (k) at level n gives $k+3$ sons at level $n+1$ labelled $(k+1), \cdots, (1), (0_2), (0_1)$ and $k+2$ sons at level $n+j_l$, $1 \leq l \leq m$, with labels $(\overline{k}), \cdots, (\overline{1}), (\overline{0_2}), (\overline{0_1})$, if $k \geq 1$, or only one son with label $(\overline{0_2})$ at level $n+j_l$, $1 \leq l \leq m$, if $k = 0$. The generating algorithm associates a lattice path in $F^{[\mathcal{P}_{j,j}]}$ to a sequence of labels obtained by means of the succession rule (4.2). This give a construction for the set $F^{[\mathcal{P}_{j,j}]}$ according to the number of rise steps or equivalently the number of ones.

The axiom (0) is associated to the empty path ε.
A path $\omega \in F$, with n rise steps and such that its endpoint has ordinate k, generates $k+3$ paths with $n+1$ rise steps, according to the first production of (4.2) having $k+1, \ldots, 1, 0, 0$ as endpoint ordinates, respectively.

In the similar way described in Section 2.2.1, the first $k+2$ paths are obtained by adding to ω a sequence of steps consisting of one rise step followed by $k+1-h$, $0 \leq h \leq k+1$, fall steps (see Figure 2.1 in Section 2.2.1). Each 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 and having a rise step as last step are never obtained. These paths have the label (0_1) of the first production in (4.2).

Let us denote by $L_l = (x_{L_l}, y_{L_l})$ and $R_l = (x_{R_l}, y_{R_l})$ the initial and last point of a pattern $\mathfrak{p}_l$, respectively, (see Figure 4.11). We define a *marked forbidden pattern* $\mathfrak{p}_l$ as a pattern $\mathfrak{p}_l = x\rho_l\overline{x} \in \mathcal{P}$, where $\rho_l \in F$ and $|\rho_l|_1 = |\rho_l|_0$, whose steps cannot be split, that is they must always be contained all together in that defined sequence. We say that a point is strictly contained in a given marked forbidden pattern $\mathfrak{p}_l$ if it is in $\mathfrak{p}_l$ and it is different from both L_l and R_l.

We denote a marked forbidden pattern $\mathfrak{p}_l$ by drawing its minimal bounding rectangle B_l. A rectangle B_l is like a *black box*, in the sense that it masks the included pattern $\mathfrak{p}_l$. A cut operation, i.e the procedure which splits a given path into two subpaths, is not possible within a marked forbidden pattern $\mathfrak{p}_l$. After a cut operation, it is not allowed to switch any rise step with a fall one, and viceversa, inside a marked forbidden pattern, but it can be translated.

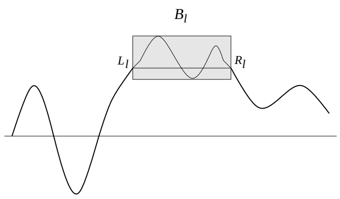


Figure 4.11: A marked forbidden pattern

A path $\omega \in F$, with n rise steps and such that its endpoint has ordinate 0, provides one path, with $n + j_l$ rise steps, $1 \leq l \leq m$, according to the $(0) \overset{j_l}{\rightsquigarrow} (\overline{0_2})$ production of (4.2). The obtained path has 0 as endpoint ordinate and it is obtained by adding to ω a sequence of steps consisting of the marked forbidden pattern $\mathfrak{p}_\mathfrak{l}$ in bounding rectangle B_l, see Figure 4.12.

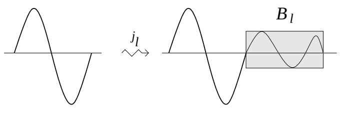


Figure 4.12: The mapping associated to $(0) \overset{j_l}{\rightsquigarrow} (\overline{0_2})$ of (4.2)

A path $\omega \in F$, with n rise steps and such that its endpoint has ordinate $k \geq 1$, provides $k + 2$ paths, with $n + j_l$ rise steps, $1 \leq l \leq m$, according to the $(k) \overset{j_l}{\rightsquigarrow} (\overline{k})(\overline{k-1}) \cdots (\overline{1})(\overline{0_2})(\overline{0_1})$ production of (4.2), having $k, \ldots, 1, 0, 0$ as endpoint ordinate, respectively. The first $k + 1$ labels are obtained by adding to ω a sequence of steps consisting of the marked forbidden pattern $\mathfrak{p}_\mathfrak{l}$ followed by $k - h$, $0 \leq h \leq k$, fall steps, see Figure 4.13.

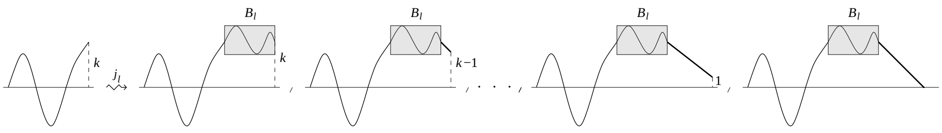


Figure 4.13: The mapping associated to $(k) \overset{j_l}{\rightsquigarrow} (\overline{k}) \ldots (\overline{1})(\overline{0_2})$ of (4.2)

At this point the label (0_1) due to the productions of (4.2) is associated with paths which either do not contain marked forbidden patterns in its rightmost suffix and end on the x-axis by a rise step or having the initial point L_l in the rightmost marked forbidden pattern at ordinate less than 0.

In order to obtain the path labelled by (0_1) according to the first production of (4.2), we consider the path ω' obtained from ω by adding a sequence of steps consisting of one rise step followed by k fall steps. In order to obtain the path labelled by (0_1) according to every one of the other productions of (4.2), we consider the paths ω' obtained from ω by adding a sequence of steps consisting of the marked forbidden pattern $\mathfrak{p}_l$ followed by $k-1$ 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.

As in Section 2.2.1, 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_1) 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 2.3 is Section 2.2.1).

If the suffix φ' contains marked points, then the desired label (0_1) is associated to the path obtained by applying cut and paste actions described in Section 2.2.1.

At this point, we have the complete mapping defined by the succession rule (4.2).

4.2.2 Proving the construction

The above construction generates 2^C copies of each path having C forbidden patterns such that 2^{C-1} instances are coded by a sequence of labels ending by a marked label, say $(\overline{k})$, and contain an odd number of marked forbidden patterns, and 2^{C-1} instances are coded by a sequence of labels ending by a non-marked label, say (k), and contain an even number of marked forbidden patterns. This is due to the fact that when a path is obtained according to the first production of (4.2) then no marked forbidden pattern is added. Moreover, when a path is obtained according to the other productions of (4.2) exactly one marked forbidden pattern is added. In any case, the actions performed to obtain the label (0_1) do not change the number of marked forbidden patterns in the path itself.

Theorem 5 *Let $\mathcal{P}_{j,j} = \{\mathfrak{p}_1(j_1,j_1), \mathfrak{p}_2(j_2,j_2), \ldots, \mathfrak{p}_m(j_m,j_m)\}$ be a cross-bifix-free set of patterns, none include in any other, such that each pattern $\mathfrak{p}_l(j_l,j_l) = \mathfrak{p}_l$ begins with a rise step and $|\mathfrak{p}_l|_1 = |\mathfrak{p}_l|_0 = j_l$, $1 \leq l \leq m$. The generating tree of the paths in $F^{[\mathcal{P}_{j,j}]}$, according to the number of rise steps, is isomorphic to the tree having its root labelled (0) and recursively defined by the succession rule (4.2).*

The proof of the Theorem 5 is analogous to the proof of Theorem 4 and it is omitted for brevity.

4.2.3 Enumeration of $F^{[\mathcal{P}_{j,j}]}$

As in Section 4.1, in order to obtain the enumeration of the class $F^{[\mathcal{P}_{j,j}]}$ according to the number of rise steps, we use a standard method, called ECO-method, for the enumeration of combinatorial objects which admit recursive descriptions in terms of generating trees, see [8, 32].

let Z be the set of paths whose instances are coded by a sequence of labels in the generating tree ending by a non-marked zero, S be the set of paths whose instances are coded by a sequence of labels ending by a marked zero, N be the set of paths whose instances are coded by a sequence of labels ending by a non-marked $k \geq 1$ and M be the set of paths whose instances are coded by a sequence of labels ending by a marked $k \geq 1$. Then $F^{[\mathcal{P}_{j,j}]} = (Z \backslash S) \cup (N \backslash M)$.

The succession rule (4.2) can be written as:

$$(4.3) \qquad \begin{cases} (0) \\ (0) \overset{1}{\rightsquigarrow} (1)(0_2)(0_1) \\ (k) \overset{1}{\rightsquigarrow} (k+1)(k)\cdots(1)(0_2)(0_1) & k \geq 1 \\ (0) \overset{j_1}{\rightsquigarrow} (\overline{0_2}) \\ (k) \overset{j_1}{\rightsquigarrow} (\overline{k})(\overline{k-1})\cdots(\overline{1})(\overline{0_2})(\overline{0_1}) & k \geq 1 \\ \vdots \\ (0) \overset{j_m}{\rightsquigarrow} (\overline{0_2}) \\ (k) \overset{j_m}{\rightsquigarrow} (\overline{k})(\overline{k-1})\cdots(\overline{1})(\overline{0_2})(\overline{0_1}) & k \geq 1 \end{cases}$$

Let us denote by $n(\omega)$ the number of rise steps of a path $\omega \in F$ and by $h(\omega)$ the last point's ordinate of ω itself. From the succession rule (4.3) we have:

$$\begin{aligned} Z(x,1) &= 1 + 2xZ(x,1) + 2xN(x,1) + (x^{j_1} + \cdots + x^{j_m})S(x,1) + \\ &\quad +2(x^{j_1} + \cdots + x^{j_m})M(x,1), \\ S(x,1) &= 2xS(x,1) + 2xM(x,1) + (x^{j_1} + \cdots + x^{j_m})Z(x,1) + \\ &\quad +2(x^{j_1} + \cdots + x^{j_m})N(x,1), \end{aligned}$$

$$\begin{aligned}
N(x,y) &= xyZ(x,1) + \sum_{\omega\in N}\sum_{i=1}^{h(\omega)+1} x^{n(\omega)+1}y^i + \sum_{\omega\in M}\sum_{i=1}^{h(\omega)} x^{n(\omega)+j_1}y^i + \cdots \\
&\cdots + \sum_{\omega\in M}\sum_{i=1}^{h(\omega)} x^{n(\omega)+j_m}y^i, \\
M(x,y) &= xyS(x,1) + \sum_{\omega\in M}\sum_{i=1}^{h(\omega)+1} x^{n(\omega)+1}y^i + \sum_{\omega\in N}\sum_{i=1}^{h(\omega)} x^{n(\omega)+j_1}y^i + \cdots \\
&\cdots + \sum_{\omega\in N}\sum_{i=1}^{h(\omega)} x^{n(\omega)+j_m}y^i.
\end{aligned}$$

Since $\sum_{\omega\in N}\sum_{i=1}^{h(\omega)+1} x^{n(\omega)+1}y^i = \sum_{\omega\in N} x^{n(\omega)+1}\left(\frac{y^{h(\omega)+2}-y}{y-1}\right) =$

$= \frac{xy^2}{y-1}N(x,y) - \frac{xy}{y-1}N(x,1)$ going on in the same way with the other terms, then we obtain:

$$\begin{aligned}
N(x,y) &= xyZ(x,1) + \frac{xy^2}{y-1}N(x,y) - \frac{xy}{y-1}N(x,1) + \\
&+ \frac{(x^{j_1}+\ldots+x^{j_m})y}{y-1}M(x,y) - \frac{(x^{j_1}+\ldots+x^{j_m})y}{y-1}M(x,1), \\
M(x,y) &= xyS(x,1) + \frac{xy^2}{y-1}M(x,y) - \frac{xy}{y-1}M(x,1) + \\
&+ \frac{(x^{j_1}+\ldots+x^{j_m})y}{y-1}N(x,y) - \frac{(x^{j_1}+\ldots+x^{j_m})y}{y-1}N(x,1).
\end{aligned}$$

Since $T(x,y) = N(x,y) - M(x,y)$ then:

$$\begin{aligned}
T(x,y) &= xy(Z(x,1)-S(x,1)) + \frac{(xy^2-(x^{j_1}+\ldots+x^{j_m})y)}{y-1}T(x,y) - \\
&- \frac{(xy-(x^{j_1}+\ldots+x^{j_m})y)}{y-1}T(x,1)
\end{aligned}$$

that is

$$T(x,y)(y-1-xy^2+(x^{j_1}+\ldots+x^{j_m})y) =$$
$$= xy(y-1)(Z(x,1)-S(x,1)) - (xy-(x^{j_1}+\ldots+x^{j_m})y)T(x,1).$$

Let

$$y_0(x) = \frac{1+x^{j_1}+\ldots+x^{j_m}-\sqrt{(x^{j_1}+\ldots+x^{j_m}+1)^2-4x}}{2x}$$

be a solution of $xy^2 - (x^{j_1} + \ldots + x^{j_m} + 1)y + 1 = 0$. Then we have the desired equation according to the number of ones, only:

$$T(x,1) = \frac{y_0(x) - 1}{1 - x^{j_1 - 1} + \ldots + x^{j_m - 1}}(Z(x,1) - S(x,1)).$$

Since

$$W(x,1) = Z(x,1) - S(x,1) = \frac{1 + 2x(1 - x^{j_1-1} + \ldots + x^{j_m-1})T(x,1)}{1 - 2x + x^{j_1} + \ldots + x^{j_m}},$$

then we have:

$$T(x,1) = \frac{y_0(x) - 1}{(1 - x^{j_1-1} - \cdots - x^{j_m-1})(1 + x^{j_1} + \cdots + x^{j_m} - 2xy_0(x))},$$

$$W(x,1) = \frac{1}{(1 + x^{j_1} + \cdots + x^{j_m} - 2xy_0(x))}.$$

Therefore the generating function $F^{jj}_{j_1,\cdots,j_m}(x) = T(x,1) + W(x,1)$ for the words $\omega \in F^{[\mathcal{P}_{j,j}]}$ according to the number of ones is:

$$F^{jj}_{j_1,\cdots,j_m}(x) = \frac{y_0(x) - x^{j_1-1} - \cdots - x^{j_m-1}}{(1 - x^{j_1-1} - \cdots - x^{j_m-1})(1 + x^{j_1} + \cdots + x^{j_m} - 2xy_0(x))}.$$

Let us remark that the generating function $F^{jj}_{j_1,\ldots,j_m}(x)$ depends only on the number of ones in each forbidden pattern, so it enumerates all the sets of binary words in F avoiding cross-bifix-free forbidden patterns with $j_1, \ldots, j_m$ ones (or zeroes) independently from their shapes.

Example 5 *For any set* $\mathcal{P}_{j,j} = \{\mathfrak{p}_1(4,4), \mathfrak{p}_2(5,5)\}$, *the first numbers of the sequence enumerating the binary words in* $F^{[\mathcal{P}_{j,j}]}$, *according to the number of ones, are:* $1, 3, 10, 35, 125, 453, 1663, 6166, 23037, \cdots$

being

$$F^{jj}_{4,5}(x) \quad = \quad \frac{\frac{1 + x^4 + x^5 - \sqrt{(x^4+x^5+1)^2 - 4x}}{2x} - x^3 - x^4}{(1 - x^3 - x^4)\sqrt{(x^4 + x^5 + 1)^2 - 4x}}$$

the associated generating function.

4.3 A generating algorithm for $F^{[\mathcal{P}_{j,i}]}$

In this section we focus on the generalization of cross-bifix-free set $\mathcal{P}$ of forbidden patterns none include in any other, passing from $\mathcal{P}_j = \{\mathfrak{p}_1(j_1), \mathfrak{p}_2(j_2), \ldots, \mathfrak{p}_m(j_m)\}$ such that each $\mathfrak{p}_l(j_l)$ begins with a rise step and has j_l+1 rise steps and j_l down steps, $1 \leq l \leq m$, to $\mathcal{P}_{j,i} = \{\mathfrak{p}_1(j_1, i_1), \mathfrak{p}_2(j_2, i_2), \ldots$ $\ldots, \mathfrak{p}_m(j_m, i_m)\}$ such that each $\mathfrak{p}_l(j_l, i_l)$ begins with a rise step and has j_l rise steps and i_l down steps, with $0 < i_l < j_l$ for any l, $1 \leq l \leq m$.

It is possible to adapt the algorithm constructing the class $F^{[\mathcal{P}_j]}$ to the class $F^{[\mathcal{P}_{j,i}]}$.

As in Chapter 2, also in this case the theory of Riordan arrays is not applicable to, neither the algorithmic approach allows to obtain their generating functions. Anyway it gives us a way to construct all the objects in this class.

4.3.1 A construction for the class $F^{[\mathcal{P}_{j,i}]}$

In this section, we propose an algorithm to construct the set $F^{[\mathcal{P}_{j,i}]}$ where $\mathcal{P}_{j,i} = \{\mathfrak{p}_1(j_1, i_1), \mathfrak{p}_2(j_2, i_2), \ldots, \mathfrak{p}_m(j_m, i_m)\}$ is a cross-bifix-free set of patterns, none include in any other, such that each $\mathfrak{p}_l(j_l, i_l) = \mathfrak{p}_l$ begins with a rise step and $|\mathfrak{p}_l|_1 = j_l$ and $|\mathfrak{p}_l|_0 = i_l$ with $0 < i_l < j_l$ for each l, $1 \leq l \leq m$.

Let us observe that, since the patterns $\mathfrak{p}_l$ are cross-bifix-free and begins with a rise step, they necessarily end with a fall step. Moreover, remark that it is not required that the patterns $\mathfrak{p}_l$ have the same length.

The study presented in Section 2.4 can be extended in order to obtain the generating algorithm for the class $F^{[\mathcal{P}_{j,i}]}$.

In particular, the growth of the class, according to the number of rise steps or equivalently the number of ones, also in this case 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.

Let us denote by $L_l = (x_{L_l}, y_{L_l})$ and $R_l = (x_{R_l}, y_{R_l})$ the initial and last point of a forbidden pattern $\mathfrak{p}_l$, respectively (see Figure 4.14).

We define a *marked forbidden pattern* $\mathfrak{p}_l$ as a pattern $\mathfrak{p}_l = x\varrho_l\overline{x} \in \mathcal{P}_{j,i}$, where $\varrho_l \in F$ such that $|\varrho_l|_1 = j_l - 1$ and $|\varrho_l|_0 = i_l - 1$, whose steps cannot be split, that is they must always be contained all together in that defined sequence. We say that a point is strictly contained in a given marked forbidden pattern $\mathfrak{p}_l$ if it is in $\mathfrak{p}_l$ and it is different from both L_l and R_l.

We denote a marked forbidden pattern $\mathfrak{p}_l$ by drawing its minimal bounding rectangle B_l.

A rectangle B_l is like a *black box*, in the sense that it masks the included pattern $\mathfrak{p}_l$. A cut operation, i.e the procedure which splits a given path into two subpaths, is not possible within a marked forbidden pattern $\mathfrak{p}_l$. After a cut operation, it is not allowed to switch any rise step with a fall one, and viceversa, inside a marked forbidden pattern, but it can be translated.

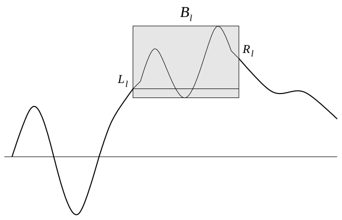


Figure 4.14: A marked forbidden pattern

In order to study the enumeration and the construction for the class $F^{[\mathcal{P}]}$, we have to distinguish two cases depending on the shape of the paths in F.

Definition 3 *A path ω in F is a* Δ-path *if:*

- *it ends on the x-axis (see Figure 4.15.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 $\mathfrak{p}_{\mathfrak{l}}$ (see Figure 4.15.b)) or not (see Figure 4.15.c)). If φ contains marked forbidden patterns $\mathfrak{p}_{\mathfrak{l}}$, then $y_{L_l} \geq 0$.*

Definition 4 *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 $\mathfrak{p}_{\mathfrak{l}}$ having ordinates $y_{L_l} < 0$ and $y_{R_l} > 0$ (see Figure 4.15.d)).*

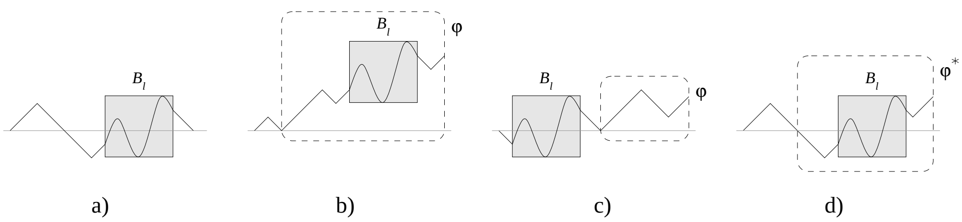


Figure 4.15: Some examples of paths in F

Δ-paths in F

For each Δ-path ω in F having k as the ordinate of its endpoint, we apply the succession rule (4.4), for each $k \geq 0$:

$$(4.4) \qquad \begin{cases} (0) \\ (k) \stackrel{1}{\rightsquigarrow} (k+1)(k)\cdots(2)(1)(0)^2 \\ (k) \stackrel{j_1}{\rightsquigarrow} (\overline{k+j_1-i_1})\ldots(\overline{s_1})(\overline{s_1-1})^2\ldots(\overline{1})^{s_1}(\overline{0})^{s_1+1} \\ (k) \stackrel{j_2}{\rightsquigarrow} (\overline{k+j_2-i_2})\ldots(\overline{s_2})(\overline{s_2-1})^2\ldots(\overline{1})^{s_2}(\overline{0})^{s_2+1} \\ \vdots \\ (k) \stackrel{j_m}{\rightsquigarrow} (\overline{k+j_m-i_m})\ldots(\overline{s_m})(\overline{s_m-1})^2\ldots(\overline{1})^{s_m}(\overline{0})^{s_m+1} \end{cases}$$

In the productions of (4.4), the parameter s_l, with $s_l \geq 0$ for each l, $1 \leq l \leq m$, is related to the shape of the Δ-path ω and the way to find s_l 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 (4.4).

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 (4.4) having $k+1, k, \ldots, 1, 0, 0$ as endpoint ordinate, respectively.

As in Section 2.2.1, the first $k+2$ 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 2.1 in Section 2.2.1.

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 (4.4).

In order to obtain the first label (0) according to the first production of (4.4), we consider the 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.

In the similar way described in Section 2.2.1, 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 2.3 in Section 2.2.1.

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 - described in Section 2.2.1 - 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, called α, and the path running from the initial point in φ' to z, called β, see Figure 2.5 in Section 2.2.1.

This last mapping can be inverted as in Section 2.2.1. In particular, 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 P 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 P to the endpoint of the path and α is the path φ^* running from the endpoint of d to P, see Figure 2.6 in Section 2.2.1.

Let the parameter s_l be fixed for each l, $1 \leq l \leq m$, a Δ-path $\omega \in F$, with n rise steps and such that its endpoint has ordinate k, provides $1 + k + j_l - i_l + \sum_{f_l=1}^{s_l} f_l$ lattice paths, with $n + j_l$ rise steps, according to the production $(k) \overset{j_l}{\rightsquigarrow} (\overline{k + j_l - i_l}) \dots \overline{(s_l)}(\overline{s_l - 1})^2 \dots (\overline{1})^{s_l}(\overline{0})^{s_l+1}$ of (4.4). The first $1 + k + j_l - i_l$ lattice paths have $k + j_l - i_l, \dots, s_l, s_l - 1, \dots, 1, 0$ as endpoint ordinate, respectively, and concerning the remaining $\sum_{f_l=1}^{s_l} f_l$ lattice paths each f_l of them has $s_l - f_l$ as endpoint ordinate, for each f_l, $0 \leq f_l \leq s_l$.

The first $1 + k + j_l - i_l$ lattice paths are obtained by adding to ω a sequence of steps consisting of the marked forbidden pattern $\mathfrak{p}_l$ followed by $k + j_l - i_l - h$ fall steps, for each h, $0 \leq h \leq k + j_l - i_l$, (see Figure 4.16).

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 $\sum_{f_l=1}^{s_l} f_l$ marked labels according to the production $(k) \overset{j_l}{\rightsquigarrow} (\overline{k + j_l - i_l}) \dots$ $\dots \overline{(s_l)}(\overline{s_l - 1})^2 \dots (\overline{1})^{s_l}(\overline{0})^{s_l+1}$ of (4.4), must give lattice paths having the rightmost marked forbidden pattern $\mathfrak{p}_l$ with ordinate $y_{L_l} < 0$.

In order to obtain such $\sum_{f_l=1}^{s_l} f_l$ marked labels, 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 $\mathfrak{p}_l$ followed by $k + j_l - i_l - f_l$ fall steps, for each f_l, $1 \leq f_l \leq s_l$. Therefore, we consider the just obtained paths labelled with $(\overline{f_l})$, for each f_l, $1 \leq f_l \leq s_l$, which are represented in Figure 4.16. By applying the previous actions, a path ω'' can be written as $\omega'' = \omega\mathfrak{p}_l\overline{x}^{k+j_l-i_l-f_l} = v\varphi\mathfrak{p}_l\overline{x}^{k+j_l-i_l-f_l} = v\varphi''$, $1 \leq f_l \leq s_l$, where φ'' is the rightmost suffix in w'' beginning from the x-axis and strictly remaining above the x-axis (see Figure 4.17).

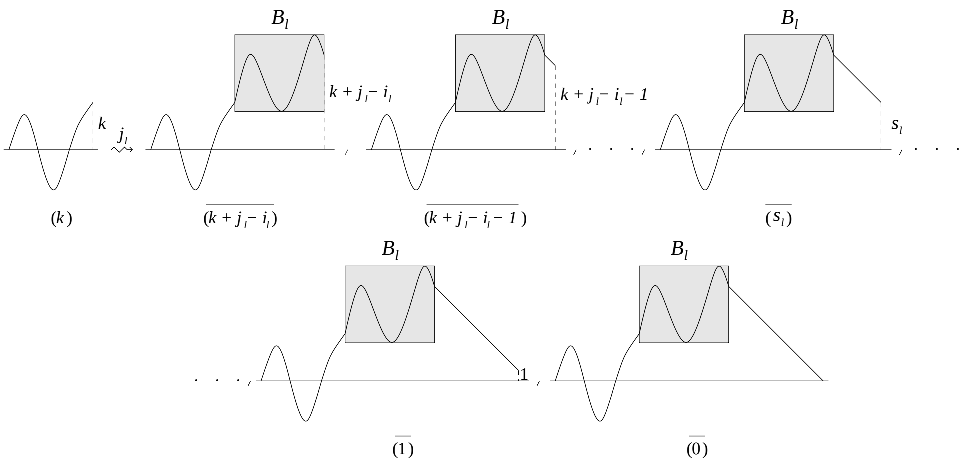


Figure 4.16: The mapping associated to $(k) \overset{j_l}{\rightsquigarrow} (\overline{k + j_l - i_l}) \ldots \overline{(s_l)} \ldots (\overline{1})(\overline{0})$ of (4.4)

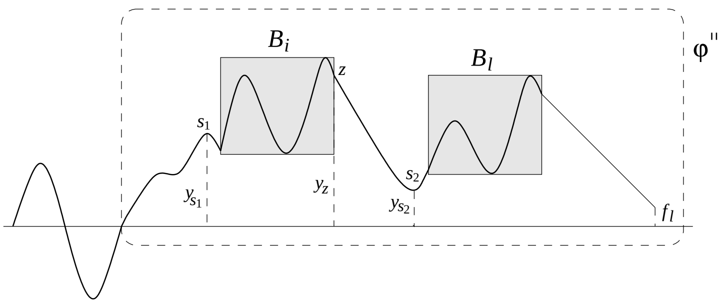


Figure 4.17: A graphical representation of the path $\omega'' = \omega \mathfrak{p}_l \overline{x}^{k+j_l-i_l-f_l} = v\varphi''$, $1 \leq f_l \leq s_l$

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_l in the production $(k) \overset{j_l}{\rightsquigarrow} (\overline{k + j_l - i_l}) \ldots \ldots \overline{(s_l)}(\overline{s_l - 1})^2 \ldots (\overline{1})^{s_l}(\overline{0})^{s_l+1}$ of (4.4) is $s_l = min\{y_z - y_{s_1}, y_{s_2}\}$. When z is contained in the suffix $\mathfrak{p}_l \overline{x}^{k+j_l-i_l-f_l}$ of ω'', $1 \leq f_l \leq s_l$, 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 P (see Figure 4.18).

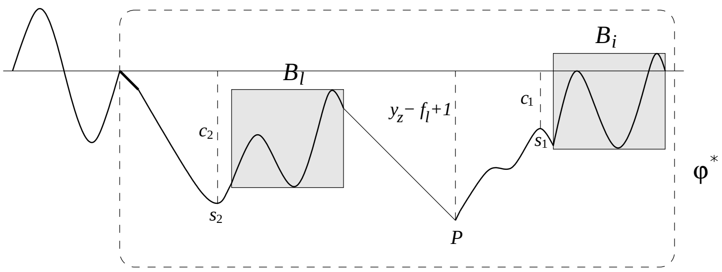


Figure 4.18: A graphical representation of the path $v\varphi^*$ obtained by applying the cut and paste actions to the path ω''

Remind that, from the reverse of the cut and paste actions, the point P 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 P 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 P 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 $c_1 = y_z - f_l + 1 - y_{s_1}$, must be greater than 0, that is $f_l < y_z - y_{s_1} + 1$. 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 $y_z - f_l + 1$, that is $f_l \leq y_{s_2}$. So $s_l = min\{y_z - y_{s_1}, y_{s_2}\}$ assures that the two conditions are verified as s_l is the upper value which can get f_l.

By performing the cut and paste actions on each ω'', we obtain s_l paths labelled $\overline{(f_l - 1)}$ for each f_l, $1 \leq f_l \leq s_l$. By adding g_l fall steps for each g_l, $0 < g_l \leq f_l - 1$, to each of such paths (see Figure 4.19), we obtain the complete mapping associated to the production $(k) \overset{j_l}{\rightsquigarrow} \overline{(k + j_l - i_l)} \ldots$ $\ldots \overline{(s_l)} \overline{(s_l - 1)}^2 \ldots \overline{(1)}^{s_l} \overline{(0)}^{s_l+1}$ of (4.4).

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 $\mathfrak{p}_l$ followed by f'_l fall steps, for each f'_l, $0 \leq f'_l < k + j_l - i_l - s_l$, we have already obtained paths.

Figure 4.20 shows the complete mapping associated to the succession rule (4.4) with cross-bifix-free set $\mathcal{P}_{j,i} = \{x^3\overline{x}^2, x^3\overline{x}x\overline{x}\}$.

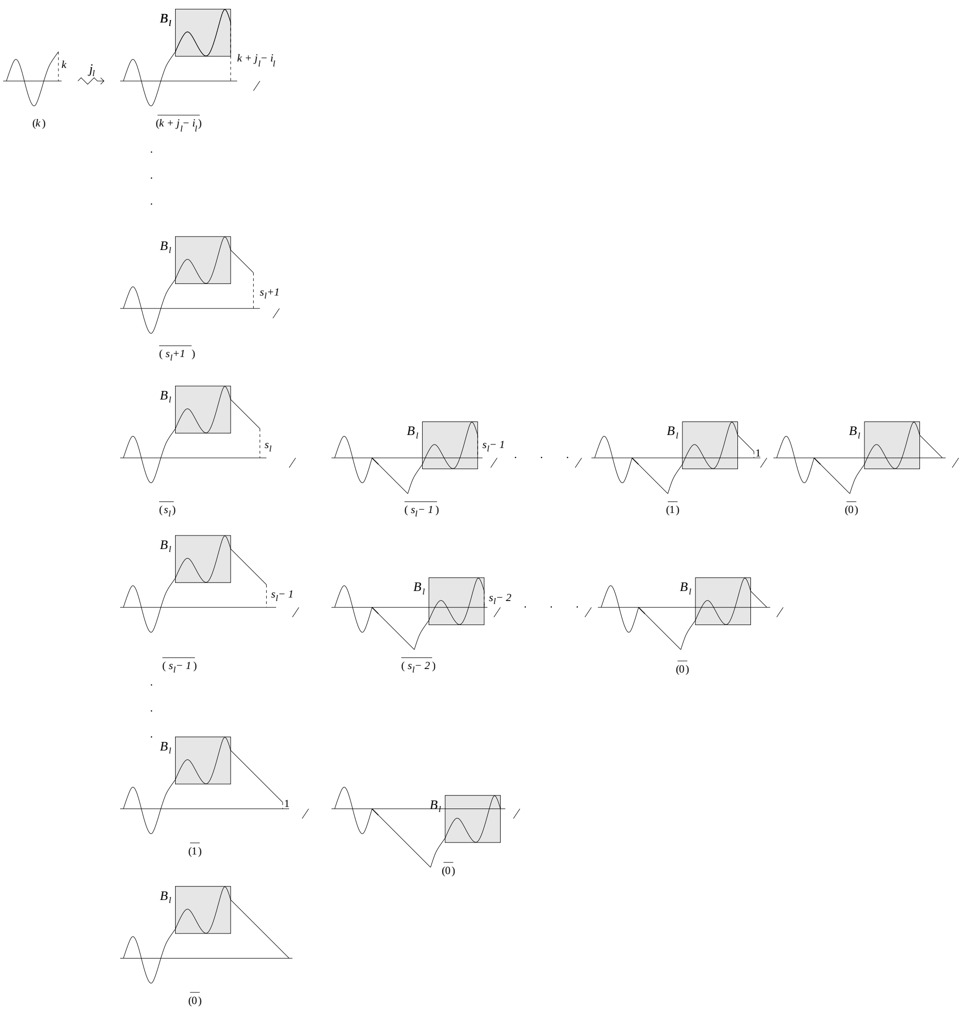


Figure 4.19: The mapping associated to $(k) \overset{j_l}{\rightsquigarrow} (\overline{k+j_l-i_l})\dots(\overline{s_l})(\overline{s_l-1})^2\dots$ $\dots(\overline{1})^{s_l}(\overline{0})^{s_l+1}$ of (4.4)

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 $k \geq 1$:

$$(4.5) \qquad \begin{cases} (k) \overset{1}{\rightsquigarrow} (k+1)(k)\cdots(2)(1)(0) \\ (k) \overset{j_1}{\rightsquigarrow} (\overline{k+j_1-i_1})(\overline{k+j_1-i_1-1})\cdots(\overline{2})(\overline{1})(\overline{0}) \\ (k) \overset{j_2}{\rightsquigarrow} (\overline{k+j_2-i_2})(\overline{k+j_2-i_2-1})\cdots(\overline{2})(\overline{1})(\overline{0}) \\ \quad\vdots \\ (k) \overset{j_m}{\rightsquigarrow} (\overline{k+j_m-i_m})(\overline{k+j_m-i_m-1})\cdots(\overline{2})(\overline{1})(\overline{0}) \end{cases}$$

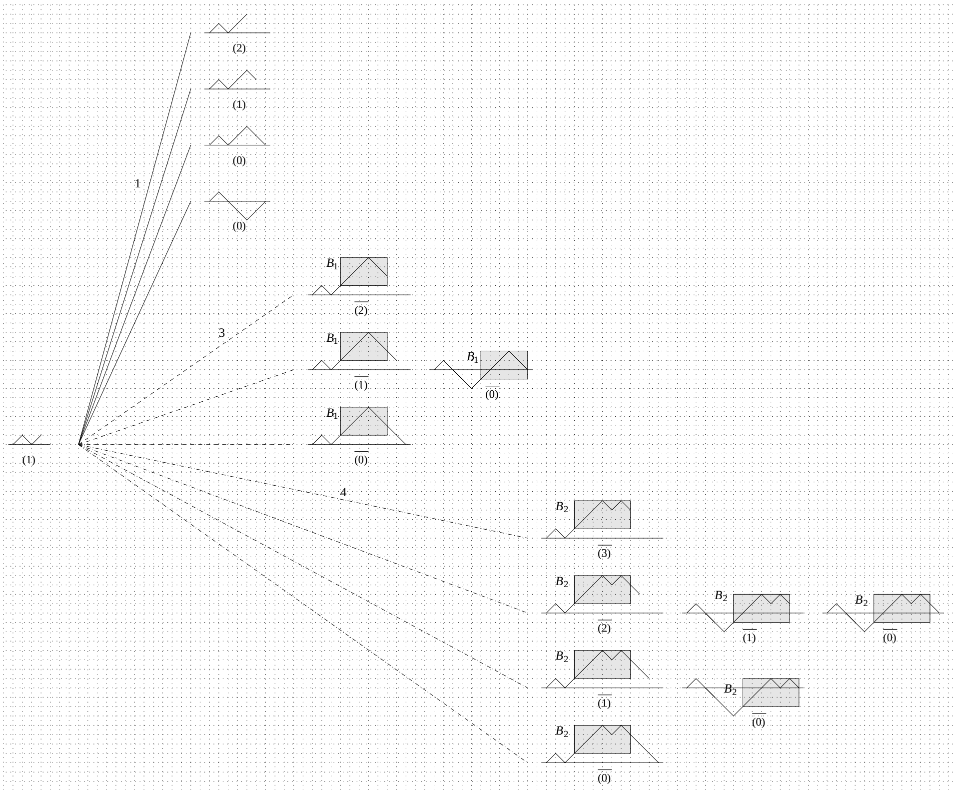


Figure 4.20: the complete mapping associated to the succession rule (4.4), being $\mathcal{P}_{j,i} = \{x^3\overline{x}^2, x^3\overline{x}x\overline{x}\}$

A Γ-path $\omega \in F$, with n rise steps and such that its endpoint has ordinate k, provides $k+2$ lattice paths, with $n+1$ rise steps, according to the first production of (4.5) having $k+1, \ldots, 1, 0$ 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, $0 \leq h \leq k+1$.

Moreover, a Γ-path $\omega \in F$, with n rise steps and such that its endpoint has ordinate k, provides $1+k+j_l-i_l$ lattice paths, with $n+j_l$ rise steps, according to the production $(k) \overset{j_l}{\rightsquigarrow} (\overline{k+j_l-i_l})(\overline{k+j_l-i_l-1})\cdots(\overline{2})(\overline{1})(\overline{0})$ of (4.5) having $k+j_l-i_l, \ldots, 2, 1, 0$ as endpoint ordinate, respectively, for each l, $1 \leq l \leq m$. These labels are obtained by adding to ω a sequence of steps consisting of the marked forbidden pattern $\mathfrak{p}_l$ followed by $k+j_l-i_l-h$ fall steps, $0 \leq h \leq k+j_l-i_l$, for each l, $1 \leq l \leq m$.

4.3.2 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 patterns, 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 patterns.

This observation is due to the fact that when a path is obtained either according to the first production of (4.4) or according to the first production of (4.5) then no marked forbidden pattern is added.

Moreover when a path is obtained either according to the production $(k) \overset{j_l}{\rightsquigarrow} (\overline{k+j_l-i_l}) \dots \overline{(s_l)}(\overline{s_l-1})^2 \dots (\overline{1})^{s_l}(\overline{0})^{s_l+1}$ of (4.4) or according to the production $(k) \overset{j_l}{\rightsquigarrow} (\overline{k+j_l-i_l})(\overline{k+j_l-i_l-1}) \cdots (\overline{2})(\overline{1})(\overline{0})$ of (4.5) for each l, $1 \leq l \leq m$, 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 (4.4) or the $\sum_{f_l=1}^{s_l} f_l$ marked labels, according to the production $(k) \overset{j_l}{\rightsquigarrow} (\overline{k+j_l-i_l}) \dots \overline{(s_l)}(\overline{s_l-1})^2 \dots (\overline{1})^{s_l}(\overline{0})^{s_l+1}$ of (4.4), for each l, $1 \leq l \leq m$, do not change the number of marked forbidden patterns in the path.

Theorem 6 *Let $\mathcal{P}_{j,i} = \{\mathfrak{p}_1(j_1, i_1), \mathfrak{p}_2(j_2, i_2), \dots, \mathfrak{p}_m(j_m, i_m)\}$ be a cross-bifix-free set of patterns, none include in any other, such that each pattern $\mathfrak{p}_l(j_l, i_l) = \mathfrak{p}_l$, $1 \leq l \leq m$, starts with a rise step and $|\mathfrak{p}_l|_1 = j_l$ and $|\mathfrak{p}_l|_0 = i_l$ with $0 < i_l < j_l$. The generating tree of the paths in $F^{[\mathcal{P}_{j,i}]}$, according to the number of rise steps, is isomorphic to the tree having the root labelled (0) and recursively defined by the succession rule (4.4), related to the shape of the path $\omega \in F$, and the succession rule (4.5).*

Proof. We have to show that the algorithm described in the previous pages is a construction for the set $F^{[\mathcal{P}_{j,i}]}$, 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 $F \backslash F^{[\mathcal{P}_{j,i}]}$ 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 $F \backslash F^{[\mathcal{P}_{j,i}]}$ 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, if the last label of the code associated to ω is (k) then 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 (4.4).

In order to prove the second assertion we have to distinguish two cases depending on whether the last forbidden pattern $\mathfrak{p}_\ell$ is marked or not.

Let y_{L_ℓ} and y_{R_ℓ} be the ordinate of the initial point and ordinate the last point of $\mathfrak{p}_\ell$, respectively and h_ℓ be the largest ordinate in $\mathfrak{p}_\ell$.

First case: the last forbidden pattern $\mathfrak{p}_\ell = x\varrho_\ell\overline{x}$ in ω is marked. In the following we represent the marked forbidden pattern $\mathfrak{p}_\ell$ by its minimal bounding rectangle B_ℓ.

We consider the following subcases:

$y_{L_\ell} \geq 0$: The path ω in $F \backslash F^{[\mathcal{P}_{j,i}]}$ can be written as $\omega = \mu \mathfrak{p}_\ell \, \eta \, \nu$, where $\mu \in F$ and $\eta, \nu \in F^{[\mathcal{P}_{j,i}]}$ (see Figure 4.21).

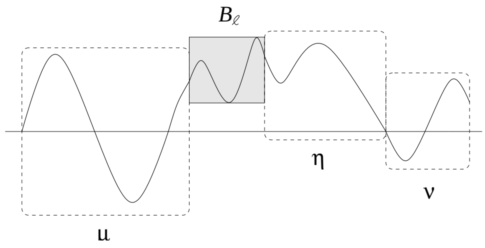


Figure 4.21: A representation of the path ω in the case $y_{L_\ell} \geq 0$

If μ is a Δ-path then the path $\overline{\omega}$ which kills ω is obtained by adding on μ the path $x \varrho_\ell \overline{x} \eta$ by applying consecutive and appropriate mappings of the first production of (4.4). Otherwise, if μ is a Γ-path then the path $\overline{\omega}$ which kills ω is obtained by adding on μ the path $x \varrho_\ell \overline{x} \eta$ by applying consecutive and appropriate mappings of the first production of (4.5). In each case the path ν in $\overline{\omega}$ is obtained as in ω.

$y_{L_\ell} < 0$: In this case we distinguish the following two subcases: $h_\ell \geq 0$ and $h_\ell < 0$

$h_\ell \geq 0$: The path ω in $F \backslash F^{[\mathcal{P}_{j,i}]}$ can be written as $\omega = \mu \overline{x} \gamma \mathfrak{p}_\ell \, \nu$, where $\mu, \gamma \in F$, $\nu \in F^{[\mathcal{P}_{j,i}]}$ (see Figure 4.22).

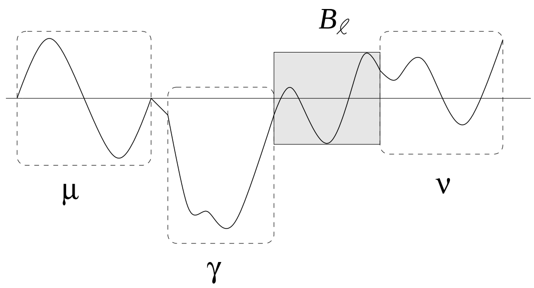


Figure 4.22: A representation of the path ω in the case $y_{L_\ell} < 0$ with $h_\ell \geq 0$

In this case exists at least one point intersecting the x-axis which is contained in $\mathfrak{p}_\ell$.

Let ς be the prefix of the forbidden (but no marked) pattern $\mathfrak{p}_\ell$ running from the point L_ℓ to the leftmost point in which $\mathfrak{p}_\ell$ meets with the x-axis. The path $\overline{\omega}$ which kills ω is obtained by adding to $\mu\overline{x}\gamma\varsigma$ the path $\varsigma'\nu$ by applying consecutive and appropriate mappings of the first production of (4.4), where ς' is the suffix of $x\varrho_\ell\overline{x}$ running from the endpoint of ς to the endpoint of $\mathfrak{p}_\ell$.

$h_\ell < 0$: The path ω in $F \backslash F^{[\mathcal{P}_{j,i}]}$ can be written as $\omega = \mu\overline{x}\gamma\mathfrak{p}_\ell\eta x\nu$, where $\mu, \gamma \in F$ and $\eta, \nu \in F^{[\mathcal{P}_{j,i}]}$ (see Figure 4.23).

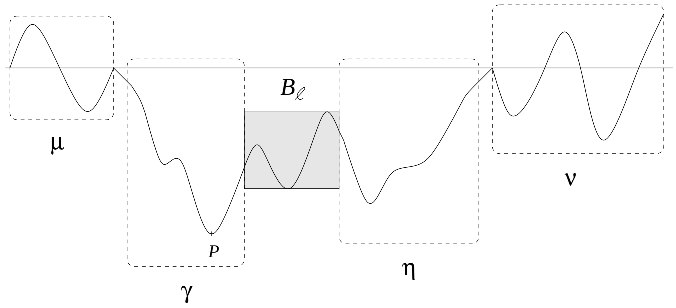


Figure 4.23: A representation of the path ω in the case $y_{L_\ell} < 0$ with $h_\ell < 0$

We observe that the path γ can contain marked forbidden patterns, with endpoints at ordinate less than 0, or not. If the path γ contains no marked forbidden patterns, then it remains strictly under the x-axis, otherwise the marked forbidden patterns intersect the x-axis when its largest ordinate is greater then 0. Moreover, the path η remains strictly under the x-axis. We distinguish two subcases.

In the first one the path γ contains no marked forbidden patterns and remains strictly under the x-axis. The path ϖ which kills ω is obtained by performing on μ the following: add the path $\overline{x}\gamma x\varrho_\ell\overline{x}\eta x$ by applying consecutive and appropriate mappings of the first production of (4.4), apply the actions giving the label (0) in case of no marked forbidden patterns. The path ν in ϖ is obtained as in ω.

In the latter subcase the path γ contains at least one marked forbidden pattern. We consider the rightmost point P of the path $\overline{x}\gamma\mathfrak{p}_\ell\eta x$ with lowest ordinate. The path ϖ which kills ω is obtained by performing on μ the following: add the path in $\gamma x\varrho_\ell\overline{x}\eta x$ running from P to the endpoint of the path by applying consecutive and appropriate mappings of the productions of (4.4), add the path in $\gamma x\varrho_\ell\overline{x}\eta x$ running from its initial point to P by applying consecutive and appropriate mappings of the production of (4.4), apply the cut and paste actions in case of marked forbidden patterns. Obviously, the last forbidden pattern in the path must be generated by applying consecutive and appropriate mappings of the first production of (4.4). The path ν in ϖ is obtained as in ω.

Second case: the last forbidden pattern $\mathfrak{p}_\ell$ in ω is not a marked forbidden pattern. We consider the subcases: $y_{L_\ell} \geq 0$ and $y_{L_\ell} < 0$.

$y_{L_\ell} \geq 0$: The path ω in $F \backslash F^{[\mathcal{P}_{j,i}]}$ can be written as $\omega = \mu \mathfrak{p}_\ell \eta \nu$, where $\mu \in F$ and $\eta, \nu \in F^{[\mathcal{P}_{j,i}]}$ (see Figure 4.24).

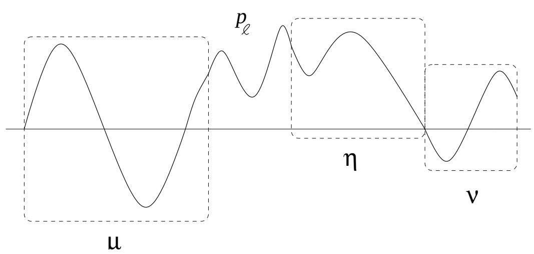


Figure 4.24: A representation of the path ω in the case $y_{L_\ell} \geq 0$

If μ is a Δ-path then the path $\overline{\omega}$ which kills the path ω is obtained by adding on μ the path $\mathfrak{p}_\ell \nu$ by applying an appropriate mapping of $(k) \overset{j_\ell}{\rightsquigarrow} (\overline{k + j_\ell - i_\ell}) \ldots (\overline{s_\ell}) \ldots (\overline{1})(\overline{0})$ of (4.4), and consecutive and appropriate mappings of the first production of (4.4).

Otherwise, if μ is a Γ-path then the path $\overline{\omega}$ which kills ω is obtained by adding on μ the path $\mathfrak{p}_\ell \nu$ by applying an appropriate mapping of $(k) \overset{j_\ell}{\rightsquigarrow} (\overline{k + j_\ell - i_\ell}) \ldots (\overline{s_\ell}) \ldots (\overline{1})(\overline{0})$ of (4.5), and consecutive and appropriate mappings of the first production of (4.5). In each case the path ν in $\overline{\omega}$ is obtained as in ω.

$y_{L_\ell} < 0$: In this case we distinguish the following two subcases: $y_{R_\ell} > 0$ and $y_{R_\ell} \leq 0$.

$y_{R_\ell} > 0$: The path ω in $F \backslash F^{[\mathcal{P}_{j,i}]}$ can be written as $\omega = \mu \overline{x} \gamma \mathfrak{p}_\ell \overline{x}^f \nu$, where $\mu, \gamma \in F$, $\nu \in F^{[\mathcal{P}_{j,i}]}$ and $0 \leq f \leq y_{R_\ell}$ (see Figure 4.25).

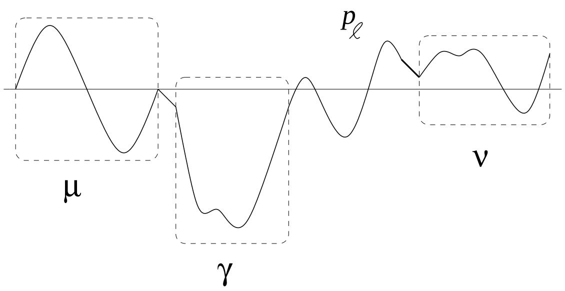


Figure 4.25: A representation of the path ω in the case $y_{L_\ell} < 0$ with $y_{R_\ell} > 0$

Let P be the rightmost point of the path $\overline{x} \gamma \mathfrak{p}_\ell$ with lowest ordinate. The path $\overline{\omega}$ which kills ω is obtained by performing on μ the following: add the path in $\gamma\mathfrak{p}_\ell$ running from P to the endpoint of the path by applying consecutive and appropriate mappings of (4.4), add the path in $\gamma\mathfrak{p}_\ell$ running from its initial point to P by applying consecutive and appropriate mappings of (4.4), apply the cut and paste actions in case of marked points and add the path $\overline{x}^f$. Obviously the last forbidden pattern in the path must be generated by applying the mapping of production $(k) \stackrel{j_\ell}{\rightsquigarrow} (\overline{k+j_\ell-i_\ell})\dots(\overline{s_\ell})\dots(\overline{1})(\overline{0})$ of (4.4). The path ν in $\overline{\omega}$ is obtained as in ω. Note that, in case $0 \leq f < y_{R_\ell}$ any prefix of ν in $\overline{\omega}$ which running from the end of the path $\mu\overline{x}\gamma\mathfrak{p}_\ell\overline{x}^f$ to the x-axis is obtained by applying the mapping associated to the first production of (4.5).

$y_{R_\ell} \leq 0$: The path ω in $F\backslash F^{[\mathcal{P}_{j,i}]}$ can be written as $\omega = \mu\overline{x}\gamma\mathfrak{p}_\ell\,\eta x\nu$, where $\mu, \gamma \in F$ and $\eta, \nu \in F^{[\mathcal{P}_{j,i}]}$ (see Figure 4.26).

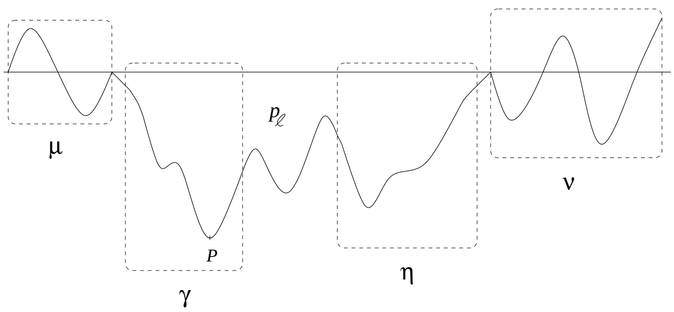


Figure 4.26: A representation of the path ω in the case $y_{L_\ell} < 0$ with $y_{R_\ell} \leq 0$

We consider the rightmost point P of the path $\overline{x}\gamma\mathfrak{p}_\ell\,\eta x$ with lowest ordinate. The path $\overline{\omega}$ which kills ω is obtained by performing on μ the following actions: add the path in $\gamma\mathfrak{p}_\ell\eta x$ running from P to its endpoint by applying consecutive and appropriate mappings of the productions of (4.4), add the path in $\gamma\mathfrak{p}_\ell\eta x$ running from its initial point to P by applying consecutive and appropriate mappings of the productions of (4.4), apply the cut and paste actions giving the label (0) in case of marked forbidden patterns. Obviously, the last marked forbidden pattern $\mathfrak{p}_\ell$ in $\overline{\omega}$ is generated by an appropriate mapping of the production $(k) \stackrel{j_\ell}{\rightsquigarrow} (\overline{k+j_\ell-i_\ell})\dots(\overline{s_\ell})\dots(\overline{1})(\overline{0})$. The path ν in $\overline{\omega}$ is obtained as in ω. Note that, if $y_{R_\ell} = 0$ then the path ω in $F\backslash F^{[\mathcal{P}_{j,i}]}$ can be written as $\omega = \mu\overline{x}\gamma\mathfrak{p}_\ell\,\nu$ and operations above described are applied on the path $\gamma\mathfrak{p}_\ell$.

We observe that for each path ω in $F\backslash F^{[\mathcal{P}_{j,i}]}$ 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 $F\backslash F^{[\mathcal{P}_{j,i}]}$ having n rise steps, containing C forbidden patterns and having last label $(\bar{k})$ of the associated code. The paths ω and ϖ have the same shape, exactly the same number and positions of the forbidden patterns except for the last one which is marked in ϖ 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 ϖ which kills a given path ω by applying two distinct procedures. ■

5
Positivity problem

This chapter shows as the succession rules can be applied in a different field from pattern avoidance. Indeed, we present a method to pass from a recurrence relation having constant coefficients (in short, a C-recurrence) to a finite succession rule defining the same number sequence. We also discuss the applicability of our method as a test for the positivity of a number sequence.

5.1 From C-sequences to succession rules

The main purpose of our research is to develop a general formal method to translate a given recurrence relation into a succession rule defining the same number sequence. By abuse of notation, in this case we will say that the recurrence relation and the succession rule are *equivalent*.

As a first step we deal with linear recurrence relations with integer coefficients [19, 27]. Following Zeilberger [83], we will address to these as *C-finite recurrence relations*, and to the defined sequences as *C-finite sequences*.

This section is organized as follows.

i) First we deal with C-sequences of the form:

$$f_n = a_1 f_{n-1} + a_2 f_{n-2} + \cdots + a_k f_{n-k} \qquad a_i \in \mathbb{Z}, 1 \leq i \leq k \tag{5.1}$$

with *default* initial conditions, i.e. $f_0 = 1$ and $f_h = 0$ for all $h < 0$.

We translate the given C-finite recurrence relation into an extended succession rule, possibly using both jumps and marked labels (Section 5.1.1).

ii) Then, we recursively eliminate jumps and marked labels from such an extended succession rule, thus obtaining a finite succession rule equivalent to the previous one (Section 5.1.2).

We remark that steps i) and ii) can be applied independently of the positivity of $\{f_n\}_{n\geq 0}$, but at this step we cannot be sure that all the labels of the obtained rule are nonnegative integers.

iii) We state a condition to ensure that the labels of the obtained succession rule are all nonnegative. If such a condition holds, then the sequence $\{f_n\}_{n\geq 0}$ has all positive terms, thus we refer to this as *positivity condition* (Section 5.1.3).

iv) We show how our method can be extended to C-sequences with generic initial conditions (Section 5.1.4).

5.1.1 C-sequences with default initial conditions

Let us consider a C-finite recurrence relation expressed as in (5.1), with default initial conditions and the related C-sequence $\{f_n\}_{n\geq 0}$. We recall that the generating function of $\{f_n\}_{n\geq 0}$ is rational, and precisely it is

$$f(x) = \sum_{n\geq 0} f_n x^n = \frac{1}{1 - a_1 x - a_2 x^2 - \cdots - a_k x^k}. \tag{5.2}$$

The first step of our method consists into translating the C-finite recurrence relation (5.1) into an extended succession rule. The translation is rather straightforward, since in practice it is just an equivalent way to represent the recurrence relation.

Proposition 7 *The recurrence relation (5.1) with default initial conditions is equivalent to the following extended succession rule:*

$$\begin{cases} (a_1) & \\ (a_1) & \overset{1}{\rightsquigarrow} (a_1)^{a_1} \\ (a_1) & \overset{2}{\rightsquigarrow} (a_1)^{a_2} \\ \vdots & \\ (a_1) & \overset{k}{\rightsquigarrow} (a_1)^{a_k} \end{cases} \tag{5.3}$$

where $(a_1)^{-a_i} = (\overline{a_1})^{a_i}$ *if* $a_i > 0$, $1 \leq i \leq k$.

For example, the recurrence relation $f_n = 3f_{n-1} + 2f_{n-2} - f_{n-3}$ with default initial conditions, defines the sequence $1, 3, 11, 38, 133, 464, 1620, 5655, \ldots,$ and it is equivalent to the following extended succession rule:

$$\begin{cases} (3) & \\ (3) & \overset{1}{\rightsquigarrow} (3)^3 \\ (3) & \overset{2}{\rightsquigarrow} (3)^2 \\ (3) & \overset{3}{\rightsquigarrow} (\overline{3}) \end{cases} \tag{5.4}$$

Figure 5.1 shows the first few levels of the generating tree associated to the jumping and marked succession rule (5.4).

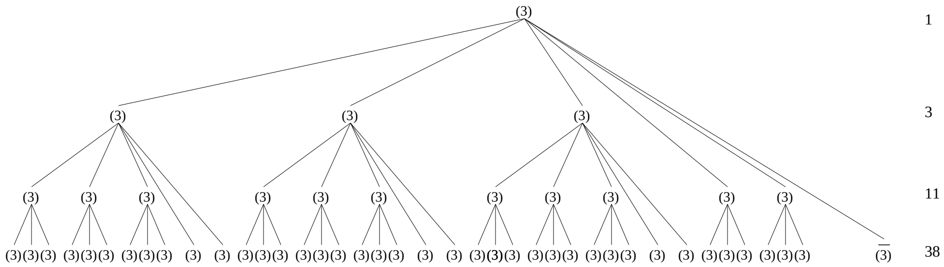


Figure 5.1: Some levels of the generating tree associated to the extended succession rule (5.4)

5.1.2 Elimination of jumps and marked labels

The successive step of our method consists into recursively eliminating jumps from the extended succession rule (5.3) in order to obtain a finite succession rule which is equivalent to the previous one. Once jumps have been eliminated we will deal with marked labels.

Proposition 8 *The succession rule:*

$$\left\{ \begin{array}{cl} (a_1) & \\ (a_1) & \rightsquigarrow (a_1+a_2)(a_1)^{a_1-1} \\ (a_1+a_2) & \rightsquigarrow (a_1+a_2+a_3)(a_1)^{a_1+a_2-1} \\ \vdots & \\ (\sum_{l=1}^{k-1} a_l) & \rightsquigarrow (\sum_{l=1}^{k} a_l)(a_1)^{(\sum_{l=1}^{k-1} a_l)-1} \\ (\sum_{l=1}^{k} a_l) & \rightsquigarrow (\sum_{l=1}^{k} a_l)(a_1)^{(\sum_{l=1}^{k} a_l)-1} \end{array} \right. \tag{5.5}$$

is equivalent to the recurrence relation $f_n = a_1 f_{n-1} + a_2 f_{n-2} + \cdots + a_k f_{n-k}$*,* $a_i \in \mathbb{Z}, 1 \leq i \leq k$*, with default initial conditions.*

Proof. Let $A_k(x)$ be the generating function of the label $(\sum_{l=1}^{k} a_l)$ related to the succession rule (5.5). We have:

$$A_1(x) = 1 + (a_1-1)xA_1(x) + (a_1+a_2-1)xA_2(x) + \ldots$$

$$\cdots + (a_1+a_2+\cdots+a_k-1)xA_k(x);$$

$$A_2(x) = xA_1(x);$$

$$A_3(x) = xA_2(x) = x^2A_1(x);$$

$$\vdots$$

$$A_{k-1}(x) = xA_{k-2}(x) = x^{k-2}A_1(x);$$

$$A_k(x) = xA_{k-1}(x) + xA_k(x) = \tfrac{x^{k-1}}{1-x}A_1(x).$$

Therefore,

$$A_1(x) = 1 + x(a_1 - 1)A_1(x) + x^2(a_1 + a_2 - 1)A_1(x) + \dots$$

$$\dots + \tfrac{x^k}{1-x}(a_1 + a_2 + \dots + a_k - 1)A_1(x),$$

and we obtain the generating function:

$$A_1(x) = \frac{1-x}{1 - a_1x - a_2x^2 - \dots - a_kx^k} \ .$$

At this point we can consider the generating function determined by the succession rule (5.5) as following:

$$A_1(x) + A_2(x) + \dots + A_{k-1}(x) + A_k(x) =$$

$$= A_1(x) + xA_1(x) + \dots + x^{k-2}A_1(x) + \frac{x^{k-1}}{1-x}A_1(x) =$$

$$= \frac{(1-x) + x(1-x) + \dots + x^{k-2}(1-x) + x^{k-1}}{1 - a_1x - a_2x^2 - \dots - a_kx^k} =$$

$$= \frac{1}{1 - a_1x - a_2x^2 - \dots - a_kx^k} \ .$$

■

Following the previous statement, the extended succession rule (5.4) – determined in the previous section – can be translated into the succession rule (5.6) and Figure 5.2 shows some levels of the associated generating tree.

$$\left\{ \begin{array}{ll} (3) & \\ (3) & \rightsquigarrow (5)(3)^2 \\ (5) & \rightsquigarrow (4)(3)^4 \\ (4) & \rightsquigarrow (4)(3)^3 \end{array} \right. \tag{5.6}$$

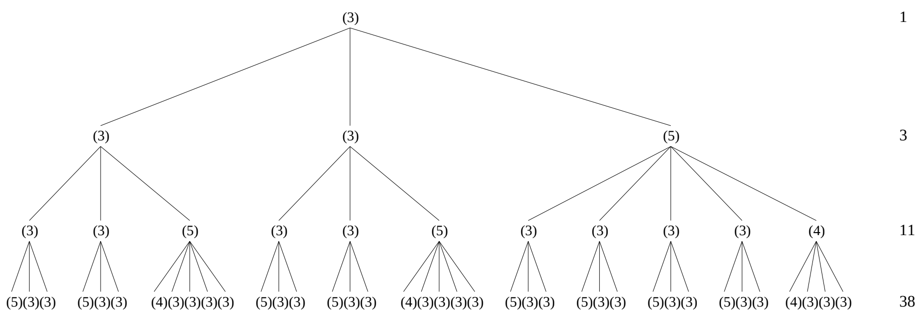


Figure 5.2: Four levels of the generating tree associated to the succession rule (5.6)

We observe that the previously obtained succession rule is an ordinary finite succession rule, but it may happen that the value of the label $(\sum_{l=1}^{i} a_l)$ is negative, for some i with $i \leq k$, then the succession rule (5.5) contains marked labels.

For example, the recurrence relation $f_n = 5f_{n-1} - 6f_{n-2} + 2f_{n-3}$, with default initial conditions, which defines the sequence $1, 5, 19, 67, 231, 791, 2703, 9231, \ldots$, (sequence A035344 in the The On-Line Encyclopedia of Integer Sequences [71]) is equivalent to the following succession rule:

$$\left\{ \begin{array}{ll} (5) & \\ (5) & \rightsquigarrow (-1)(5)^4 \\ (-1) & \rightsquigarrow (1)(\overline{5})^2 \\ (1) & \rightsquigarrow (1) \end{array} \right. \tag{5.7}$$

Figure 5.3 shows some levels of the generating tree associated to the extended succession rule (5.7), which is represented using a "compact notation", i.e., by convention, the number of nodes at a given level n is obtained by means of the algebraic sum of the exponents of the labels lying at level n.

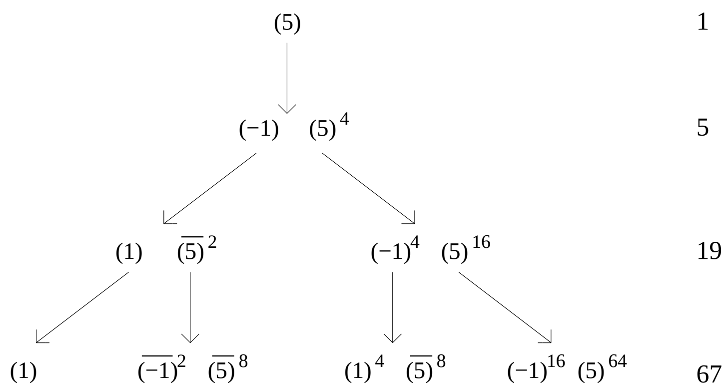


Figure 5.3: Compact notation for the generating tree associated to the succession rule (5.7)

Therefore our next goal is to remove all possible marked labels from the succession rule. We observe that in order to obtain this goal, the recurrence relation $f_n = a_1 f_{n-1} + a_2 f_{n-2} + \cdots + a_k f_{n-k}$ with default initial conditions needs $a_1 > 0$. We assume that this condition holds throughout the rest of the present section.

In order to furnish a clearer description of our method, we start considering the case $k = 2$.

Proposition 9 *The C-finite recurrence $f_n = a_1 f_{n-1} + a_2 f_{n-2}$, with default initial conditions, and having $a_1 > 0$, is equivalent to*

$$
\left\{
\begin{array}{ll}
(a_1) & \\
(a_1) & \rightsquigarrow (0)^{q_2}(r_2)(a_1)^{a_1-(q_2+1)} \\
(r_2) & \rightsquigarrow \Big((0)^{q_2}(r_2)\Big)^{q_2}(0)^{q_2}(r_2)(a_1)^{r_2-(q_2+1)^2}
\end{array}
\right.
\tag{5.8}
$$

where, by convention, the label (0) does not produce any son, and q_2, r_2 are defined as follows:

- *if $a_1 + a_2 \leq 0$ then $q_2, r_2 > 0$ such that $|a_1 + a_2| = q_2 a_1 - r_2$;*

- *otherwise $q_2 = 0$ and $r_2 = a_1 + a_2$.*

Proof. We have to distinguish two cases: in the first one $a_1 + a_2 \leq 0$ and in the second one $a_1 + a_2 > 0$.

If $a_1 + a_2 \leq 0$, we have to prove that the generating tree associated to the succession rule (5.8) is obtained by performing some actions on the generating tree associated to the extended succession rule (5.9) which is obviously equivalent to the recurrence $f_n = a_1 f_{n-1} + a_2 f_{n-2}$ having $a_1 > 0$ and $a_2 < 0$, with $f_0 = 1$ and $f_h = 0$ for each $h < 0$.

$$
\left\{
\begin{array}{ll}
(a_1) & \\
(a_1) & \stackrel{1}{\rightsquigarrow} (a_1)^{a_1} \\
(a_1) & \stackrel{2}{\rightsquigarrow} (a_1)^{a_2}
\end{array}
\right.
\tag{5.9}
$$

The proof consists in eliminating jumps and marked labels at each level of the generating tree associated to succession rule (5.9), sketched in Figure 5.4, by modifying the structure of the generating tree, still maintaining f_n nodes at level n, for each n.

Let (a_1) be a label at a given level n. We denote by B_1 the set of a_1 labels (a_1) at level $n+1$ and by B_2 the set of a_2 labels (a_1) at level $n+2$, see Figure 5.4. We remark that $(a_1)^{a_2} = \underbrace{\overline{(a_1)} \ldots \overline{(a_1)}}_{-a_2}$.

In order to eliminate both jumps and marked labels in B_2 at level 2 produced by the root (a_1) at level 0, we have to consider the set of a_1 labels (a_1) in B_1 at level 2 obtained by (a_1) which lie at level 1.

At level 2, each label (a_1) in a given set B_1 kills one and only one marked label $\overline{(a_1)}$ in B_2. At this point $|a_1+a_2|$ labels $\overline{(a_1)}$ in B_2 always exist at level 2.

In order to eliminate such marked labels we have to consider more than a single set B_1 of label (a_1) at level 2. Let q_2 be a sufficient number of sets B_1 at level 2 able to kill all the labels $\overline{(a_1)}$ in B_2 at level 2. Therefore $|a_1 + a_2| = q_2 a_1 - r_2$ with $q_2, r_2 > 0$.

By setting q_2 labels (a_1) at level 1 equal to (0) and one more label (a_1) to (r_2), we have the desired number of labels (a_1) at level 2.

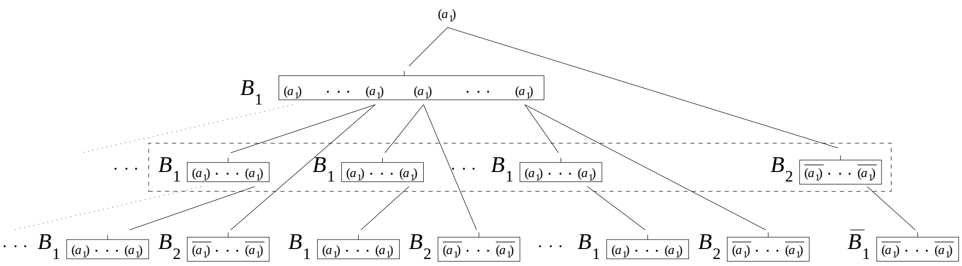


Figure 5.4: Step 1

Note that the marked labels at level 2 are not generated and the labels (a_1) at level 1 are revised in order to have the right number of labels at level 2, see Figure 5.5.

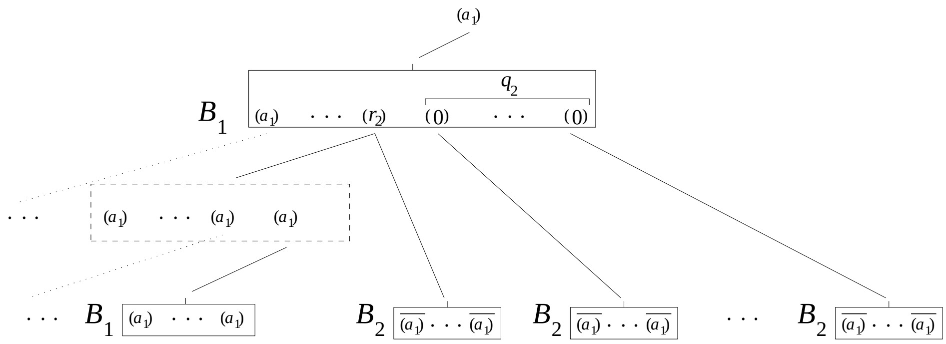


Figure 5.5: Step 2

Moreover, when a label (a_1) kills a marked label $\overline{(a_1)}$ at a given level n, then the subtree, having such label (a_1) as its root, kills the subtree having $\overline{(a_1)}$ as its root. So, at level 2 when a label (a_1) of B_1 kills a label $\overline{(a_1)}$ of B_2 then the two subtrees having such labels as their roots are eliminated too, see Figure 5.5.

On the other hand, the $q_2 + 1$ sets B_2 at level 3 obtained by the $q_2 + 1$ labels at level 1, once labelled with (a_1) and now having value $r_2, 0, \dots, 0$, respectively, are always present in the tree, see Figure 5.5.

In order to eliminate such undesired marked labels we can only set the production of (r_2). As a set B_2 at a given level is eliminated by using $q_2 + 1$ labels at previous level then (r_2) must give $(r_2)\underbrace{(0)\dots(0)}_{q_2}$ exactly $q_2 + 1$ times. This explains the first part of the production rule of the label (r_2) in rule (5.8).

Since (r_2) has r_2 sons then the remaining $r_2-(q_2+1)^2$ labels are set to be equal to (a_1) as in the previous case, see Figure 5.6.

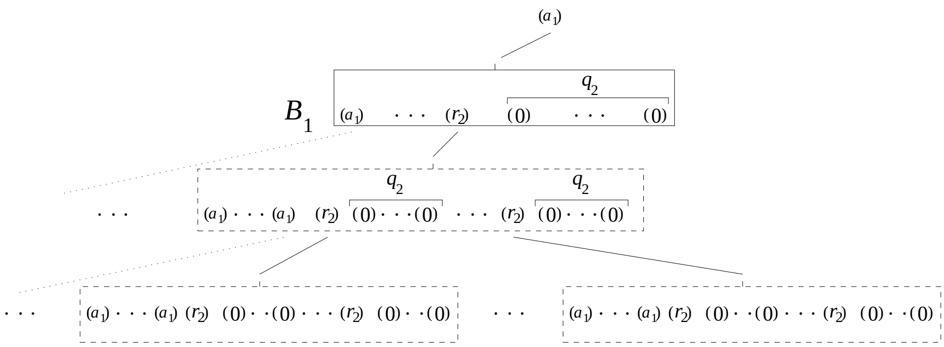


Figure 5.6: Step 3

By the way, the modified q_2+1 labels having value $r_2, 0, \ldots, 0$, respectively, at a given level n, produce the labels $\Big((0)^{q_2}(r_2)\Big)^{q_2+1}(a_1)^{r_2-(q_2+1)^2}$ at level $n+1$. Just as obtained for levels 1 and 2, the labels $\Big((0)^{q_2}(r_2)\Big)^{q_2+1}$ automatically annihilate the remaining q_2+1 sets B_2 of marked labels at level $n+2$, once obtained by the modified q_2+1 labels at level n, see Figure 5.6.

Till now we have modified a portion T of the total generating tree in a way that it does not contain any marked label. Note that, the remaining labels (a_1) will be the roots of subtrees which are all isomorphic to T.

The value f_n defined by the tree associated to the extended succession rule (5.9), is given by the difference between the number of non-marked and marked labels.

The just described algorithm modifies the number of generated non-marked labels and sets to 0 the number of marked ones in a way that f_n is unchanged, for each n, so the succession rule (5.8) is equivalent to the recurrence $f_n = a_1 f_{n-1} + a_2 f_{n-2}$.

In the case $a_1+a_2>0$ we have marked labels only if $a_2<0$. In this case a single set B_1 is sufficient to kill all the marked labels in B_2 at level 2.

By the way, both in the case $a_2<0$ and $a_2>0$ we have that $q_2=0$ and $r_2=a_1+a_2$, and the succession rule (5.8) has the same form of the rule (5.5) which is equivalent to the recurrence $f_n = a_1 f_{n-1} + a_2 f_{n-2}$ having $a_1>0$ and $a_2\in\mathbb{Z}$, with $f_0=1$ and $f_h=0$ for each $h<0$. ■

The statement of Proposition 9 can be naturally extended to the general case $k>2$.

Proposition 10 *The C-sequence $\{f_n\}_n$ satisfying $f_n = a_1 f_{n-1} + a_2 f_{n-2} + \cdots + a_k f_{n-k}$, with default initial conditions and $a_1 > 0$ is equivalent to*

$$(5.10)\qquad \begin{cases} (a_1) & \\ (a_1) & \rightsquigarrow (0)^{q_2}(r_2)(a_1)^{a_1-(q_2+1)} \\ (r_2) & \rightsquigarrow \Big((0)^{q_2}(r_2)\Big)^{q_2}(0)^{q_3}(r_3)(a_1)^{r_2-(q_2(q_2+1)+q_3+1)} \\ \vdots & \\ (r_i) & \rightsquigarrow \Big((0)^{q_2}(r_2)\Big)^{q_i}(0)^{q_{i+1}}(r_{i+1})(a_1)^{r_i-(q_i(q_2+1)+q_{i+1}+1)} \\ \vdots & \\ (r_k) & \rightsquigarrow \Big((0)^{q_2}(r_2)\Big)^{q_k}(0)^{q_k}(r_k)(a_1)^{r_k-(q_k(q_2+1)+q_k+1)} \end{cases}$$

where the parameters q_i and r_i, with $2 \leq i \leq k$, can be determined in the following way:

- if $\sum_{l=1}^{i} a_l \leq 0$ then $q_i, r_i > 0$ such that $|\sum_{l=1}^{i} a_l| = q_i a_1 - r_i$,

- otherwise $q_i = 0$ and $r_i = \sum_{l=1}^{i} a_l$.

The proof of the Proposition 10 is quite similar to the proof of Proposition 9. It has the same level of difficulty but it is more cumbersome, so it is omitted for brevity.

Using Proposition 10, we can translate the previously considered recurrence relation $f_n = 5f_{n-1} - 6f_{n-2} + 2f_{n-3}$, with default initial conditions, into the following ordinary succession rule:

$$(5.11)\qquad \begin{cases} (5) & \\ (5) & \rightsquigarrow (0)(4)(5)^3 \\ (4) & \rightsquigarrow (0)(4)(1)(5) \\ (1) & \rightsquigarrow (1) \end{cases}$$

being $q_2 = 1$, $r_2 = 4$, $q_3 = 0$ and $r_3 = 1$.

5.1.3 Positivity condition

The statement of Proposition 10 is indeed a tool to translate C-recurrences into finite succession rules. However this property turns out to be effectively applicable only when the labels of the succession rule are all positive, and the reader can easily observe that Proposition 10 does not give us an instrument to test whether this happens or not.

In particular, if the labels of the succession rule are all positive then the terms of the C-sequence are all positive. It is then interesting to relate our problem with the so called *positivity problem*, which we have already mentioned in the Introduction.

Positivity Problem: given a C-finite sequence $\{f_n\}_{n\geq 0}$, establish if all its terms are positive.

We recall that the problem was originally proposed as an open problem in [17], and then re-presented in [66] (Theorems 12.1-12.2, pages 73-74), but no general solution has been found yet.

Moreover, the positivity problem can be solved for a large class of C-finite sequences, precisely for $\mathbb{N}$-rational sequences. We recall that the class of $\mathbb{N}$-rational series is precisely the class of the generating functions of regular languages, and that Soittola's Theorem [72] states that the problem of establishing whether a rational generating function is $\mathbb{N}$-rational is decidable.

Let us start examining the case of C-recurrences of degree 2. So, let $f_n = a_1 f_{n-1} + a_2 f_{n-2}$ be a recurrence relation, with $a_1 > 0$ and $a_2 \in \mathbb{Z}$.

By referring to the succession rule (5.8), precisely to the case $a_1 + a_2 \leq 0$, we observe that the succession rule equivalent to the recurrence relation is an ordinary rule (i.e., it has all positive labels) if and only if the following condition is verified:

$$\left\{\begin{array}{l} a_1 - (q_2+1) \geq 0 \\ r_2 - (q_2+1)^2 \geq 0 \end{array}\right. \tag{5.12}$$

As $r_2 = q_2 a_1 - |a_1 + a_2| = q_2 a_1 + a_1 + a_2$ then $r_2 - (q_2+1)^2 \geq 0$ means ${q_2}^2 + (2 - a_1) q_2 + 1 - a_1 - a_2 \leq 0$. This inequality has solution if and only if ${a_1}^2 + 4a_2 \geq 0$, and this is clearly a necessary and sufficient condition to ensure the positivity of all the terms of f_n [11].

Let us now consider a generic C-recurrence of degree k. Using a similar reasoning, and following Proposition 10 we can prove:

Corollary 1 *Let us consider the recurrence relation* $f_n = a_1 f_{n-1} + a_2 f_{n-2} + \cdots + a_k f_{n-k}$ *having* $a_1 > 0$ *and* $a_i \in \mathbb{Z}$, $2 \leq i \leq k$, *with* $f_0 = 1$ *and* $f_h = 0$ *for each* $h < 0$. *If*

$$\left\{\begin{array}{l} a_1 - (q_2+1) \geq 0 \\ r_2 - (q_2(q_2+1) + q_3 + 1) \geq 0 \\ \vdots \\ r_i - (q_i(q_2+1) + q_{i+1} + 1) \geq 0 \quad , \quad 3 \leq i \leq k-1 \\ \vdots \\ r_k - (q_k(q_2+1) + q_k + 1) \geq 0 \end{array}\right. \tag{5.13}$$

then $f_n > 0$ *for all* n.

As $r_i = \sum_{l=1}^{i} a_l + q_i a_1$, $2 \leq i \leq k$, then the system (5.13) can be rewritten in the following form:

$$(5.14)\quad \begin{cases} a_1 - (q_2+1) \geq 0 \\ \vdots \\ \sum_{l=1}^{i} a_l + q_i a_1 - (q_i(q_2+1) + q_{i+1} + 1) \geq 0 \quad , \quad 2 \leq i \leq k-1 \\ \vdots \\ \sum_{l=1}^{k} a_l + q_k a_1 - (q_k(q_2+1) + q_k + 1) \geq 0. \end{cases}$$

As previously mentioned, condition (5.14) ensures that all the labels of the succession rules equivalent to the given C-recurrence relation are positive, hence all the terms f_n are positive. Thus it can be viewed as a sufficient condition to test the positivity of a given C-recurrence relation.

Unfortunately, this is not a necessary condition to test positivity, then there are cases of positive C-sequences for which our method fails to prove positivity. A simple example is given by any positive non $\mathbb{N}$-rational C-sequence. The reader can find an instance of such sequences in [47].
It would be more interesting to give an example of a $\mathbb{N}$-rational C-sequence for which our method is not able to prove positivity, but we have not been able to find any such example.

Clearly, any C-sequence satisfying the positivity condition has a $\mathbb{N}$-rational generating function (in fact, any finite succession rule may be regarded as a finite state automaton), thus our method can be suitably used to test the $\mathbb{N}$-rationality of a sequence.
Though it is not our intention to deepen the computational complexity of our test, we remark that, despite the methods presented in [19, 47, 62], our method does not deal with calculating polynomial roots.

In order to give an idea of the computational cost to solve the system (5.14) we consider the worst case that is when $\sum_{l=1}^{i} a_l \leq 0$, $2 \leq i \leq k$, and the system itself has no solution.

In this case all the possible values for each q_i, $2 \leq i \leq k$, must be checked in order to conclude that the system (5.14) does not admit any solution.

As q_2 can range in the close set $[1, a_1 - 1]$ and q_{i+1} in $[1, \sum_{l=1}^{i} a_l + q_i a_1 - (q_i(q_2+1) - 1]$ then we have

$$1 + (a_1 - 1) \prod_{i=2}^{k-1} \left(\sum_{l=1}^{i} a_l + a_1 q_i - q_i(q_2+1) - 1 \right)$$

where the first 1 accounts the check to verify $\sum_{l=1}^{k} a_l + q_k a_1 - (q_k(q_2+1) + q_k + 1) \geq 0$.

An average complexity study of our test is a further development. Anyway, the referred experimental results give a sufficiently short computational time to test condition (5.14).

5.1.4 Generic initial conditions

It is now possible to use the statement of Proposition 10 to treat the case of C-recurrence relations with generic initial conditions. The following result is obtained by simply adapting the productions of the labels in the first levels of the generating tree to the given initial conditions, then using the productions of Proposition 10. So, we have two sets of productions: the ones stating the initial conditions, and the remaining ones defining all the other levels.

Proposition 11 *Let us consider the C-finite recurrence relation* $f_n = a_1 f_{n-1} + a_2 f_{n-2} + \cdots + a_k f_{n-k}$, $a_i \in \mathbb{Z}$, $1 \leq i \leq k$, *and let us assume that the initial conditions are* $f_0 = 1$ *and* $f_i = h_i$, *with* $h_i \in \mathbb{Z}$, $1 \leq i < k$, *then it can be translated into the following extended succession rule:*

$$\left\{\begin{array}{lll} (h_1) & & \\ (h_1) & \overset{1}{\rightsquigarrow} & (a_1)^{h_1} \\ \vdots & & \\ (h_1) & \overset{i}{\rightsquigarrow} & (a_1)^{h_i - \sum_{j=1}^{i-1} h_j a_{i-j}} \quad , \quad 1 < i < k \\ \vdots & & \\ (h_1) & \overset{k}{\rightsquigarrow} & (a_1)^{a_k} \\ & & \\ (a_1) & \overset{1}{\rightsquigarrow} & (a_1)^{a_1} \\ (a_1) & \overset{2}{\rightsquigarrow} & (a_1)^{a_2} \\ \vdots & & \\ (a_1) & \overset{k}{\rightsquigarrow} & (a_1)^{a_k} \end{array}\right. \tag{5.15}$$

For example, the recurrence relation $f_n = 3f_{n-1} + 2f_{n-2} - f_{n-3}$ with $f_0 = 1$, $f_1 = 2$ and $f_2 = 3$, which defines the sequence $1, 2, 3, 12, 40, 141, 491, 1715, \ldots$, is equivalent to the following extended succession rule:

$$\left\{\begin{array}{lll} (2) & & \\ (2) & \overset{1}{\rightsquigarrow} & (3)^2 \\ (2) & \overset{2}{\rightsquigarrow} & (\overline{3})^3 \\ (2) & \overset{3}{\rightsquigarrow} & (\overline{3}) \\ & & \\ (3) & \overset{1}{\rightsquigarrow} & (3)^3 \\ (3) & \overset{2}{\rightsquigarrow} & (3)^2 \\ (3) & \overset{3}{\rightsquigarrow} & (\overline{3}) \end{array}\right. \tag{5.16}$$

Figure 5.7 shows some levels of the generating tree associated to the extended succession rule (5.16).

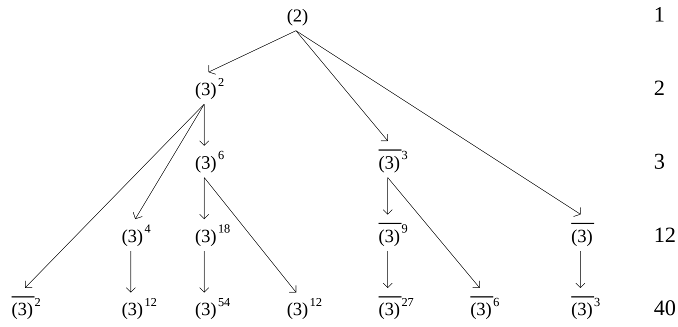


Figure 5.7: Compact notation for the generating tree associated to the succession rule (5.16)

Following the described method in Section 5.1.2 to eliminate jumps and marked labels, we can translate the extended succession rule (5.16) into the ordinary succession rule (5.17), where the labels (3), $(3)_1$ and $(3)_2$ are different labels with different productions.

$$
\left\{
\begin{array}{lll}
(2) & & \\
(2) & \rightsquigarrow & (0)(3)_1 \\
(3)_1 & \rightsquigarrow & (6)(3)^2 \\
(6) & \rightsquigarrow & (3)^5(3)_2 \\
(3)_2 & \rightsquigarrow & (4)(3)^2 \\
 & & \\
(3) & \rightsquigarrow & (5)(3)^2 \\
(5) & \rightsquigarrow & (4)(3)^4 \\
(4) & \rightsquigarrow & (4)(3)^3
\end{array}
\right. \tag{5.17}
$$

6 Conclusions

In Chapter 2 we propose an unified algorithmic approach to study binary words, having the number of ones greater than or equal to the number of zeroes, and avoiding a fixed pattern $\mathfrak{p}$. Initially, we consider $\mathfrak{p} = \mathfrak{p}(j) = 1^{j+1}0^j$, $j \geq 1$, then we pass to $\mathfrak{p}(j,j) = 1^j0^j$, $j \geq 1$, and finally the approach is generalized to $\mathfrak{p}(j,i) = 1^j0^i$, $0 < i < j$.

In Chapter 4, using the results of Chapter 3 we extend the results obtained in Chapter 2 passing from the avoidance of a single pattern to a set of patterns. The used technics always involve the generation of all words and then the delation of the undesired ones. The main result on this chapter is that the construction and the enumeration of the studied class of words does not depend on the shape of the avoided patterns themselves, but only on the number of ones in the patterns.

Further studies could take into consideration other parameters of these structures instead of the number of ones. These studies could be the first step to investigate a possible uniform constructive algorithm for pattern avoidance in words. In order to do that other forbidden patterns should be take into consideration like the ones having the number of 1's less then the number of 0's.

The words studied in this thesis have the restriction that the number of ones is greater than or equal to the number of zeroes, it could be interesting to investigate if the proposed algorithmic approach can be slightly modified in order to consider different restrictions on the words. These studies could be carry out considering an alphabet having cardinality greater than 2.

In Chapter 3, we introduce a general constructing method for the sets of cross-bifix-free binary words of fixed length n based upon the study of lattice paths on the Cartesian plane. This approach enables us to obtain the cross-bifix-free set having greater cardinality than the ones proposed in [4]. Moreover, we prove that this set is a non-expandable cross-bifix-free set. The non-expandable property is obviously a necessary condition to obtain a maximal cross-bifix-free set, anyway we are not able to find and prove a sufficient condition.

Further studies could prove that this set is maximal or not by finding a new set having greater cardinality of this one. Anyway, it should be challenging to determine if the problem of proving that, a given non-expandable cross-bifix-free set is maximal, is decidable or not. In the first case an accurate study of the algorithmic complexity would be required. This question arise from the fact that the problem of determining the maximal cross-bifix-free set is strictly lied to the problem of finding the maximum size of clique in a given graph.

Successive studies should take into consideration the general study of cross-bifix-free sets on an alphabet having cardinality grater than 2.

Another interesting research line is to investigate the relations between the distributed sequences in [50] and the proposed set.

In Chapter 5 we present a general method to translate a given C-finite recurrence relation into an ordinary succession rule and we have proposed a sufficient condition for testing the positivity of a given C-finite sequence. A further development could take into consideration the average complexity necessary to prove the positivity of a given C-finite sequence. Afterwards, it should be interesting to develop the study concerning the C-recurrence relations with generic initial conditions in order to examine in depth the potentiality of our method.

Moreover, we would like to show that some of our ideas can be applied to the case of holonomic integer sequences, i.e., those satisfying a linear recurrence relation with polynomial coefficients. Just to have a simple example, let us consider the *involutions* of n, enumerated by the sequence $\{f_n\}$ defined by the holonomic recurrence relation

$$f_n = f_{n-1} + (n-1)f_{n-2}, \tag{6.1}$$

with $f_0 = 1$, $f_1 = 1$ (sequence A000085 in the The On-Line Encyclopedia of Integer Sequences [71]).

We easily observe that, using the same argument of Proposition 7, we can translate the recurrence relation (6.1) into an infinite succession rule (possibly having marked labels and jumps), where now we adopt the convention that a generic label (k) is placed at the level k of the generating tree:

$$\left\{\begin{array}{lll} (0) & & \\ (k) & \stackrel{1}{\rightsquigarrow} & (k+1) \\ (k) & \stackrel{2}{\rightsquigarrow} & (k+2)^{k+1} \end{array}\right. \tag{6.2}$$

The successive step is to find a way how to convert such a rule into an ordinary succession rule. Referring to (6.2), this can be done by eliminating "by hand" marked labels and jumps, then re-writing the (ordinary) rule as follows:

$$\left\{\begin{array}{ll} (1) & \\ (k) & \rightsquigarrow (k-1)^{k-1}(k+1) \end{array}\right. \tag{6.3}$$

We believe that such a method should be formalized in some further works, and then applied to automatically convert the obtained rule into an ordinary succession rule. Moreover, from this method, we could also develop a more general criterion for proving the positivity of an holonomic sequence.

Finally, it could be interesting to define a language, possibly with forbidden patterns, starting from a generic recurrence relation.

Bibliography

[1] A. Apostolico, M. Atallah. Compact recognizers of episode sequences. *Information and Computation* 174 (2), 180–192, (2002). (Cited on page 2).

[2] S. Banchelli, E. Barcucci, E. Grazzini, E. Pergola. Exhaustive generation of combinatorial objects using ECO. *Acta Informatica* 40 (8), 585–602, (2004). (Cited on page 8).

[3] D. Baccherini, D. Merlini, R. Sprugnoli. Binary words excluding a pattern and proper Riordan arrays. *Discrete Mathematics* 307, 1021–1037, (2007). (Cited on pages 13, 19, and 21).

[4] D. Bajic. On Construction of Cross-Bifix-Free Kernel Sets. *2nd MCM COST 2100* TD (07), 237, Lisbon, Portugal, (2007). (Cited on pages 57, 58, and 115).

[5] D. Bajic, D. Drajic. Duration of search for a fixed pattern in random data: Distribution function and variance. *Electronics letters* 31 (8), 631–632, (1995). (Cited on page 57).

[6] D. Bajic, C. Stefanovic. On the search for a sequence from a predefined set of sequences in random and framed data streams. *IEEE Transactions on Commununications* 60 (1), 189–197, (2012). (Cited on page 58).

[7] D. Bajic, J. Stojanovic. Distributed Sequences and Search Process. *IEEE International Conference on Communications ICC2004*, 514–518, Paris, (2004). (Cited on pages 57 and 58).

[8] C. Banderier, M. Bousquet-Mélou, A. Denise, P. Flajolet, D. Gardy, D. Gouyou-Beauchamps. Generating functions for generating trees. *Discrete Mathematics* 246, 29–55, (2002). (Cited on pages 2, 33, 35, 38, 77, 79, and 84).

[9] E. Barcucci, A. Del Lungo, A. Frosini, S. Rinaldi. A technology for reverse-engineering a combinatorial problem from a rational generating function. *Advances in Applied Mathematics* 26 (2), 129–153, (2001). (Cited on page 3).

[10] E. Barcucci, A. Del Lungo, E. Pergola, R. Pinzani. Random generation of trees and other combinatorial objects. *Theoretical Computer Science* 218, 219–232, (1999). (Cited on pages 7 and 8).

[11] E. Barcucci, S. Rinaldi. Some linear recurrences and their combinatorial interpretation by means of regular languages. *Theoretical Computer Science* 255, 679–686, (2001). (Cited on page 110).

[12] E. Barcucci, A. Del Lungo, E. Pergola, R. Pinzani. ECO: a methodology for the Enumeration of Combinatorial Objects. *Journal of Difference Equations and Applications* 5, 435–490, (1999). (Cited on pages 2, 7, and 8).

[13] R. H. Barker. Group synchronizing of binary digital systems. *Communication theory*, W. jackson, Ed. London, U.K.: Butterworth, 273–287, (1953). (Cited on page 57).

[14] D. Beauquier, M. Nivat. On traslating one polyomino to tile the plane. *Discrete and Computational Geometry* 6, 575–592, (1991). (Cited on page 1).

[15] A. Bernini, E. Grazzini, E. Pergola, R. Pinzani. A general exhaustive generation algorithm for Gray structures, *Acta Informatica* 44 (5), 361–376, (2007). (Cited on page 8).

[16] J. Berstel. Axel Thue's papers on repetitions on words: traslation. *Publication du LaCIM* 19, (1994). (Cited on page 1).

[17] J. Berstel, M. Mignotte. Deux propriétés décidables des suites récurrentes linéaires. *Bulletin de la Socité Mathématique de France* 104 (2), 175–184, (1976). (Cited on pages 3 and 110).

[18] J. Berstel, D. Perrin, C. Reutenauer. Codes and Automata (Encyclopedia of Mathematics and its Applications). Cambridge University Press, Cambridge 2009. (Cited on page 58).

[19] J. Berstel, C. Reutenauer. Another proof of Soittola's Theorem. *Theoretical Computer Science* 393, 196–203, (2008). (Cited on pages 3, 101, and 111).

[20] P. Bose, J. F. Buss, A. Lubiw. Pattern matching for permutations. *Lecture Notes in Computer Science* 709, 200–209, (1993). (Cited on page 1).

[21] M. Bousquet-Melou. Percolation Models and Animals. *European Journal of Combinatorics* 17, 343–369, (1996). (Cited on page 1).

[22] F. R. K. Chung, R. L. Graham, V. E. Hoggatt, M. Kleimann. The number of Baxter permutations. *Journal of Combinatorial Theory Series A* 24, 382–394, (1978). (Cited on page 7).

[23] L. Comtet. Advanced Combinatorics: The Art of Finite and Infinite Expansions, D. Reidel Publishing Company, (1974). (Cited on page 60).

[24] J. H. Conway, J. C. Lagarias. Tiling with polyominoes and combinatorial group theory. *Journal of Combinatorial Theory Series A* 53, 183–208 (1990). (Cited on page 1).

[25] S. Corteel. Series generatrices exponentielles pour les eco-systemes signes. *Proceedings of the 12-th International Conference on Formal Power Series and Algebraic Combinatorics*, Moscow (2000). (Cited on pages 2 and 8).

[26] M. Crochemore, C. Hancart, T. Lecroq. Algorithms on strings. Cambridge University Press, Cambridge (2007). (Cited on page 58).

[27] E. Duchi, A. Frosini, R. Pinzani, S. Rinaldi. A note on rational succession rules. *Journal of Integer Sequences* 6, Article 03.1.7, (2003). (Cited on pages 3 and 101).

[28] S. Dulucq, O. Guibert. Mots de piles, tableaux stardards et permutations de Baxter. *Proceedings of the 6-th International Conference on Formal Power Series and Algebraic Combinatorics*, Dimacs, 119–128, (1994). (Cited on page 7).

[29] S. Dulucq, S. Gire, J. West. Permutations with forbidden subsequences and nonseparable planar maps. *Discrete Mathematics* 153, 85-105, (1996). (Cited on page 7).

[30] L. Ferrari, E. Pergola, R. Pinzani, S. Rinaldi. An algebraic characterization of the set of succession rules. *Theoretical Computer Science* 281, 351–367, (2002). (Cited on page 2).

[31] L. Ferrari, E. Pergola, R. Pinzani, S. Rinaldi. Jumping succession rules and their generating functions. *Discrete Mathematics* 271, 29–50, (2003). (Cited on pages 2 and 10).

[32] L. Ferrari, R. Pinzani. A linear operator approach to succession rules. *Linear Algebra and Its Applications* 348, 231–246, (2002). (Cited on pages 33, 38, 77, and 84).

[33] P. Flajolet, J. Francon, J. Vuillemin. Sequences of operations analysis for dynamic data structures. *Journal of Algorithms* 1, 111–141, (1980). (Cited on page 1).

[34] P. Flajolet, R. Sedgewick. Analytic Combinatorics. Cambridge University Press, Cambridge (2009). (Cited on page 1).

[35] P. Flajolet, W. Szpankowski, B. Valle. Hidden word statistics, *Journal of the ACM* 53 (1), 147–183, (2006). (Cited on page 2).

[36] A. M. Garsia. Combinatorics of the free Lie algebra and symmetric group. Analysis, et cetera, Academic Press, Boston, MA, 309-382, (1990). (Cited on page 1).

[37] I. Gessel. Symmetric functions and P-recursiveness. *Journal of Combinatorial Theory Series A* 53, 257–285, (1990). (Cited on page 1).

[38] I. Gessel. Rational functions with nonnegative integer coefficients. In *The 50th seminaire Lotharingien de Combinatoire*, page Domaine Saint-Jacques, March (2003). Unpublished, available at Gessel's homepage. (Cited on page 4).

[39] S. Gerhold. Sequences: Non-Holonomicity and Inequalities. *Ph.D. Thesis.* (Cited on page 4).

[40] L. J. Guibas, M. Odlyzko, Long repetitive patterns in random sequences. *Zeitschrift für Wahrscheinlichkeitstheorie* 53, 241–262, (1980). (Cited on pages 3 and 13).

[41] L. J. Guibas, M. Odlyzko. String overlaps, pattern matching, and non-transitive games. *Journal of Combinatorial Theory Series A* 30, 183–208, (1981). (Cited on pages 3 and 13).

[42] V. Halava, T. Harju, M. Hirvensalo. Positivity of second order linear recurrent sequences. *Discrete Applied Mathematics* 154, 447–451, (2006). (Cited on page 4).

[43] T. Harju, D. Nowotka. Counting bordered and primitive words with a fixed weight. *Theoretical Computer Science* 340, 273–279, (2005). (Cited on page 58).

[44] R. E. Horton. Erosioned development of systems and their drainege basins, hydrophysical approach to quantitative morphology. *Geological Society of America Bulletin* 56, 275–379, (1945). (Cited on page 1).

[45] K. H. Kim, M. S. Putcha, F. W. Roush. Some combinatorial properties of free semigroups. *Journal of the London Mathematical Society* 16 (2), 397–402, (1977). (Cited on page 3).

[46] D. E. Knuth. The art of computer programming, vol. 3, Addison-Wesley, Reading, MA (1973). (Cited on page 1).

[47] C. Koutschan. Regular languages and their generating functions: The inverse problem. *Theoretical Computer Science* 391, 65–74, (2008). (Cited on pages 3 and 111).

[48] S. Kumar, E. H. Spafford. A pattern matching model for misuse intrusion detection. In *Computer Security*, 11–21, (1994). (Cited on page 2).

[49] V. Laohakosol, P. Tangsupphathawat. Positivity of third order linear recurrence sequences. *Discrete Applied Mathematics* 157, 3239–3248, (2009). (Cited on page 4).

[50] A. J. de Lind van Wijngaarden, T. J. Willink. Frame Synchronization Using Distributed Sequences. *IEEE Transactions on Commununications* 48 (12), 2127–2138, (2000). (Cited on pages 57, 58, and 116).

[51] J. Lindner. Binary sequences up to length 40 with best possible autocorrelation function. *Electronics letters* 11, 507, 1975. (Cited on page 57).

[52] M. Lothaire. Combinatorics and words. G.C. Rota Editor, Encyclopedia of Mathemathics and its Applications 17, Addison-Wesley, Reading, MA (1983). (Cited on page 1).

[53] T. A. MacMahon. The mechanical design of trees. *Scientific American* 233, 92–102, (1975). (Cited on page 1).

[54] D. Merlini, R. Sprugnoli. Algebraic aspects of some Riordan arrays related to binary words avoiding a pattern. *Theoretical Computer Science* 412 (27), 2988–3001, (2011). (Cited on page 19).

[55] D. Merlini, R. Sprugnoli, M. C. Verri. An Algebra for proper generating tree. In *Algorithms, trees, combinatorics and probabilities.* Trends in Mathematics, Mathematics and Computer Science, 127–139, (2000). (Cited on pages 11 and 21).

[56] D. Merlini, D. G. Rogers, R. Sprugnoli, M. C. Verri. On some alternative characterizations of Riordan arrays. *Canadian Journal of Mathematics* 49 (2), 301–320, (1997). (Cited on pages 17, 18, and 19).

[57] D. Merlini, M. C. Verri. Generating trees and proper Riordan Arrays. *Discrete Mathematics* 218, 167–183, (2003). (Cited on pages 19 and 21).

[58] Mitiko Go. Statistical mechanics of biopolymers and its applications to the melting transition of polynucleotides. *Journal of the Physical Society of Japan* 23, (1967). (Cited on page 1).

[59] P. T. Nielsen. On the Expected Duration of a Search for a Fixed Pattern in Random Data. *IEEE Transactions on Information Theory* IT-29, 702–704, (1973). (Cited on page 57).

[60] P. T. Nielsen. A Note on Bifix-Free Sequences. *IEEE Transactions on Information Theory* IT-29, 704–706, (1973). (Cited on page 58).

[61] G. Percheron. Principles and methods of the graph theoretical analysis of natural binary arborescences. *Journal of Theoretical Biology* 99, 509–552, (1982). (Cited on page 1).

[62] D. Perrin. On positive matrices. *Theoretical Computer Science* 94 (2), 357–366, (1992). (Cited on pages 3 and 111).

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

[64] D. Rotem. Stack sortable permutations. *Discrete Mathematics* 33, 185–196, (1981). (Cited on page 1).

[65] B. E. Sagan. The Symmetric Group: Representations, Combinatorial Algorithms, and Symmetric Functions. Springer, 2nd Edition (2009). (Cited on page 1).

[66] A. Salomaa, M. Soittola. Automata-Theoretic Aspects of Formal Power Series, Springer-Verlag, New York (1978). (Cited on pages 3 and 110).

[67] R. A. Scholtz. Frame synchronization techniques. *IEEE Transactions on Commununications* COM-20, 1204–1213, (1980). (Cited on page 57).

[68] M. P. Schützenberger. On context-free languages and push-down automata. *Information and Control* 6, 246–264, (1963). (Cited on pages 1 and 16).

[69] R. Sedgewick, P. Flajolet. An Introduction to the Analysis of Algorithms. Chapman-Hall, London (1995). (Cited on pages 1, 3, and 13).

[70] D. Shen, W. Zhang, D. P. Reed, A. B. Lippman. On frame synchronization for multiple access channels. *ICC Workshops 2009. IEEE International Conference on ICC 2009*, 1–5, (2009). (Cited on page 58).

[71] N. J. A. Sloane. On-line encyclopedia of integer sequences. http://oeis.org/. (Cited on pages 8, 58, 105, and 116).

[72] M. Soittola. Positive rational sequences. *Theoretical Computer Science* 2 (3), 317–322, (1976). (Cited on pages 3 and 110).

[73] R. P. Stanley. Enumerative Combinatorics, Vol. 2. Cambridge University Press. Cambridge (1999). (Cited on pages 8, 20, and 60).

[74] R. P. Stanley. Catalan addendum. Available at http://www-math.mit.edu/∼rstan/ec/catadd.pdf. (Cited on page 8).

[75] A. N. Strahler. Hypsometric (area-altitude) analysis of erosional topology. *Geological Society of America Bulletin* 63, 1117–1142, (1952). (Cited on page 1).

[76] R. A. Sulanke. Moments of generalized Motzkin paths. *Journal of Integer Sequences* 3, Article 00.1.1, (2000). (Cited on page 10).

[77] M. Vanchaussade De Chaumont, X. G. Viennot. Enumeration of RNAs secondary structures by complexity. Mathematics in medicine and biology. *Lecture Notes in Biomaths* 57, Springer-Verlag, Berlin, 360–365, (1985). (Cited on page 1).

[78] M. Waterman. Secondary structure of singled stranded nucleid acids. Studies on foundations and combinatorics. *Advances in Mathematics. Supplementary Studies* 1, 167–212, (1978). (Cited on page 1).

[79] M. Waterman. Introduction to Computational Biology. Addison-Wesley, Reading, MA (1995). (Cited on page 2).

[80] J. West. Sorting twice through a stack. *Theoretical Computer Science* 117, 303–313, (1993). (Cited on page 1).

[81] J. West. Generating trees and the Catalan and Schröder numbers. *Discrete Mathematics* 146, 247–262, (1995). (Cited on page 7).

[82] M. J. Woldenberg. A structural taxonomy of spatial hierarchies. *Proceedings on Colston papers* 22, Butterwoth Scient. Publ., London, 147–175, (1970). (Cited on page 1).

[83] D. Zeilberger. A holonomic systems approach to special functions identities. *Journal of Computational and Applied Mathematics* 32, 321–368, (1990). (Cited on pages 3 and 101).