

1 Analytic combinatorics of lattice paths with forbidden patterns,  
2 the vectorial kernel method,  
3 and generating functions for pushdown automata

4 Andrei Asinowski<sup>12\*</sup> Axel Bacher<sup>3†</sup> Cyril Banderier<sup>3‡</sup> Bernhard Gittenberger<sup>1§</sup>

5 <sup>1</sup> Technische Universität Wien, Institut für Diskrete Mathematik und Geometrie, Austria

6 <sup>2</sup> Alpen-Adria Universität Klagenfurt, Institut für Mathematik, Austria

7 <sup>3</sup> Université Paris 13, LIPN, UMR CNRS 7030, France

8 Version of August 15, 2019 (to appear in *Algorithmica*)

9 \*\*\*

10 *We dedicate this article to the memory of Philippe Flajolet,*  
11 *our cheerful and inspiring mentor, founder of analytic combinatorics.*

12 \*\*\*

13 **Abstract**

14  
15 In this article we develop a *vectorial kernel method* – a powerful method which  
16 solves in a unified framework all the problems related to the enumeration of words  
17 generated by a pushdown automaton. We apply it for the enumeration of lattice paths  
18 that avoid a fixed word (a *pattern*), or for counting the occurrences of a given pattern.  
19 We unify results from numerous articles concerning patterns like peaks, valleys, humps,  
20 etc., in Dyck and Motzkin paths. This refines the study by Banderier and Flajolet from  
21 2002 on enumeration and asymptotics of lattice paths: we extend here their results to  
22 pattern-avoiding walks/bridges/meanders/excursions. We show that the autocorrelation  
23 polynomial of this forbidden pattern, as introduced by Guibas and Odlyzko in 1981 in  
24 the context of rational languages, still plays a crucial role for our algebraic languages.  
25 *En passant*, our results give the enumeration of some classes of self-avoiding walks, and  
26 prove several conjectures from the On-Line Encyclopedia of Integer Sequences.

27 Finally, we also give the trivariate generating function (length, final altitude, number  
28 of occurrences of the pattern  $p$ ), and we prove that the number of occurrences is normally  
29 distributed and linear with respect to the length of the walk: this is what Flajolet and  
30 Sedgewick call an instance of *Borges's theorem*.

31 **Keywords:** Lattice paths, Dyck paths, Motzkin paths, Łukasiewicz paths, pattern  
32 avoidance, autocorrelation, finite automata, Markov chains, pushdown automata, generat-  
33 ing functions, Wiener–Hopf factorization, kernel method, asymptotic analysis, Gaussian  
34 limit law, Borges' theorem

---


\*<https://me.aau.at/~anasinowski/>


†<https://lipn.fr/~bacher/>


‡<https://lipn.fr/~banderier>

§<https://dmg.tuwien.ac.at/gittenberger/>

 <https://orcid.org/0000-0002-0689-0775>

 <https://orcid.org/0000-0002-9789-7074>

 <https://orcid.org/0000-0003-0755-3022>

 <https://orcid.org/0000-0002-2639-8227>

## 35 Contents

36	<b>1 Introduction</b>	<b>3</b>
37	<b>2 Definitions, notations, autocorrelation polynomial</b>	<b>4</b>
38	<b>3 Lattice paths with forbidden patterns</b>	<b>7</b>
39	<b>4 Automaton, adjacency matrix <math>A</math>, and kernel <math>K =  I - tA </math></b>	<b>9</b>
40	4.1 The automaton and its adjacency matrix . . . . .	9
41	4.2 Algebraic properties of the kernel: link with the autocorrelation polynomial . .	10
42	4.3 Analytic properties of the kernel: Newton polygons and geometry of branches .	14
43	<b>5 Proofs of the generating functions for walks, bridges, meanders, and excursions</b>	<b>17</b>
44	<b>6 Asymptotics of lattice paths avoiding a given pattern</b>	<b>24</b>
45	<b>7 Limit law for the number of occurrences of a pattern</b>	<b>30</b>
46	<b>8 Examples, pushdown automata</b>	<b>32</b>
47	<b>9 Conclusion</b>	<b>38</b>

# 1 Introduction

Combinatorial structures having a rational or an algebraic generating function play a key role in many fields: computer science (e.g. for the analysis of algorithms involving trees, lists, words), computational geometry (integer points in polytopes, maps, graph decomposition), bioinformatics (RNA structure, pattern matching), number theory (integer compositions, automatic sequences and modular properties, integer solutions of varieties), probability theory (Markov chains, directed random walks); see e.g. [8, 22, 42, 78]. Rational functions are often the trace of a structure encodable by an automaton, while algebraic generating functions are often the trace of a structure which has a tree-like recursive specification (typically, a context-free grammar), or which satisfies a functional equation solvable by variants of the *kernel method*.

One of the origin of this method goes back to 1968, when Knuth introduced it to enumerate permutations sortable by a stack; see the solution to Exercise 2.2.1–4 in *The Art of Computer Programming* ([52, pp. 536–537]), which presents a “new method for solving the ballot problem”. For this problem, the solution involves the root of a quadratic polynomial (the so-called kernel). The method was later extended to more general equations (see e.g. [9, 24, 39, 40]). We refer to [14] for more on the long history and the numerous evolutions of this method, which found many applications e.g. for planar maps, permutations, lattice paths, directed animals, polymers, may it be in combinatorics, statistical mechanics, computer algebra, or in probability theory.

In this article, we show how a new extension of this method, which we call *vectorial kernel method*, allows us to solve the enumeration of languages generated by any pushdown automaton.

Since the seminal article by Chomsky and Schützenberger on the link between context-free grammars and algebraic functions [27], which also holds for pushdown automata [76], many articles tackled the enumeration of combinatorial structures via a formal language approach. See e.g. [36, 56, 61] for such an approach on the so-called *generalized Dyck languages*. The words generated by these languages are in bijection with directed lattice paths, and we show how these fundamental objects can be enumerated when they have the additional constraint to avoid a given pattern. For sure, such a class of objects can be described as the intersection of a context-free language and a rational language; therefore, classical closure properties imply that they are directly generated by another (but huge and clumsy) context-free language. Unfortunately, despite the fact that the algebraic system associated with the corresponding context-free grammar is *in theory* solvable by a resultant computation or by Gröbner bases, this leads *in practice* to equations which are so big that no current computer could handle them in memory, even for generalized Dyck languages with only 20 different letters.

Our vectorial kernel method offers a generic and efficient way to tackle such enumeration and bypass these intractable equations. Our approach thus generalizes the enumeration and asymptotics obtained by Banderier and Flajolet [9] for classical lattice paths to lattice paths avoiding a given pattern. This work continues the tradition of investigation of enumerative and asymptotic properties of lattice paths via analytic combinatorics [9, 10, 12, 14]. This allows us to unify the considerations of many articles which investigated natural patterns like peaks, valleys, humps, *etc.*, in Dyck and Motzkin paths, corresponding patterns in trees, compositions, . . . ; see e.g. [5, 16, 25, 33, 38, 53, 57, 60, 62, 66, 69] and all the examples mentioned in our Section 8.

## 2 Definitions, notations, autocorrelation polynomial

Let  $\mathcal{S}$ , the *set of steps*, be some finite subset of  $\mathbb{Z}$ , that contains at least one negative and at least one positive number<sup>1</sup>. A *lattice path with steps from  $\mathcal{S}$*  is a finite word  $w = [v_1, v_2, \dots, v_n]$  in which all letters  $v_i$  belong to  $\mathcal{S}$ , visualized as a directed polygonal line in the plane, which starts at the origin and is formed by successive appending of vectors  $(1, v_1), (1, v_2), \dots, (1, v_n)$ . The letters that form the path are referred to as its *steps*. The *length* of  $w = [v_1, v_2, \dots, v_n]$ , to be denoted by  $|w|$ , is the number of steps in  $w$  (that is,  $n$ ). The *final altitude* of  $w$ , to be denoted by  $\text{alt}(w)$ , is the sum of all steps in  $w$  (that is,  $v_1 + v_2 + \dots + v_n$ ). Visually,  $(|w|, \text{alt}(w))$  is the point where  $w$  terminates.

Under this setting, it is usual to consider two possible restrictions: that the whole path be (weakly) above the  $x$ -axis and that the final altitude be 0 (that is, the path terminates at the  $x$ -axis). Consequently, one considers four classes of lattice paths (see Table 1 page 6 for an illustration):

1. A *walk* is any path as described above.

The corresponding generating function is denoted by  $W(t, u)$ .

Nota bene: in all our generating functions, the variable  $t$  encodes the length of the paths, and the variable  $u$  the final altitude.

2. A *bridge* is a path that terminates at the  $x$ -axis.

The corresponding generating function is denoted by  $B(t)$ . One has<sup>2</sup>  $B(t) = [u^0]W(t, u)$ .

3. A *meander* is a path that stays (weakly) above the  $x$ -axis.

The corresponding generating function is denoted by  $M(t, u)$ .

4. An *excursion* is a path that stays (weakly) above the  $x$ -axis and also terminates at the  $x$ -axis. In other words, an excursion satisfies both restrictions.

The corresponding generating function is denoted by  $E(t)$ . One has  $E(t) = M(t, 0)$ .

For each of these classes, Banderier and Flajolet [9] gave general expressions for the corresponding generating functions and the asymptotics of their coefficients. The *step polynomial* of the set of steps  $\mathcal{S}$ , denoted by  $P(u)$ , is defined by

$$P(u) = \sum_{s \in \mathcal{S}} u^s. \quad (1)$$

The smallest (negative) number in  $\mathcal{S}$  is denoted by  $-c$ , and the largest (positive) number in  $\mathcal{S}$  is denoted by  $d$ : that is<sup>3</sup>, if one orders the terms of  $P(u)$  by the powers of  $u$ , one has  $P(u) = u^{-c} + \dots + u^d$ .

<sup>1</sup>Without this assumption, the corresponding models are easy to solve: they are leading to rational generating functions. Our enumeration results still hold in these easy cases, but are then just classical folklore.

<sup>2</sup>The notation  $[t^n]F(t)$  stands for the coefficient of  $t^n$  in the power series  $F(t)$ .

<sup>3</sup>Some weights (or probabilities) could be associated with each step. Most of the results would generalize accordingly, we omit them in this article to keep readability.

120 We extend the study of Banderier and Flajolet by considering lattice paths with step set  $\mathcal{S}$   
 121 that avoid a certain *pattern* of length  $\ell$ , that is, an a priori fixed path

$$122 \quad p = [a_1, a_2, \dots, a_\ell] \quad (\text{where the } a_i \text{'s belong to } \mathcal{S}). \quad (2)$$

123 To be precise, we define an *occurrence* of  $p$  in a lattice path  $w$  as a (contiguous) substring  
 124 which coincides with  $p$ . If there is no occurrence of  $p$  in  $w$ , we say that  $w$  *avoids*  $p$ . For  
 125 example, the path  $[1, 1, 1, 2, -3, 1, 2]$  has two occurrences of  $[1, 1]$ , two occurrences of  $[1, 2]$ ,  
 126 and it avoids  $[2, 1]$ .

127 Before we state our results, we introduce some notations.

128 A *presuffix* of  $p$  is a non-empty string that occurs in  $p$  both as a prefix and as a suffix. In  
 129 particular, the whole word  $p$  is a (trivial) presuffix of itself. If  $p$  has one or several non-trivial  
 130 presuffixes, we say that  $p$  exhibits an *autocorrelation* phenomenon. For example, the pattern  $p =$   
 131  $[1, 1, -2, 1, -2]$  has no autocorrelation. In contrast, the pattern  $p = [1, 1, 2, -3, 1, 1, 2, -3, 1, 1]$   
 132 has three non-trivial presuffixes:  $[1]$ ,  $[1, 1]$ , and  $[1, 1, 2, -3, 1, 1]$ , and thus in this case we have  
 133 autocorrelation.

134 While analysing the Boyer–Moore string searching algorithm and properties of periodic words,  
 135 Guibas and Odlyzko introduced in 1981 [45] what turns out to be one of the key characters of  
 136 our article, the autocorrelation polynomial<sup>4</sup> of the pattern  $p$ . For any given word  $p$ , let  $\mathcal{Q}$  be  
 137 the set of its presuffixes; the *autocorrelation polynomial* of  $p$  is

$$138 \quad R(t, u) := \sum_{q \in \mathcal{Q}} t^{|\bar{q}|} u^{\text{alt}(\bar{q})}, \quad (3)$$

139 where  $\bar{q}$  denotes the complement of  $q$  in  $p$  (i.e.,  $p = q\bar{q}$ ). The choice of the letter  $R$  to denote  
 140 this polynomial is a mnemonic of the fact that it encodes the *relations* of the pattern  $p$ .

141 For example, consider the pattern  $p = [1, 1, 2, 3, 1, 1, 2, 3, 1, 1]$ . Its four presuffixes produce  
 142 four terms of  $R(t, u)$  as follows:

presuffix $q$	length of its complement $ \bar{q} $	final altitude of its complement $\text{alt}(\bar{q})$
$[1]$	9	15
$[1, 1]$	8	14
$[1, 1, 2, 3, 1, 1]$	4	7
$[1, 1, 2, 3, 1, 1, 2, 3, 1, 1]$	0	0

144 Therefore, for this  $p$  we have  $R(t, u) = 1 + t^4 u^7 + t^8 u^{14} + t^9 u^{15}$ .

145 Notice that if for some  $p$  no autocorrelation occurs, then we have  $\mathcal{Q} = \{p\}$  and therefore  
 146  $R(t, u) = 1$ .

<sup>4</sup>A similar notion also appears in the work of Schützenberger on synchronizing words [77]. It should also be added that the autocorrelations of a pattern allow to explain some famous paradoxes, e.g. why, in a random sequence of coin tossings, the pattern HTT is likely to occur much sooner (after 8 tosses on average) than the pattern HHH (needing 14 tosses on average): “All patterns are not born equal!” as Flajolet and Sedgewick pleasantly write in [42, Example IV.11].

Finally, we define the *kernel* as the following Laurent polynomial:

$$K(t, u) := (1 - tP(u))R(t, u) + t^{|p|}u^{\text{alt}(p)}. \tag{4}$$

It will be shown in Proposition 4.4 of Section 4.2 that each root  $u(t)$  of  $K(t, u) = 0$  is either *small* (meaning  $\lim_{t \rightarrow 0} u(t) = 0$ ) or *large* (meaning  $\lim_{t \rightarrow 0} |u(t)| = \infty$ ). Moreover, the number of small roots (to be denoted by  $e$ ) is the absolute value of the lowest power of  $u$ , and the number of large roots (to be denoted by  $f$ ) is the highest power of  $u$  in  $K(t, u)$ . In particular, if  $R(t, u) = 1$ , then we have  $e = \max\{c, -\text{alt}(p)\}$ , and  $f = \max\{d, \text{alt}(p)\}$ .

Equipped with these definitions, we can give a first illustration of some of our results in Table 1. From now on, we use the notations  $W/B/M/E$  for generating functions enumerating paths constrained to avoid a pattern  $p$ .

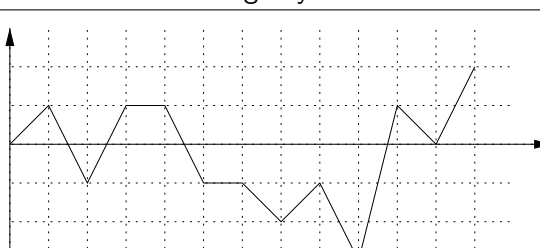
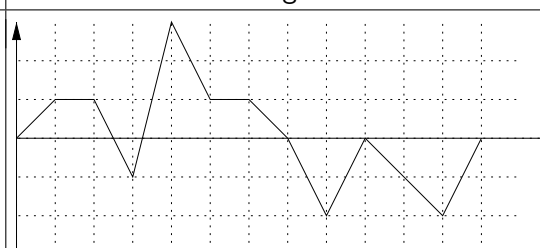
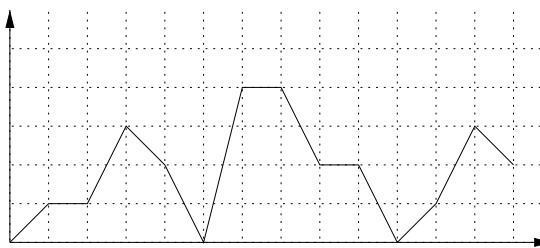
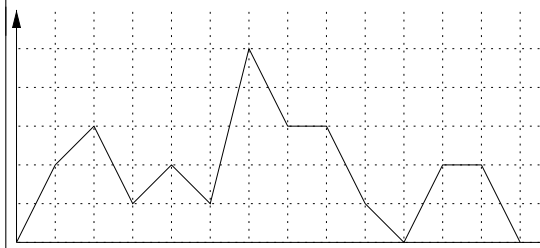
	ending anywhere	ending at 0
on $\mathbb{Z}$	 <p>walks</p> $W(t, u) = \frac{R(t, u)}{K(t, u)}$	 <p>bridges</p> $B(t) = - \sum_{i=1}^e \frac{u'_i(t)}{u_i(t)} \frac{R(t, u_i)}{K_t(t, u_i)}$
on $\mathbb{N}$	 <p>meanders</p> $M(t, u) = \frac{R(t, u)}{K(t, u)} \prod_{i=1}^e (u - u_i(t))$	 <p>excursions</p> $E(t) = \frac{(-1)^{e+1}}{t} \prod_{i=1}^e u_i(t)$

Table 1: **Summary of our results.** For the four classes of paths and for any set of steps encoded by  $P(u)$ , we find the generating functions of such lattice paths that avoid a pattern  $p$ . The formulas involve the autocorrelation polynomial  $R(t, u)$  of  $p$ , and the  $e$  small roots  $u_i(t)$  of the kernel  $K(t, u)$ . For meanders and excursions, the table shows the formula for the special case when  $p$  is a meander itself; in the general case, these formulas might have a different prefactor (see Theorem 3.2 below).

In the next section, we present in a more detailed way the generating functions of all these paths.

### 3 Lattice paths with forbidden patterns

In this section, we state our main results for the generating functions  $W/B/M/E$  of walks, bridges, meanders, excursions, constrained to avoid a given pattern  $p$ . In all the theorems, we denote by  $\ell$  the length of the pattern  $p$ , and by  $\text{alt}(p)$  its final altitude.

**Theorem 3.1** (Generating function of walks, generic pattern). *Let  $\mathcal{S}$  be a set of steps, and let  $p$  be a pattern with steps from  $\mathcal{S}$ .*

1. The bivariate generating function for **walks** avoiding the pattern  $p$  is

$$W(t, u) = \frac{R(t, u)}{K(t, u)}. \quad (5)$$

If one does not keep track of the final altitude, this yields

$$W(t) = W(t, 1) = \frac{1}{1 - tP(1) + t^\ell/R(t, 1)}. \quad (6)$$

2. The generating function for **bridges** avoiding the pattern  $p$  is

$$B(t) = - \sum_{i=1}^e \frac{u'_i(t)}{u_i(t)} \frac{R(t, u_i)}{K_t(t, u_i)}, \quad (7)$$

where  $u_1(t), \dots, u_e(t)$  are the small roots of the kernel  $K(t, u)$ , as defined in (4).

In order to find the generating functions of meanders and excursions, we shall use a novel approach that we call **vectorial kernel method**. With this method, we obtain formulas that generalize those from the Banderier–Flajolet study (in which they considered classical paths, without the additional pattern constraint).

**Theorem 3.2** (Generating function of meanders and excursions, generic pattern). *The bivariate generating function of **meanders** avoiding the pattern  $p$  is*

$$M(t, u) = \frac{G(t, u)}{u^e K(t, u)} \prod_{i=1}^e (u - u_i(t)), \quad (8)$$

where  $u_1(t), \dots, u_e(t)$  are the small roots of  $K(t, u) = 0$ ,  $G(t, u)$  is a polynomial in  $u$  (and a formal power series in  $t$ ) which will be characterized in (27).

For **excursions**, just take the above closed form for  $u \rightarrow 0$ .

Although we present here this formula under the paradigm of patterns in lattice paths, it holds in fact much more generally for the enumeration of words generated by any pushdown automaton. This shall become transparent via the proof given Section 5 and via the examples from Section 8.

We now introduce two classes of patterns for which the factor  $G(t, u)$  in Formula (8) has a very nice simple shape.

**Definition 3.3** (Quasimeanders, reversed meanders). *A **quasimeander** is a lattice path which does not cross the  $x$ -axis, except, possibly, at the last step. A **reversed meander** is a lattice path whose terminal point has a (strictly) smaller  $y$ -coordinate than all other points. (Notice that the empty path is both a quasimeander and a reversed meander.)*

199 **Theorem 3.4** (Generating function of meanders, quasimeander pattern subcase). *Let  $p$  be a*  
 200 *quasimeander.*

201 1. *The bivariate generating function of **meanders** avoiding the pattern  $p$  is*

$$202 \quad M(t, u) = \frac{R(t, u)}{u^c K(t, u)} \prod_{i=1}^c (u - u_i(t)), \quad (9)$$

203 *where  $u_1(t), \dots, u_c(t)$  are the small roots<sup>5</sup> of  $K(t, u) = 0$ .*

204 *If one does not keep track of the final altitude, this yields*

$$205 \quad M(t) = M(t, 1) = \frac{R(t, 1)}{K(t, 1)} \prod_{i=1}^c (1 - u_i(t)). \quad (10)$$

206 2. *The generating function for **excursions** avoiding the pattern  $p$  is*

$$207 \quad E(t) = M(t, 0) = \begin{cases} \frac{(-1)^c}{-t} \prod_{i=1}^c u_i(t) & \text{if } \text{alt}(p) > -c, \\ \frac{(-1)^c}{t^\ell - t} \prod_{i=1}^c u_i(t) & \text{if } \text{alt}(p) = -c. \end{cases} \quad (11)$$

208 **Theorem 3.5** (Generating function of meanders, reversed meander pattern subcase). *Let  $p$  be*  
 209 *a reversed meander.*

210 1. *The bivariate generating function for **meanders** avoiding the pattern  $p$  is*

$$211 \quad M(t, u) = \frac{1}{u^e K(t, u)} \prod_{i=1}^e (u - u_i(t)), \quad (12)$$

212 *where  $u_1(t), \dots, u_e(t)$  are the small roots of  $K(t, u) = 0$ .*

213 *If one does not keep track of the final altitude, this yields*

$$214 \quad M(t) = M(t, 1) = \frac{1}{K(t, 1)} \prod_{i=1}^e (1 - u_i(t)). \quad (13)$$

215 2. *The generating function for **excursions** avoiding the pattern  $p$  is*

$$216 \quad E(t) = M(t, 0) = \frac{(-1)^e}{D(t)} \prod_{i=1}^e u_i(t), \quad (14)$$

217 *where  $D(t) := [u^0]u^e K(t, u)$  is either some power of  $t$ , or a difference of two powers of  $t$*   
 218 *(similarly to (11), but with more cases that will be specified in the proof in Section 5).*

219 **Remark 3.6** (Compatibility with the literature on classical lattice paths). Notice that for these  
 220 four classes of lattice paths, if one forbids a pattern of length 1 or if one uses symbolic weights  
 221 for each step, this recovers the formulas from Banderier and Flajolet [9].

222 The proof will involve the adjacency matrix of a finite automaton which encodes accumulating  
 223 the prefixes of the forbidden pattern. This will be introduced in the next section.

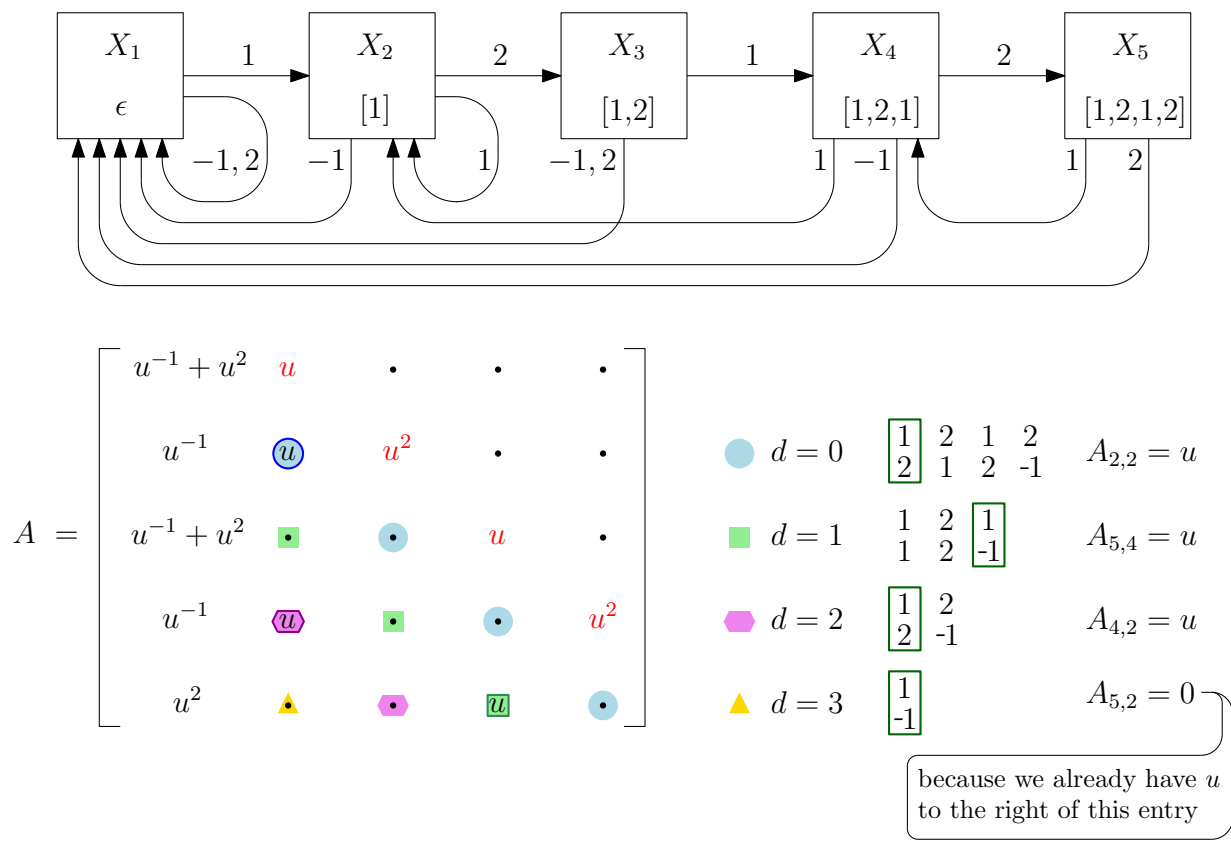
<sup>5</sup>It will be shown in the proof that in this case we have  $e = c$ .



## 224 4 Automaton, adjacency matrix $A$ , and kernel $K = |I - tA|$

### 225 4.1 The automaton and its adjacency matrix

226 In this section, we introduce an automaton which will allow us to tackle pattern avoidance. As  
 227 explained in Section 2, any path is seen as word (the steps are transformed into alphabet letters).  
 228 Let  $p = [a_1, \dots, a_\ell]$  be the “forbidden” pattern. Sharing the spirit of the Knuth–Morris–Pratt  
 229 algorithm, we build an automaton with  $\ell$  states, where each state corresponds to a proper  
 230 prefix of  $p$  collected so far. We label these states  $X_1, \dots, X_\ell$ : the first state is labelled by the  
 231 empty word (namely,  $X_1 = \epsilon$ ) and the next states are labelled by proper prefixes of  $p$  (namely,  
 232  $X_i = [a_1, a_2, \dots, a_{i-1}]$ ). If the automaton reads a path  $w$ , then it ends in the state labelled by  
 233 the longest prefix of  $p$  that coincides with a suffix of  $w$ . Note that the automaton is completely  
 234 determined by  $P(u)$  and  $p$ .



236 Figure 1: The automaton and the adjacency matrix  $A$  for  $\mathcal{S} = \{-1, 1, 2\}$  and the pattern  
 237  $p = [1, 2, 1, 2, -1]$ . (The 0 entries of  $A$  are replaced by dots.) The powers of  $u$  in the  
 238 superdiagonal (in red) correspond to the pattern. For each other diagonal (consisting of entries  
 239 surrounded in colour), there is only one nonzero entry, according to a property illustrated in the  
 240 bottom right part of the figure and explained hereafter.

Let us describe the transitions of this automaton more precisely. For  $i, j \in \{1, \dots, \ell\}$ , we have an arrow labelled  $\lambda$  from the state  $X_i$  to the state  $X_j$  if  $j$  is the maximum number such that  $X_j$  is a suffix of  $X_i\lambda$ . Its adjacency matrix (also called transition or transfer matrix by some authors) will be denoted by  $A$ : it is an  $\ell \times \ell$  matrix, and its  $(i, j)$  entry is the sum of all terms  $u^\lambda$  such that there is an arrow labelled  $\lambda$  from  $X_i$  to  $X_j$ . See Figure 1 for an example. The following general properties of  $A$  are an easy consequence of its combinatorial definition:

- For all  $i, j$  such that  $j > i + 1$ , we have  $A_{i,j} = 0$ .
- The superdiagonal entries are  $A_{i,i+1} = u^{a_i}$ .
- For each  $j$  such that  $2 \leq j \leq \ell$ , every entry in the  $j$ -th column is either 0 or  $u^{a_{j-1}}$ , depending on a procedure explained below.
- The first column is such that all rows sum to  $P(u)$ , except for the last row, where this entry is  $P(u) - u^{a_\ell}$  (because the transition with  $a_\ell$  is forbidden in this row, as it would create an occurrence of the pattern  $p$ ).

The entries of  $A$  along the diagonals surrounded in colour can be determined by the following procedure, which is also illustrated in the bottom right of Figure 1. For  $d = 0, \dots, \ell - 2$ :

1. If  $[a_1, a_2, \dots, a_{\ell-d-1}] = [a_{d+2}, a_{d+3}, \dots, a_\ell]$ , then all the entries  $A_{i,j}$  with  $i - j = d$ ,  $j \geq 2$ , are 0.
2. Otherwise, if  $k$  is the smallest number such that  $a_k \neq a_{d+1+k}$ , then  $A_{d+k+1, k+1} = u^k$ , unless a smaller  $d$  yielded  $u^k$  in the same row, to the right of this position (if this happens,  $A_{d+k+1, k+1} = 0$ ).

This more intimate knowledge of the structure of the adjacency matrix  $A$  will help us to compute some related determinants.

## 4.2 Algebraic properties of the kernel: link with the autocorrelation polynomial

It is well known that the matrix  $1/(I - tA) = \text{adj}(I - tA)/\det(I - tA)$  (where  $I$  is the identity matrix) plays an important role in the enumeration of walks. We will see that this adjoint and this determinant also play a fundamental role in the enumeration of meanders. In fact, the role of  $\det(I - tA)$  in our study is the analogue of the role played by  $1 - tP(u)$  in the study of Banderier and Flajolet [9], but, as we shall see, the situation is more involved in our case.

**Proposition 4.1** (Structure of the kernel). *Let  $\mathcal{S}$  be a set of steps, and let  $p$  be a pattern with steps from  $\mathcal{S}$ . Then the adjacency matrix  $A$  of the automaton satisfies*

$$(1 \ 0 \ \dots \ 0) \text{adj}(I - tA) (1 \ 1 \ \dots \ 1)^\top = R(t, u), \quad (15)$$

$$\det(I - tA) = K(t, u) = (1 - tP(u))R(t, u) + t^{|p|}u^{\text{alt}(p)}, \quad (16)$$

where  $R(t, u)$  and  $K(t, u)$  are the kernel and the autocorrelation polynomial, as defined in (3) and (4). In particular, in the case without autocorrelation we have  $\det(I - tA) = 1 - tP(u) + t^{|p|}u^{\text{alt}(p)}$ , and the sum of the entries in the first row of  $\text{adj}(I - tA)$  is 1.

*Proof.* Equations (15) and (16) will be proven in the course of the proof of Theorem 3.1.  $\square$

278 Another important quantity related to the adjacency matrix is what we call the *autocorrelation*  
 279 *vector*, defined as  $\vec{v} := \text{adj}(I - tA) \cdot \vec{\mathbf{1}}$ , where  $\vec{\mathbf{1}}$  denotes the column vector  $(1, 1, \dots, 1)^\top$  of  
 280 size  $\ell \times 1$ . In Proposition 4.2 we give a simple combinatorial description of this vector  $\vec{v}$ ; in  
 281 particular, this description has the advantage of smaller computational cost than getting it via a  
 282 case by case matrix inversion.

**Proposition 4.2** (Structure of the autocorrelation vector). *The autocorrelation vector  $\vec{v} := \text{adj}(I - tA) \cdot \vec{\mathbf{1}}$  satisfies*

$$\vec{v} = R(t, u)\vec{\mathbf{1}} - \vec{s}(t, u),$$

283 where the  $j$ -th entry of  $\vec{s}(t, u)$  is a polynomial, the generating function of a finite set  $S_j$  of  
 284 walks (which we call the “subtracted set” of walks). The subtracted set  $S_j$  is the finite set of  
 285 walks of length smaller than  $\ell$  having the factorization  $w.a_\ell.\bar{q}$ , where

- 286 1.  $w$  is any walk starting in state  $X_j$  and ending in state  $X_\ell$ ,
- 287 2.  $a_\ell$  is the last step of the pattern  $p$ ,
- 288 3.  $\bar{q}$  corresponds to a term of  $R(t, u)$  (that is,  $\bar{q}$  is the complement of some presuffix  $q$  of  $p$ ,  
 289 i.e.  $p = q\bar{q}$ ).

290 As this proposition is a little bit hard to digest, we give an example after the proof (Ex. 4.3),  
 291 and the reader will see that the situation is in fact simpler than what she may have feared.

292 *Proof.* First note that  $(I - tA)\vec{v} = K(t, u)\vec{\mathbf{1}}$  and  $A\vec{\mathbf{1}} = P(u)\vec{\mathbf{1}} - u^{a_\ell}\vec{e}_\ell$ , where  $\vec{e}_\ell = (0, \dots, 0, 1)^\top$ .  
 293 Now define  $\vec{s}$  by  $\vec{s} := R(t, u)\vec{\mathbf{1}} - \vec{v}$ . Then

$$\begin{aligned} 294 (I - tA)\vec{s} &= -(I - tA)\vec{v} + (I - tA)R(t, u)\vec{\mathbf{1}} \\ 295 &= -K(t, u)\vec{\mathbf{1}} + R(t, u)\vec{\mathbf{1}} - tR(t, u)A\vec{\mathbf{1}} \\ 296 &= tu^{a_\ell}R(t, u)\vec{e}_\ell - t^\ell u^{\text{alt}(p)}\vec{\mathbf{1}} \end{aligned}$$

297 where  $\vec{e}_\ell = (0, \dots, 0, 1)$ . This implies

$$298 \vec{s} = (I - tA)^{-1}\vec{e}_\ell tu^{a_\ell}R(t, u) - (I - tA)^{-1}\vec{\mathbf{1}} t^\ell u^{\text{alt}(p)}. \quad (17)$$

299 In this form, it is now easy to give a combinatorial interpretation to the column vector  $\vec{s}$ :  
 300 its  $j$ -th component is the generating function of certain lattice paths for which the associated  
 301 automaton is in state  $X_j$  before the walk does its first step.

302 The entries of  $(I - tA)^{-1}$  are the generating functions of lattice paths not containing the  
 303 pattern  $p = [a_1, a_2, \dots, a_\ell]$ , specifically: the  $(i, j)$ -entry of this matrix is the generating function  
 304 of such walks that start in state  $X_i$  and end in state  $X_j$ . Therefore, the  $j$ -th component of  
 305  $(I - tA)^{-1}\vec{e}_\ell tu^{a_\ell}R(t, u)$  is the generating function of all lattice paths that are composed as  
 306 follows: first, start with a  $p$ -avoiding walk  $w$  that begins in the state  $X_j$  and ends in state  $X_\ell$ ,  
 307 followed by the single step  $a_\ell$ , and then finally by some complement  $\bar{q}$  of some presuffix  $q$  of  
 308 the pattern  $p$ . Note that if the walk  $w$  is long enough and  $\bar{q}$  is not the empty sequence, then  
 309 adding the step  $a_\ell$  to  $w$  makes it end with  $q$ . So,  $w.a_\ell.\bar{q}$  has an occurrence of  $p$  at the very end.  
 310 Furthermore, having added the step  $a_\ell$  to  $w$  an occurrence of  $p$  is created unless  $w$  was shorter  
 311 than  $\ell$ .

312 The second term,  $(I - tA)^{-1} \vec{\mathbf{1}} t^\ell u^{\text{alt}(p)}$ , is the generating function of all walks avoiding  $p$  to  
 313 which we add  $p$  at their end. We may have a single occurrence of  $p$  at the end of the walk or  
 314 two overlapping patterns if the first part of the walk ended with some prefix of  $p$ , which needs  
 315 only a presuffix to complete  $p$ ; the completion is done by the final occurrence of  $p$ .

316 We observed that the two terms on the right-hand side of (17) correspond to sets of lattice  
 317 paths. Clearly, any walk ending with  $p$  and having only one occurrence of  $p$  is in both sets. And  
 318 so is any walk having two overlapping occurrences of  $p$  at its very end. However, the first set  
 319 contains in addition paths being too short for having an occurrence of  $p$ : they are precisely  
 320 the set of paths described in the assertion. Since this set is finite, its generating function is a  
 321 polynomial.  $\square$

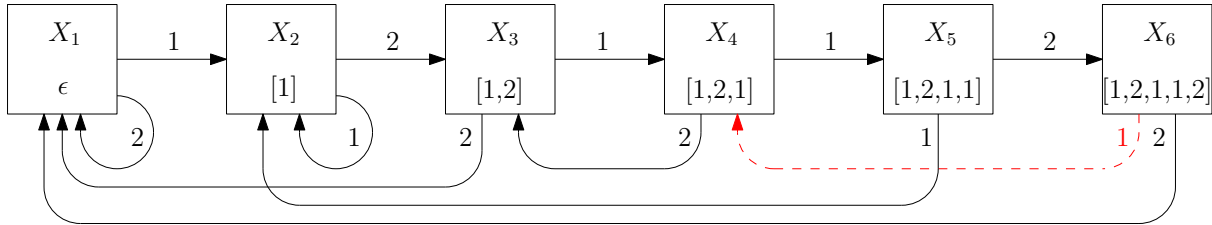
322 **Example 4.3.** Consider as an example the lattice path model with some step set  $\mathcal{S} \supseteq \{1, 2\}$   
 323 and the pattern  $p = [1, 2, 1, 1, 2, 1]$ . The autocorrelation polynomial is  $R(t, u) = 1 + t^3 u^4 + t^5 u^7$ ,  
 324 and the autocorrelation vector  $\vec{v} = \text{adj}(I - tA) \cdot \vec{\mathbf{1}}$  has the following structure:

$$\begin{aligned}
 \vec{v} &= \begin{pmatrix} 1 + t^3 u^4 + t^5 u^7 \\ 1 + t^3 u^4 \\ 1 + t^3 u^4 - t^4 u^5 + t^5 u^7 \\ 1 \\ 1 - t^2 u^3 + t^3 u^4 \\ 1 - tu + t^3 u^4 - t^4 u^5 + t^5 u^7 \end{pmatrix} = (1 + t^3 u^4 + t^5 u^7) \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 0 \\ t^5 u^7 \\ t^4 u^5 \\ t^3 u^4 + t^5 u^7 \\ t^2 u^3 + t^5 u^7 \\ tu + t^4 u^5 \end{pmatrix}. \quad (18)
 \end{aligned}$$

$\uparrow$   
 $R(t, u)$

$\uparrow$   
 $\vec{s}(t, u)$

328 Now, let us interpret the polynomials subtracted from  $R(t, u)$  according to Proposition 4.2. To  
 329 this end, it is convenient to introduce a variant of the automaton defined at the beginning of  
 330 this section (see Figure 1), in which we add a transition with the letter  $a_\ell$  from the last state  
 331 to the state indexed by the longest presuffix which is shorter than the whole pattern (in our  
 332 case, it is the state labelled by  $[1, 2, 1]$ ). Marking this new transition allows us to count the  
 333 occurrences of the pattern  $p$ , instead of just counting walks avoiding this pattern. (Section 7  
 334 will be dedicated to this topic.) This gives the following automaton.



335 Figure 2: Illustration to Example 4.3: the automaton marking occurrences of the pattern 121121.

336 With this automaton, it is easy to exhaustively list the walks having the factorization  $w.a_\ell.\bar{q}$   
 337 mentioned in Proposition 4.2 (n.b.: the suffix  $w$  does not contain the pattern  $p$ , i.e., it is a  
 338 word generated by the automaton *without* the added red transition); this gives a combinatorial  
 339 explanation of the components of the vector  $\vec{s}(t, u)$ .

- 340 • For the first component of  $\vec{s}$ , the subtracted set of walks is in fact empty. Indeed, each  
 341 path  $w$  from  $X_1$  to  $X_6$ , followed by the step  $a_\ell$ , has at least length 6. So, there are no  
 342 walks that satisfy the condition of being shorter than  $\ell$ . Therefore, we have  $\vec{s}_1 = 0$  (which  
 343 holds in full generality, implying  $\vec{v}_1 = R(t, u)$  for any pattern).
- 344 • For the second component of  $\vec{s}$ , we start in state  $X_2$ . The length of the prefactor  $w$  has  
 345 to be smaller than  $\ell - 1$ , and the only way to reach  $X_6$  in strictly fewer than 5 steps is  
 346 the path  $X_2 \rightarrow X_3 \rightarrow X_4 \rightarrow X_5 \rightarrow X_6$ , which generates 2112, to which we append the  
 347 last letter  $a_\ell = 1$ . Then the only possible  $\bar{q}$  is the empty one (otherwise the path will  
 348 be too long). Thus, the subtracted set of walks is  $\mathcal{S}_2 = \{2112.1.\varepsilon\}$ , and its generating  
 349 function is  $t^5u^7$ .
- 350 • For the third component of  $\vec{s}$ , the only sufficiently short walk from  $X_3$  to  $X_6$  is  $X_3 \rightarrow$   
 351  $X_4 \rightarrow X_5 \rightarrow X_6$ , and the only possible choice of  $\bar{q}$  is  $\varepsilon$ . Thus, the subtracted set of walks  
 352 is  $\mathcal{S}_3 = \{112.1.\varepsilon\}$ , and its generating function is  $t^4u^5$ .
- 353 • For the fourth component of  $\vec{s}$ , we have *two* ways to go from  $X_4$  to  $X_6$ :  $X_4 \rightarrow X_5 \rightarrow X_6$   
 354 and  $X_4 \rightarrow X_3 \rightarrow X_4 \rightarrow X_5 \rightarrow X_6$ . In both cases, the only possible choice of  $\bar{q}$  is  $\varepsilon$ .  
 355 Thus, the subtracted set of walks is  $\mathcal{S}_4 = \{12.1.\varepsilon, 2112.1.\varepsilon\}$ , and its generating function  
 356 is  $t^3u^4 + t^5u^7$ .
- 357 • For the fifth component of  $\vec{s}$ , starting in  $X_5$ , only one short path to  $X_6$  is possible:  
 358  $X_5 \rightarrow X_6$ , but in this case we have two choices for  $\bar{q}$ :  $\varepsilon$  and 121. This gives the  
 359 subtracted set of walks  $\mathcal{S}_5 = \{2.1.\varepsilon, 2.1.121\}$ , and its generating function is  $t^2u^3 + t^5u^7$ .
- 360 • For the sixth component of  $\vec{s}$ , starting in  $X_6$ , the subtracted set of walks is  $\mathcal{S}_6 =$   
 361  $\{\varepsilon.1.\varepsilon, \varepsilon.1.121\}$ , and its generating function is  $tu + t^4u^5$ .

362 In conclusion, we combinatorially got  $\vec{s} = (0, t^5u^7, t^4u^5, t^3u^4 + t^5u^7, t^2u^3 + t^5u^7, tu + t^4u^5)^\top$ ,  
 363 which agrees with Formula (18). ■

364 This autocorrelation vector  $\vec{v}$  plays a role for many problems of lattice path pattern enumera-  
 365 tion: it allows rewriting the initial functional equation associated to the problem in a condensed  
 366 system, with less unknowns. We shall see in the proof of Theorem 3.2 (for the generating  
 367 function of meanders) that  $\vec{v}$  will act as an *eigenvector* for a system playing a key role in our  
 368 enumeration.

### 4.3 Analytic properties of the kernel: Newton polygons and geometry of branches

Let us end this section with the proof of an important property of the kernel, the number of “small” and “large” roots  $u(t)$  of  $K(t, u) = 0$ :

**Proposition 4.4** (Small and large roots of the kernel  $K$ ). *All roots  $u(t)$  of  $K(t, u)$  are either small (meaning  $\lim_{t \rightarrow 0} u(t) = 0$ ) or large (meaning  $\lim_{t \rightarrow 0} |u(t)| = \infty$ ). Let  $d_K$  denote the degree of  $K(t, u)$  in  $u$ , i.e.,  $d_K = \max\{j \mid [u^j]K(t, u) \neq 0\}$ , and let  $l_K$  denote the lowest power of  $u$  in the monomials of  $K$ , i.e.,  $l_K = \min\{j \mid [u^j]K(t, u) \neq 0\}$ . Then,  $K$  has  $e$  small roots and  $f$  large roots, where  $e = \max(0, -l_K)$  and  $f = \max(0, d_K)$*

**Remark 4.5.** The nested  $\max / \min$  is needed for the cases where either  $l_K > 0$  or  $d_K < 0$ . An example for such a model is  $\mathcal{S} = \{-1, 3\}$  and  $p = [-1, -1]$ , where we have  $K = 1 - u^3t - u^2t^2$ .

*Proof.* Recall that the kernel  $K$  is defined as  $K(t, u) = (1 - tP(u))R(u) + t^\ell u^{\text{alt}(p)}$ . Now, consider  $u^e K(t, u)$ , which is a polynomial, and draw for each of its monomials  $t^{r_1} u^{r_2}$  a point  $(r_1, r_2)$  in the plane. This gives a set of points  $\mathcal{P}$ . The *Newton polygon*<sup>6</sup> of  $K$  is the boundary of the convex hull of  $\mathcal{P}$ ; of particular interest to us are the segments of this polygon that are visible when we look from the left: their slope (it is always a rational number) will give the exponent of the Puiseux expansions of each root of  $K$ . Figure 3 shows schematically all possible shapes for the case  $R(t, u) = 1$  (if  $R(t, u) \neq 1$ , the possible shapes are more diverse). Each segment (in green) which lies below the point  $(0, e)$  (they all have a negative slope) gives the Puiseux expansion of a set of small roots. Each segment (in red) which lies above the point  $(0, e)$  (they all have a positive slope) gives the Puiseux expansion of a set of large roots. We now focus on the small roots and prove the proposition for them. (The proof for large roots is similar and can be obtained from this one by the replacement  $u \mapsto \frac{1}{u}$ , which corresponds to the vertical reflection of the Newton polygon.)

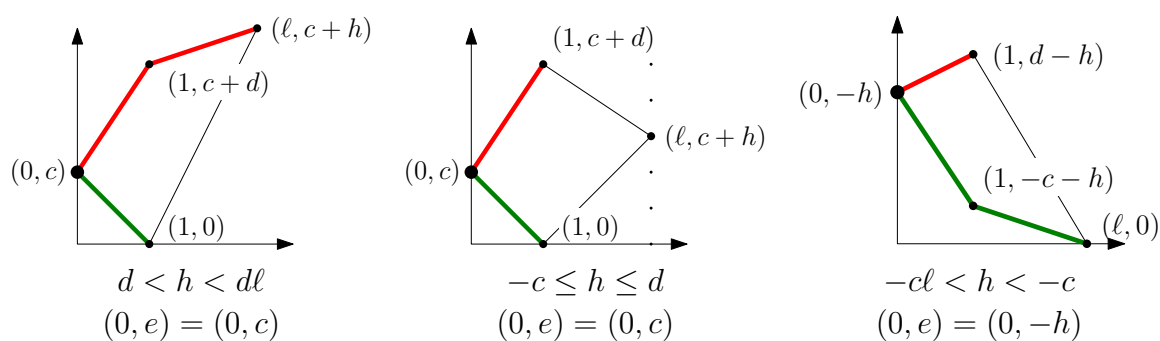


Figure 3: Three possibilities for the Newton polygon of the kernel  $K(t, u)$ . This classification depends on the final altitude  $h$  of the pattern  $p$ , and is exhaustive if  $R(t, u) = 1$ . Each point  $(i, j)$  corresponds to a monomial  $t^i u^j$  of the numerator of  $K$ . The slopes of the convex hull segments on the left give the Puiseux behaviour at 0 of the small and large roots  $u_i$  and  $v_j$ .

<sup>6</sup>See [34, pp. 106–112] or [51, Chapter 6.3] for a nice presentation of the theory of Newton polygons.

392 If the Newton polygon has no segment of negative slope (the ones drawn in green in  
 393 Figure 3), then  $K$  is a polynomial in  $u$  (and in  $t$  as well) having the constant term 1. Hence,  
 394  $K(t, u) = 1 + Q(t, u)$  where  $Q(t, u)$  is a polynomial in  $t$  and  $u$  with  $\lim_{(t,u) \rightarrow (0,0)} Q(t, u) = 0$ .  
 395 This implies that any non large functions  $u(t)$  can never compensate the constant term near  
 396  $t = 0$ , and so there are zero small roots and  $d_K$  large roots, in accordance with  $e = 0$  and  
 397  $f = d_K$ .

398 If the Newton polygon has at least one such segment of negative slope, then consider one  
 399 of them, denote it by  $\Sigma$ , and let  $-\beta/\alpha$  be its slope. The two endpoints of  $\Sigma$  correspond to  
 400 monomials of  $u^e K(t, u)$  (some other points of  $\Sigma$  also possibly do). Any two such monomials  
 401 differ by some power of  $t^\alpha u^{-\beta}$ . If  $u \sim C \cdot t^{\alpha/\beta}$  as  $t \rightarrow 0$ , then  $t^\alpha u^{-\beta} \sim C^{-\beta}$ , which is some  
 402 nonzero constant. Thus, all monomials that correspond to a point of the segment  $\Sigma$  are of  
 403 the same order of magnitude as Puiseux series in  $t$ . What about the other monomials? By  
 404 construction of the Newton polygon, their corresponding points must lie above the line containing  
 405 the segment  $\Sigma$ . Hence, such a point can be represented in the form  $t^{r_1+j} u^{r_2+k}$  where  $(r_1, r_2) \in \Sigma$   
 406 and  $k > -j\beta/\alpha$ . We now use the notation  $f(t) \approx g(t)$  for  $\lim_{t \rightarrow 0^+} f(t)/g(t) = \text{constant}$ . If  
 407  $u \approx t^{\alpha/\beta}$ , then one has

$$t^{r_1+j} u^{r_2+k} \approx t^{r_1+j+k\alpha/\beta} u^{r_2}.$$

408 Since  $k > -j\beta/\alpha$ , we have  $j + k\alpha/\beta > 0$  and thus the monomial  $t^{r_1+j} u^{r_2+k}$  has a smaller  
 409 order of magnitude than the monomials corresponding to the points of  $\mathcal{P}$  on the segment  $\Sigma$ ,  
 410 like  $t^{r_1} u^{r_2}$ . The arguments above show that if  $u \approx t^{\alpha/\beta}$ , then

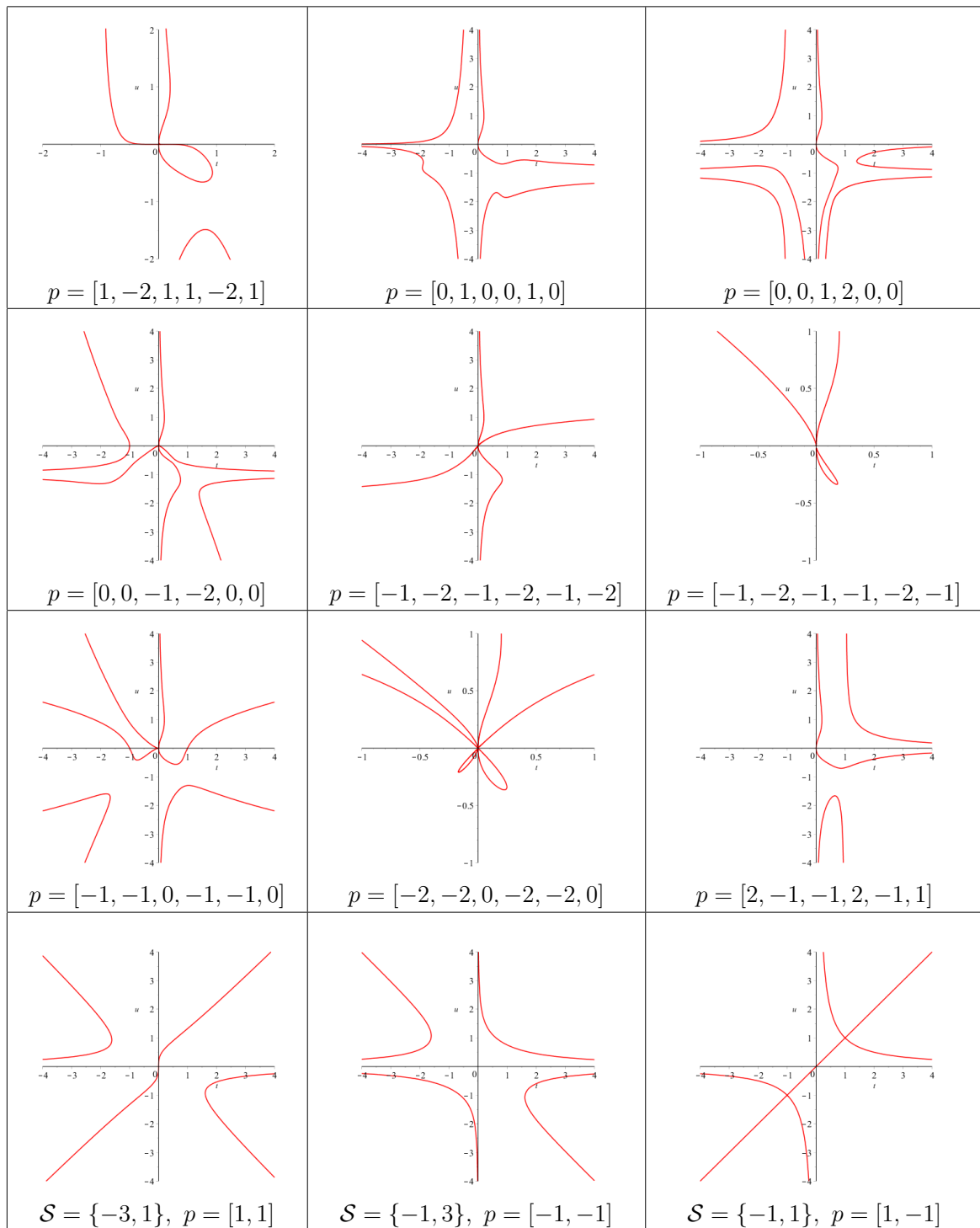
$$411 \quad u^e K(t, u) \approx \sum_{(r_1, r_2) \in \mathcal{P} \cap \Sigma} t^{r_1} u^{r_2}. \quad (19)$$

412 Now, let  $(\Sigma_1, \Sigma_2)$  denote the upper left endpoint of  $\Sigma$  and  $(\Sigma_3, \Sigma_4)$  be its lower right  
 413 endpoint. Clearly, the highest power of  $u$  in the right-hand side of (19) is  $u^{\Sigma_2}$  and the lowest  
 414 one is  $u^{\Sigma_4}$ . Thus the right-hand side of (19) is a polynomial in  $u$  that can be split into  $u^{\Sigma_4}$   
 415 and a polynomial of degree  $\Sigma_2 - \Sigma_4$  which has a nonzero constant term. Hence the number of  
 416 non-trivial solutions equals  $\Sigma_2 - \Sigma_4$ . All those solutions are small roots of the kernel. Moreover,  
 417 note that  $\Sigma_2 - \Sigma_4$  is the height of the segment  $\Sigma$ .

418 Among all monomials of the form  $t^j u^{-e}$  in  $K(t, u)$ , let  $t^\lambda u^{-e}$  be the one where  $\lambda$  is minimal.  
 419 If we draw the Newton polygon of the polynomial  $u^e K(t, u)$ , then it contains the points  $(0, e)$   
 420 and  $(\lambda, 0)$ . Assume that the polygonal line we obtain when traversing the boundary of the  
 421 Newton polygon from  $(0, e)$  to  $(\lambda, 0)$  consists of line segments  $\Sigma_1, \dots, \Sigma_r$  of respective slopes  
 422  $-\beta_1/\alpha_1, \dots, -\beta_r/\alpha_r$ . Then, the line segment  $\Sigma_j$  corresponds to a set of small roots of  $K(t, u)$   
 423 each satisfying  $u(t) \approx t^{\alpha_j/\beta_j}$  as  $t \rightarrow 0$ . The number of such roots is equal to the height of  $\Sigma_j$ ,  
 424 which is the difference of the  $y$ -coordinates of the two endpoints of  $\Sigma_j$ . In particular, this implies  
 425 that the number of small roots of  $K(t, u)$  is equal to  $e$ .

426 Using the same reasoning for the line segments (drawn in red in Figure 3) above  $(0, e)$ , we  
 427 obtain that the number of large roots of the kernel is indeed  $f$ .  $\square$

428 Table 2 on the next page gives several examples of plots of these small and large roots.  
 429 Then, equipped with all these notions, we can give the proofs of our main theorems.



430 Table 2: Several examples that demonstrate the diversity of the behaviour of real branches of  
 431  $K(t, u) = 0$ . In all the examples the set of steps is  $\mathcal{S} = \{-2, -1, 0, 1, 2\}$  (except for the last  
 432 three examples, where  $\mathcal{S}$  is indicated explicitly), and the pattern  $p$  is as indicated. The kernel  
 433  $K$  may also have some complex branches (large or small): they are not shown in the figure, but  
 434 do play a role in our formulas.



## 5 Proofs of the generating functions for walks, bridges, meanders, and excursions

In addition to our usual notations for generating functions of different classes of paths, we denote by  $W_\alpha = W_\alpha(t, u)$  (where  $1 \leq \alpha \leq \ell$ ) the bivariate generating function of those walks avoiding the pattern  $p$  that terminate in state  $\alpha$ ; similarly  $M_\alpha = M_\alpha(t, u)$  for meanders avoiding the pattern  $p$  that terminate in state  $\alpha$ .

### **Proof of Theorem 3.1 (Generating function of walks) and proof of Proposition 4.1.**

On the one hand, we have the following vectorial functional equation:

$$(W_1 \cdots W_\ell) = (1 \ 0 \ \cdots \ 0) + t(W_1 \cdots W_\ell) A,$$

$$(W_1 \cdots W_\ell) (I - tA) = (1 \ 0 \ \cdots \ 0),$$

$$(W_1 \cdots W_\ell) = (1 \ 0 \ \cdots \ 0) \frac{\text{adj}(I - tA)}{\det(I - tA)}.$$

Therefore, the generating function  $W(t, u)$ , which is the sum of the generating functions  $W_\alpha(t, u)$  over all states, is equal to

$$W(t, u) = (W_1 \cdots W_\ell) \vec{\mathbf{1}} = \frac{(1 \ 0 \ \cdots \ 0) \text{adj}(I - tA) \vec{\mathbf{1}}}{\det(I - tA)}.$$

On the other hand, the generating function for  $W(t, u)$  can be obtained using the following combinatorial argument which *en passant* also justifies the introduction of the autocorrelation polynomial, as done in the seminal work of Guibas and Odlyzko [45]. We first introduce  $W^{\{p\}}(t, u)$ , the generating function of the walks over  $\mathcal{S}$  that end with  $p$  and contain no other occurrence of  $p$ . Then we have  $W + W^{\{p\}} = 1 + tPW$  (if we add a letter from  $\mathcal{S}$  to a  $p$ -avoiding walk, then we either obtain another  $p$ -avoiding walk, or a walk with a single occurrence of  $p$  at the end), and  $Wt^\ell u^{\text{alt}(p)} = W^{\{p\}}R$  (a walk obtained from a  $p$ -avoiding walk by appending  $p$  at the end, can also be obtained from a walk ending with a single occurrence of  $p$  at the end by appending the complement of a presuffix of  $p$ ). Solving this system, we obtain  $W(t, u) = R(t, u)/K(t, u)$ .

Thus, we got two representations for  $W(t, u)$ :

$$W(t, u) = \frac{(1 \ 0 \ \cdots \ 0) \text{adj}(I - tA) \vec{\mathbf{1}}}{\det(I - tA)} = \frac{R(t, u)}{(1 - tP(u))R(t, u) + t^{\ell|u|^{\text{alt}(p)}}}. \quad (20)$$

In order to see that this is the same representation (that is, the numerators and the denominators are equal in both fractions), we notice that  $\det(I - tA)$  is a polynomial in  $t$  of degree  $\ell$  and constant term 1. This is also the case for  $(1 - tP(u))R(t, u) + t^{\ell|u|^{\text{alt}(p)}}$ , so this allows us to say that the two numerators in Formula (20) are actually equal. This gives the proof of Theorem 3.1 for walks, and also proves Proposition 4.1 on the structure of the kernel.  $\square$

We now turn to the consequences of this formula for  $W(t, u)$  when one considers bridges.

**Proof of Theorem 3.1 (Generating function of bridges).**

In order to find the univariate generating function  $B(t)$  for bridges, we need to extract the coefficient of  $[u^0]$  from  $W(t, u)$ . To this end, we assume that  $t$  is a sufficiently small fixed number, extract the coefficient of a (univariate) function by means of Cauchy's integral formula, and apply the residue theorem (recall that  $u_1, \dots, u_e$  are the small roots of  $K(t, u)$ ):

$$B(t) = [u^0]W(t, u) = \frac{1}{2\pi i} \int_{|u|=\varepsilon} \frac{W(t, u)}{u} du = \sum_i^e \operatorname{Res}_{u=u_i(t)} \frac{W(t, u)}{u}.$$

By the formula for residues of rational functions, we have

$$\begin{aligned} \operatorname{Res}_{u=u_i(t)} \frac{W(t, u)}{u} &= \operatorname{Res}_{u=u_i(t)} \frac{R(t, u)}{u((1-tP(u))R(t, u) + t^\ell u^{\operatorname{alt}(p)})} \\ &= \frac{R(t, u)}{\frac{d}{du}(u((1-tP(u))R(t, u) + t^\ell u^{\operatorname{alt}(p)}))} \Big|_{u=u_i(t)}. \end{aligned}$$

The denominator of this expression is

$$-tuP'(u)R(t, u) + u(1-tP(u))R_u(t, u) + \operatorname{alt}(p)t^\ell u^{\operatorname{alt}(p)} \Big|_{u=u_i(t)}. \quad (21)$$

Next, we differentiate  $K(t, u_i) = 0$  with respect to  $t$  and obtain an expression for  $P'(u_i(t))$ . When we substitute it into (21), we obtain (7).  $\square$

We now consider the nonnegativity constraint for the paths.

**Proof of Theorem 3.2 (Generating function of meanders).**

We have the following vectorial functional equation:

$$(M_1 \ \dots \ M_\ell) = (1 \ 0 \ \dots \ 0) + t (M_1 \ \dots \ M_\ell) A - t \{u^{<0}\}((M_1 \ \dots \ M_\ell) A),$$

or, equivalently,

$$(M_1 \ \dots \ M_\ell)(I - tA) = (1 \ 0 \ \dots \ 0) - t \{u^{<0}\}((M_1 \ \dots \ M_\ell) A), \quad (22)$$

where  $\{u^{<0}\}$  denotes all the terms in which the power of  $u$  is negative.

The right-hand side of (22) is a vector whose components are power series in  $t$  and Laurent polynomials in  $u$  (of lowest degree  $\geq -c$ ). For  $\alpha = 1, \dots, \ell$ , denote the  $\alpha$ -th component of this vector by  $F_\alpha = F_\alpha(t, u)$  (the letter  $F$  can be seen as a mnemonic for "forbidden", as these components correspond to the forbidden transitions towards a negative value as exponent of  $u$ ).

In summary, one has

$$(M_1 \ \dots \ M_\ell)(I - tA) = (F_1 \ \dots \ F_\ell). \quad (23)$$

We multiply this from the right by  $(I - tA)^{-1} \vec{\mathbf{1}} = \frac{(\operatorname{adj}(I - tA)) \vec{\mathbf{1}}}{\det(I - tA)}$ . At this point, we denote  $\vec{\mathbf{v}} = \vec{\mathbf{v}}(t, u) := (\operatorname{adj}(I - tA)) \vec{\mathbf{1}}$ , where  $\vec{\mathbf{1}}$  is the column vector  $(1 \ 1 \ \dots \ 1)^\top$ . This vector  $\vec{\mathbf{v}}$  is the autocorrelation vector we encountered in Proposition 4.2. As a direct consequence of its definition, one has

$$M(t, u) = \frac{(F_1 \ \dots \ F_\ell) \vec{\mathbf{v}}}{K(t, u)}. \quad (24)$$

495 The following step is the essential part of the vectorial kernel method. Let  $u_i = u_i(t)$  be any  
 496 small root of  $K(t, u) = \det(I - tA)$ . We plug in  $u = u_i(t)$  into (23). The matrix  $(I - tA)|_{u=u_i}$   
 497 is then singular. At this point we observe that  $\vec{v}|_{u=u_i}$  is an eigenvector of  $(I - tA)|_{u=u_i}$   
 498 belonging to the eigenvalue  $\lambda = 0$ . Indeed,  $\vec{v}|_{u=u_i} = (\text{adj}((I - tA)|_{u=u_i}))\vec{\mathbf{1}}$  is equivalent to  
 499  $(I - tA)|_{u=u_i}\vec{v}|_{u=u_i} = \det((I - tA)|_{u=u_i})\vec{\mathbf{1}}$ , which implies  $(I - tA)|_{u=u_i}\vec{v}|_{u=u_i} = 0$ . Moreover,  
 500 due to the structure of  $A$ , we have  $\text{rank}((I - tA)|_{u=u_i}) = \ell - 1$ , therefore, the dimension of  
 501 the characteristic space of  $\lambda = 0$  is 1, and  $\vec{v}|_{u=u_i}$  is the *unique* (up to scaling) eigenvector  
 502 of  $(I - tA)|_{u=u_i}$  that belongs to  $\lambda = 0$ .

503 Thus, if we multiply (23) by  $\vec{v}|_{u=u_i}$ , the left-hand side vanishes. In other words, the equation  
 504  $(F_1(t, u), \dots, F_\ell(t, u))\vec{v}(t, u) = 0$  is satisfied by every small root  $u_i(t)$  of  $K(t, u)$ .

505 Let

$$506 \quad \Phi(t, u) := u^e (F_1(t, u), \dots, F_\ell(t, u))\vec{v}(t, u). \quad (25)$$

507 Note that  $\Phi$  is a Laurent polynomial, as the  $F_i$ 's and  $\vec{v}$  are by construction Laurent polynomials  
 508 in  $u$ . What is more, since  $\Phi(t, u) = u^e M(t, u)K(t, u)$  by (24) and since  $M(t, u)$  is a power  
 509 series in  $u$ ,  $\Phi(t, u)$  has no negative powers of  $u$  and is thus a polynomial. Now, we know that  
 510 every small root  $u_i(t)$  of  $K(t, u)$  is a root of a *polynomial* equation

$$511 \quad \Phi(t, u) = 0. \quad (26)$$

512 It follows that

$$513 \quad \Phi(t, u) = G(t, u) \prod_{i=1}^e (u - u_i(t)) \quad (27)$$

514 for some  $G(t, u)$  which is a formal power series in  $t$  and a polynomial in  $u$ . We substitute this  
 515 into (24), and obtain the claimed formula

$$516 \quad M(t, u) = \frac{G(t, u)}{u^e K(t, u)} \prod_{i=1}^e (u - u_i(t)). \quad \square$$

517 If the degree of  $\Phi(t, u)$  is precisely  $e$ , the formula simplifies as  $G$  is then just the leading  
 518 term (in  $u$ ) of  $\Phi(t, u)$ . As we shall show now, this happens if  $p$  is a quasimeander (as introduced  
 519 in Definition 3.3).

520 ***Proof of Theorem 3.4 (Generating function of meanders, when  $p$  is a quasimeander).***

521 First, we notice that all the powers of  $u$  in  $R(t, u)$  are non-negative, and  $\text{alt}(p) \geq -c$ . Moreover,  
 522 if  $\text{alt}(p) = -c$ , the cancellation of terms with  $u^{-c}$  in  $K(t, u)$  is not possible<sup>7</sup>. Therefore, the  
 523 lowest power of  $u$  in  $K(t, u)$  is  $c$ , and thus we have  $e = c$  by Proposition 4.4.

524 Let us return to (22):

$$525 \quad (M_1 \cdots M_\ell)(I - tA) = (1 \ 0 \ \cdots \ 0) - t \{u^{<0}\}((M_1 \cdots M_\ell)A). \quad (28)$$

<sup>7</sup>The only exception is the case of  $p = [-c]$ . The case of patterns of length 1 is not interesting, so we assume from now on that  $\ell \geq 2$ . Yet, one can check that our formula is also valid for  $\ell = 1$ , with some adjustments in the special case  $p = [-c]$ .

526 We claim that all the components of the right-hand side, except for the first component,  
 527 are 0. Indeed, if the path arrives at state  $X_i$  with  $i > 1$ , this means that it accumulated a  
 528 non-empty prefix of  $p$ . And since  $p$  is a quasimeander,  $w$  will always remain (weakly) above the  
 529  $x$ -axis while it accumulates its non-empty prefix.

530 Therefore, we have  $\Phi = u^c (F_1 \ 0 \ \cdots \ 0) \vec{v}$ , and thus by (15),  $\Phi = u^c F_1 R$ . Since the  
 531 constant term of  $F_1$  is 1,  $u^c F_1$  is a monic polynomial in  $u$ . Therefore, we have

$$\Phi = R \cdot \prod_{i=1}^c (u - u_i).$$

532 This yields

$$533 \quad M(t, u) = \frac{R(t, u)}{u^c K(t, u)} \prod_{i=1}^c (u - u_i(t)) \quad (29)$$

534 as claimed. □

535 Let us now simplify this formula for excursions, i.e., when  $u = 0$ .

536 ***Proof of Theorem 3.4 (Generating function of excursions, when  $p$  is a quasimeander).***

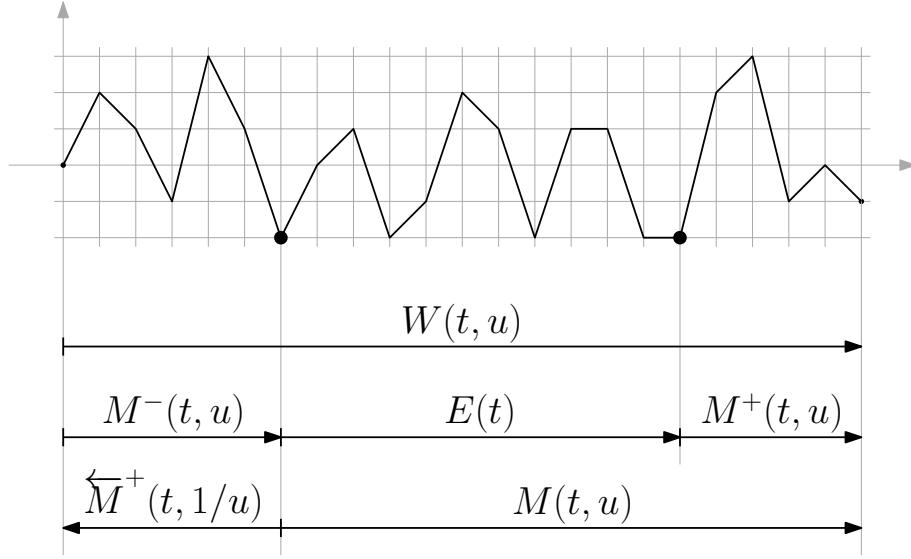
537 The generating function of excursions is given by  $E(t) = M(t, 0)$ . If  $\text{alt}(p) > -c$ , we have, as  
 538  $u$  tends to 0,  $K(t, u) \sim -tu^{-c}R(t, 0)$  from (16). If  $\text{alt}(p) = -c$ , then we have  $R(t, u) = 1$  and  
 539  $K(t, u) \sim -tu^{-c} + t^\ell u^{-c}$ . In both cases, (11) follows. □

540 We now handle the next interesting class of patterns leading to generating functions with a  
 541 nice closed form: the case of reversed meanders. Recall that a reversed meander is a lattice  
 542 path whose terminal point has a (strictly) smaller  $y$ -coordinate than all other points. Moreover,  
 543 we define a *positive meander* to be a meander that never returns to the  $x$ -axis.

544 ***Proof of Theorem 3.5 (Gen. function of meanders, when  $p$  is a reversed meander).***

545 If  $p$  is a reversed meander, then all the terms of  $R(t, u)$ , except for the monomial 1, contain  
 546 negative powers of  $u$ , and we also have  $\text{alt}(p) < 0$ . Therefore, the highest power of  $u$  in  $K$   
 547 is  $d$ , and (by Proposition 4.4) the number of large roots of  $K(t, u) = 0$  is  $d$ : we denote them as  
 548 above by  $v_1, \dots, v_d$ . The number of small roots can be in general higher than  $c$ : as usual, we  
 549 denote it by  $e$ , and the roots themselves by  $u_1, \dots, u_e$ .

550 We consider the following generalization of the *Wiener–Hopf factorization* for lattice paths.  
 551 If we split the walk  $w$  at its first and at its last left-to-right minimum, we obtain a decomposition  
 552  $w = m^- . e . m^+$ , where  $m^-$  is a reversed meander,  $e$  is a translate of an excursion, and  $m^+$  is a  
 553 translate of a positive meander. One also has the decomposition  $w = m^- . m$ , where  $m = e . m^+$   
 554 is a meander. Notice that these decompositions are unique.



555 Figure 4: The Wiener–Hopf factorization of a walk:  $W = M^-EM^+$ , a product of a reversed-  
 556 meander, an excursion, and a positive meander. See e.g. [44] for the importance of this  
 557 factorization for lattice path enumeration. This has further consequences on pattern avoidance  
 558 when one reverses the time.

559 Moreover, since  $p$  is a reversed meander, its occurrence cannot overlap the junction of two  
 560 factors. That is,  $m$  is  $p$ -avoiding if and only if its both factors are  $p$ -avoiding, and  $w$  is  $p$ -avoiding  
 561 if and only if its three factors are  $p$ -avoiding. Therefore, we have

$$562 \quad M(t, u) = E(t)M^+(t, u), \quad (30)$$

563 and

$$564 \quad W(t, u) = M^-(t, u)M(t, u) = M^-(t, u)E(t)M^+(t, u),$$

565 where  $W(t, u)$ ,  $M^-(t, u)$ ,  $E(t)$ ,  $M^+(t, u)$  are the generating functions of  $p$ -avoiding walks,  
 566 reversed meanders, excursions, positive meanders (respectively). This in particular implies

$$567 \quad M(t, u) = \frac{W(t, u)}{M^-(t, u)}. \quad (31)$$

568 By Theorem 3.1, we have  $W(t, u) = R(t, u)/K(t, u)$ . In order to find  $M^-(t, u)$ , we use a  
 569 *time reversal* argument. Namely, we notice that a path is a reversed meander if and only if its  
 570 horizontal reflection (upon translating the initial point to the origin) is a positive meander. The  
 571 precise statement is as follows. Let  $-\mathcal{S} = \{-s : s \in \mathcal{S}\}$ ; and for the pattern  $p = [a_1, a_2, \dots, a_\ell]$ ,  
 572 let  $\overleftarrow{p} = [-a_\ell, \dots, -a_2, -a_1]$ . Then there is a straightforward bijection between  $p$ -avoiding  
 573 reversed meanders with steps from  $\mathcal{S}$  and  $\overleftarrow{p}$ -avoiding positive meanders with steps from  $-\mathcal{S}$   
 574 which preserves the length and reflects the altitude. Therefore, we have

$$575 \quad M^-(t, u) = \overleftarrow{M}^+(t, 1/u), \quad (32)$$

576 where the arrow means that it is the generating function for  $\overleftarrow{p}$ -avoiding paths (positive meanders  
577 in this equation) with the step set  $-\mathcal{S}$  (rather than  $p$ -avoiding with the step set  $\mathcal{S}$ ).

578 Refer to the  $m = e.m^+$  decomposition above. As we noticed, if the pattern  $p$  is a *reversed*  
579 meander, then  $m$  is  $p$ -avoiding if and only if both  $e$  and  $m^+$  are  $p$ -avoiding. The same is true  
580 if  $p$  is a *positive* meander. Therefore, similarly to Formula (30),  $M(t, u) = E(t)M^+(t, u)$ , we  
581 also have  $\overleftarrow{M}(t, u) = \overleftarrow{E}(t)\overleftarrow{M}^+(t, u)$ . Combined with (32), this implies

$$582 \quad M^-(t, u) = \overleftarrow{M}^+(t, 1/u) = \frac{\overleftarrow{M}(t, 1/u)}{\overleftarrow{E}(t)}. \quad (33)$$

583 Since  $\overleftarrow{p}$  is a meander, (9) holds for  $\overleftarrow{M}(t, u)$  and we have

$$584 \quad \overleftarrow{M}(t, 1/u) = \frac{u^d \overleftarrow{R}(t, 1/u)}{\overleftarrow{K}(t, 1/u)} \prod_{j=1}^d \left( \frac{1}{u} - \frac{1}{v_j(t)} \right) = \frac{u^d R(t, u)}{K(t, u)} \prod_{j=1}^d \left( \frac{1}{u} - \frac{1}{v_j(t)} \right). \quad (34)$$

585 These identities are justified as follows. The equalities  $\overleftarrow{R}(t, 1/u) = R(t, u)$  and  $\overleftarrow{K}(t, 1/u) =$   
586  $K(t, u)$  can be easily derived directly, but also notice that we have  $W(t, u) = R(t, u)/K(t, u)$   
587 and  $\overleftarrow{W}(t, 1/u) = \overleftarrow{R}(t, 1/u)/\overleftarrow{K}(t, 1/u)$ , and  $W(t, u) = \overleftarrow{W}(t, 1/u)$  from the bijective horizontal  
588 reflection. Finally,  $\overleftarrow{K}(t, u)$  has  $d$  many small roots and  $e$  many large roots: if  $u_i(t)$  is a small  
589 root of  $K(t, u)$ , then  $1/u_i(t)$  is a large root of  $\overleftarrow{K}(t, u)$ ; and if  $v_j(t)$  is a large root of  $K(t, u)$ ,  
590 then  $1/v_j(t)$  is a small root of  $\overleftarrow{K}(t, u)$ .

591 Similarly, Equation (11) holds for  $\overleftarrow{E}(t)$ , and we have

$$592 \quad \overleftarrow{E}(t) = \frac{(-1)^{d+1}}{t} \prod_{j=1}^d \frac{1}{v_j(t)}. \quad (35)$$

593 Notice that the leading term of the polynomial  $u^e K(t, u)$  is  $-tu^{d+e}$  and, therefore, one has

$$594 \quad u^e K(t, u) = -t \prod_{i=1}^e (u - u_i(t)) \prod_{j=1}^d (u - v_j(t)). \quad (36)$$

595 We now substitute (34) and (35) into (33) and use (36) to obtain

$$596 \quad M^-(t, u) = \frac{(-1)^{d+1} t u^d R(t, u)}{K(t, u)} \prod_{j=1}^d \left( v_j \left( \frac{1}{u} - \frac{1}{v_j(t)} \right) \right) \quad (37)$$

$$597 \quad = \frac{-t R(t, u)}{K(t, u)} \prod_{j=1}^d (u - v_j(t)) = \frac{u^e R(t, u)}{\prod_{i=1}^e (u - u_i(t))}. \quad (38)$$

598 Finally, we substitute this into (31) and obtain

$$599 \quad M(t, u) = \frac{W(t, u)}{M^-(t, u)} = \frac{R(t, u)}{K(t, u)} \frac{1}{u^e R(t, u)} \prod_{i=1}^e (u - u_i(t)) = \frac{1}{u^e K(t, u)} \prod_{i=1}^e (u - u_i(t)). \quad (39)$$

600 □

601 **Remark 5.1.** It is interesting to notice that though  $M(t, u)$ , for  $p$  being a quasimeander (as  
 602 given in (29)), is similar to  $M(t, u)$ , for  $p$  being a reversed meander (as given in (39)), the latter  
 603 does not contain the factor  $R(t, u)$  even if  $p$  has a non-trivial autocorrelation.

604 **Remark 5.2.** It is also worth mentioning that if only the terminal point of the pattern  $p$  has  
 605 negative  $y$ -coordinate, then  $p$  is both a quasimeander and a reversed meander, and  $R = 1$ .  
 606 Therefore, we have  $M(t, u) = \frac{1}{u^e K(t, u)} \prod_{i=1}^e (u - u_i(t))$  by both Theorem 3.4 and Theorem 3.5.

607 **Proof of Theorem 3.5 (Gen. function of excursions, when  $p$  is a reversed meander).**  
 608 Excursions are given by  $M(t, 0)$ , so we need to compute  $D(t) := [u^0]u^e K(t, u)$ . To this aim,  
 609 first note that as  $p$  is a reversed meander (see Definition 3.3), one has the following facts.

- 610 • In all the terms of  $R(t, u)$ , the powers of  $u$  are non-positive.
- 611 • Moreover, if  $t^{m_1}/u^{\gamma_1}$  and  $t^{m_2}/u^{\gamma_2}$  are two distinct terms in  $R(t, u)$  such that  $0 \leq m_1 < m_2$ ,  
 612 then we have  $0 \leq \gamma_1 < \gamma_2$ .

613 Therefore, we can order the terms of  $R(t, u)$  according to the powers of  $t$ , and write  $u^e K(t, u)$   
 614 as follows:

$$u^e K(t, u) = u^e \left( \left( 1 - t \left( \frac{1}{u^c} + \dots + u^d \right) \right) \left( 1 + \dots + \frac{t^{m'}}{u^{\gamma'}} + \frac{t^m}{u^\gamma} \right) + t^\ell u^{\text{alt}(p)} \right),$$

615 where  $\frac{t^m}{u^\gamma}$  (the last term in  $R(t, u)$ ) corresponds to the longest complement of a presuffix. Now,  
 616 we have the following cases:

- 617 • Case 1:  $c + \gamma > -\text{alt}(p)$ . Then  $e = c + \gamma$  and we have  $D(t) = -t^{m+1}$ .
- 618 • Case 2:  $c + \gamma < -\text{alt}(p)$ . Then  $e = -\text{alt}(p)$  and we have  $D(t) = t^\ell$ .
- 619 • Case 3:  $c + \gamma = -\text{alt}(p)$  and  $\ell \neq m + 1$ . Then  $e = c + \gamma = -\text{alt}(p)$  and  $D(t) = t^\ell - t^{m+1}$ .
- 620 • Case 4:  $c + \gamma = -\text{alt}(p)$  and  $\ell = m + 1$ . If  $\ell \geq 2$ , then  $m \geq 1$ , and therefore  $R(t, u) \neq 1$ .  
 621 Then  $e = c + \gamma'$  and  $D(t) = -t^{m'+1}$ . As usual, we ignore the degenerate case  $\ell = 1$ .

622 In summary, we get the claim we wanted to prove, namely

$$623 \quad E(t) = M(t, 0) = \frac{(-1)^e}{D(t)} \prod_{i=1}^e u_i(t), \quad (40)$$

624 where  $D(t)$  is either some power of  $t$ , or a difference of two powers of  $t$ . □

625 Now that we have proven these closed forms for the generating functions, we can turn to  
 626 the asymptotics of their coefficients.

## 6 Asymptotics of lattice paths avoiding a given pattern

The aim of this section is to characterize the asymptotics of the number of paths (walks, bridges, meanders, excursions) with steps from  $\mathcal{S}$  avoiding a given pattern  $p$ .

In order to avoid pathological cases, we now focus on “generic” walks.

**Definition 6.1** (Generic walks). We call a constrained walk model *generic* if the following five properties hold:

- Property 1. The generating functions  $B(t)$ ,  $M(t)$  and  $E(t)$  are algebraic, not rational.
- Property 2. They have a unique dominant singularity, which is algebraic, not a pole.
- Property 3. The factor  $G(t, u)$  in Equation (8) is a polynomial in  $t$ .
- Property 4. Let  $\rho$  be the smallest positive real number such that a large branch meets a small branch at  $t = \rho$ . No large negative branch (i.e., a branch of  $K(t, u) = 0$  such that  $\lim_{t \rightarrow 0^+} u(t) = -\infty$ ) meets a small negative branch at  $t = \rho$ .
- Property 5. The smallest positive root of  $K(t, 1)$  is simple.

These properties are natural and it is easy to analyse the subcases for which they are not holding.

- For Property 1, it can be the case that the forbidden pattern leads to a degenerate model, in the sense that it is no more involving any stack. Thus, we have words generated by a *regular* automaton (hence, the generating functions are rational and the asymptotics are well understood). Example:  $\mathcal{S} = \{-1, 1\}$  with  $p = [1, -1]$  or  $p = [-1, -1]$ .
- For Property 2, it is proven in [7] that several dominant singularities appear if and only if the gcd of the pairwise differences of the steps is not 1. In this case, the asymptotics are obtained via [14, Theorem 8.8]. Moreover, polar singularities are possible, but these are easy to handle.
- For Property 3, it is satisfied in many natural cases (like e.g. in Theorems 3.4 and 3.5) and we analyse in the follow-up article [3] what happens otherwise.
- For Property 4, we conjecture that it always holds. In fact, we have a proof for many classes of walks, but some remaining cases are open. Note that it is possible to exhibit cases where one small negative root meets a large negative root, at some  $\rho' > \rho$ : this is e.g. the case for  $\mathcal{S} = \{-2, -1, 0, 1, 2\}$  with  $p = [0, 1, -2]$ . Moreover, it is also possible that two small negative roots meet at  $\rho$ : e.g. for  $\mathcal{S} = \{-2, 1\}$  with  $p = [1, -2, 1, -2]$ .
- For Property 5, an example of a double root for  $K$  is given by  $\mathcal{S} = \{-1, 1\}$  and  $p = [-1, 1]$  (this example corresponds to the very last drawing in Table 2). Double or higher multiplicity roots would just create additional subcases (trivial to handle) in the following theorems.

We observe that the behaviour of real branches of  $K(t, u) = 0$  is much more complicated and diverse than that in the Banderier–Flajolet study. To recall, in their case there are always two real *positive* branches (one small branch  $u_1$  and one large branch  $v_1$ ) that meet at a singularity point  $(t, u) = (\rho, \tau)$ , where  $u = \tau$  is the only positive number such that  $P'(\tau) = 0$ . In contrast, in our case we may have additional positive branches – even when the autocorrelation is trivial.



664 Table 2 from Section 4 illustrates that we always have a small branch and one large branch  
 665 whose shape in general resembles that of  $u_1 \cup v_1$  observed for classical paths by Banderier and  
 666 Flajolet. In one sense, the geometry of the branches of  $K$  observed for classical paths is now  
 667 perturbed by the pattern avoidance constraint: this perturbation adds new branches. In the  
 668 next section, we introduce the generating function  $W(t, u, v)$  where  $v$  encodes the number of  
 669 occurrences of the pattern  $p$ . One can then play with  $v$  like if it would encode a Boltzmann  
 670 weight/Gibbs measure (a typical point of view in statistical mechanics): moving the parameter  $v$   
 671 in a continuous way from 1 to 0 gives a rigorous explanation of this perturbation phenomenon,  
 672 and shows the coherence with the emergence of new branches. More information about these  
 673 branches (and their Puiseux expansions) can be derived from the Newton polygon associated  
 674 with the kernel (see Proposition 4.4 and [34]).

675 **Lemma 6.2** (Location and nature of the dominant singularity). *For any generic model, the*  
 676 *dominant singularity of  $B(t)$  and  $E(t)$  is  $\rho$ , the smallest real positive number such that a small*  
 677 *branch meets a large branch at  $t = \rho$ . (The branches refer to the roots of  $K(t, u) = 0$ , as*  
 678 *defined in (4)). We call these branches  $u_1$  and  $v_1$ . Additionally, their branching point is a square*  
 679 *root singularity.*

680 *Proof.* Lattice paths avoiding a given pattern can be generated by a pushdown automaton (see  
 681 Figure 1). Accordingly, they can be generated by a context-free grammar, and their generating  
 682 functions thus satisfy a “positive” system of algebraic equations (see [27]). Therefore, the  
 683 asymptotic number of words of length  $n$  in such languages is of the form  $C\rho^{-n}n^\alpha$ . When the  
 684 system is not strongly connected,  $\alpha$  is either an integer (if  $\rho$  is a pole), either a dyadic number  
 685 (if one has an iterated square root Puiseux singularity at  $\rho$ ), as proven by Banderier and Drmota  
 686 in [8]. For excursions, one has a strongly connected dependency graph (see Figure 1); the  
 687 dominant singularity  $\rho$  (or, possibly, the dominant singularities) thus behaves like a square root,  
 688 as we have generic walks (and not a degenerate case where we face a polar singularity).

689 Now, for generic walks, because of the product formula (8) for excursions, one (or several)  
 690 of the small roots have to follow this square root Puiseux behaviour. By Pringsheim’s theorem,  
 691 this has to be at a place  $0 < \rho \leq 1$ . Note that the Pólya–Fatou–Carlson theorem [26] on pure  
 692 algebraic functions with integer coefficients says that they cannot have radius of convergence 1.  
 693 Therefore, the first crossing between a small and large branch is at  $0 < \rho < 1$  (i.e.,  $\rho = 1$  or  
 694 any other root of  $t - t^\ell$ , cannot be the dominant singularity). Now, by Proposition 4.1, the  
 695 geometry of the branches implies that this branching point is at a location where a large branch  
 696 meets a small branch, because if the branching point comes from the intersection of small roots  
 697 only (see the examples 1, 6, 8 in Table 2 for such a case), then their product will be regular. So,  
 698  $\rho$  has to be the smallest real positive number where a small branch meets a large branch.

699 When one does not take into account occurrences of a pattern, the generating function of  
 700 bridges is essentially the logarithmic derivative of the generating function of excursions, and  
 701 they have the same radius of convergence (the cycle lemma, the identity  $B = 1 + Et\partial_t A$ ,  
 702 Spitzer’s and Sparre Andersen’s formulas are alter egos of this relation, see the paragraph “On  
 703 the relation between bridges and excursions” in [9, Theorem 5]). For walks with a forbidden  
 704 pattern, this simple relation is not holding anymore and there is no apparent equation linking  
 705 the two generating functions. Nevertheless, the numbers  $e_n$  of excursions and  $b_n$  of bridges  
 706 of length  $n$  still satisfy  $e_n \leq b_n \leq ne_n$  (this is easily seen by doing the  $n$  cyclic shifts of each  
 707 excursion). This implies that  $E(t)$  and  $B(t)$  have the same radius of convergence.  $\square$

708 Equipped with this additional information on the roots and the way they cross, we can  
 709 derive the following asymptotic results. Note that we use the notations  $K_t(t, u)$  for  $(\partial_t K)(t, u)$ ,  
 710 and  $K_{uu}(t, u)$  for  $(\partial_u^2 K)(t, u)$ . We start with the asymptotics of walks on  $\mathbb{Z}$  with a forbidden  
 711 pattern.

712 **Theorem 6.3** (Asymptotics of walks on  $\mathbb{Z}$ ). *Let  $\rho_K$  be the smallest positive root of  $K(t, 1)$ .  
 713 For any generic model, the asymptotic number of walks of length  $n$  avoiding a pattern  $p$  is*

$$W_n \sim -\rho_K K_t(\rho_K, 1) R(\rho_K, 1) \rho_K^{-n}.$$

714 *Proof.* This follows from the partial fraction decomposition of  $W(t) = \frac{R(t, 1)}{K(t, 1)}$ , where  $\rho_K$  is a  
 715 simple pole as the model is generic.  $\square$

716 Now, for excursions and bridges, the corresponding generating functions have an algebraic  
 717 dominant singularity; this leads to the following theorems.

718 **Theorem 6.4** (Asymptotics of excursions). *For any generic model, the asymptotic number of  
 719 excursions of length  $n$  avoiding a pattern  $p$  is*

$$E_n \sim (-1)^{e-1} \frac{Y(\rho)G(\rho, 0)}{D(\rho)} \sqrt{\frac{K_t(\rho, \tau)}{2\pi\rho K_{uu}(\rho, \tau)}} \cdot n^{-3/2} \rho^{-n},$$

720 where  $\tau := u_1(\rho)$ ,  $Y(t) := u_2(t) \cdots u_e(t)$ , and  $D(t) := [u^0]u^e K(t, u)$ .

721 *Proof.* We use the closed form given in Theorem 3.2. Since the model is generic, the product  
 722  $\frac{G(t, 0)}{D(t)} Y(t)$  is analytic for  $|t| \leq \rho$ . (Caveat: it can be the case that some small branches are not  
 723 analytic for some  $|t| < \rho$ , however, their product is then analytic.) Now, for any generic pattern,  
 724  $D(t)$  is either a monomial or of the shape given in Case 3, page 23, but, as  $\rho < 1$  (as shown in  
 725 the course of the proof of Lemma 6.2), one thus has  $D(\rho) \neq 0$ . So, the singularity and the  
 726 local behaviour of  $E(t)$  is completely determined by the singular behaviour of  $u_1(t)$ . This local  
 727 expansion of  $u_1$  is given by a local inversion of  $K(t, u)$  at  $(t, u) = (\rho, \tau)$ ; this leads to

$$728 \quad u_1(t) \sim \tau - \sqrt{\frac{2K_t(\rho, \tau)}{\rho K_{uu}(\rho, \tau)}} \sqrt{1 - \frac{t}{\rho}}, \quad \text{as } t \rightarrow \rho. \quad (41)$$

729 The claim is then reached by singularity analysis (see [42]) on the Puiseux expansion

$$730 \quad E(t) \sim E(\rho) - (-1)^{e-1} \frac{Y(\rho)G(\rho, 0)}{D(\rho)} \sqrt{\frac{2K_t(\rho, \tau)}{\rho K_{uu}(\rho, \tau)}} \sqrt{1 - \frac{t}{\rho}}. \quad \square$$

731 **Theorem 6.5** (Asymptotics of bridges). For any generic model, the asymptotic number of  
 732 bridges of length  $n$  avoiding a pattern  $p$  is

$$B_n \sim -\frac{R(\rho, \tau)}{\tau K_t(\rho, \tau)} \sqrt{\frac{K_t(\rho, 1)}{2\pi\rho K_{uu}(\rho, 1)}} \cdot n^{-1/2} \rho^{-n}.$$

733 *Proof.* We know from Lemma 6.2 that  $B(t)$  and  $E(t)$  have the same radius of convergence,  
 734 where  $B(t)$  is given by Equation (7) from Theorem 3.1. Thus, the singular behaviour of  $u_1(t)$   
 735 determines the singularity and the local behaviour of  $B(t)$ . We have therefore

$$B(t) \sim -\frac{R(t, u_1(t))}{K_t(t, u_1(t))} \frac{u_1'(t)}{u_1(t)} \quad \text{for } t \sim \rho,$$

736 with a denominator  $K_t$  which is not 0 for  $t = \rho$ . So, plugging the singular expansion of  $u_1$  into  
 737 this formula yields the result.  $\square$

738 We now introduce the notion of drift, which plays a role for the asymptotics of meanders.

739 **Definition 6.6** (Drift of a walk). For any given set of steps  $\mathcal{S}$  and forbidden pattern  $p$ , the  
 740 drift is the quantity

$$741 \quad \delta := \lim_{n \rightarrow \infty} \frac{\text{average final altitude of walks on } \mathbb{Z} \text{ of length } n}{n}.$$

742 Thus  $\delta > 0$ ,  $\delta < 0$ , or  $\delta = 0$  correspond to the fact that almost all the walks of length  $n$  on  $\mathbb{Z}$   
 743 have a final altitude of order which is either  $+\Theta(n)$ ,  $-\Theta(n)$ , or  $o(n)$ , respectively. One says  
 744 that these walks (and the corresponding meanders/excursions/bridges) have a positive, negative  
 745 or zero drift, respectively.

746 As usual, the drift is not playing a role for the asymptotics of excursions and bridges. Indeed,  
 747 the constraint to force the walk to end at altitude zero is there "killing" the drift. This is best  
 748 seen by a "time reversal" argument: under this transformation, a bridge stays a bridge (which  
 749 then avoids the reverse forbidden pattern), one thus gets the same generating function (note  
 750 that  $K(t, u)$  then becomes  $K(t, 1/u)$ ); therefore, the asymptotics *have to* be independent of  
 751 the drift. A similar reasoning holds for excursions. Now, the next theorem shows how the drift  
 752 does play a role for the asymptotics of meanders. For meanders with negative or zero drift,  
 753 the quantity  $\rho$  from Lemma 6.2 is also the radius of convergence of  $M(t)$ . For meanders with  
 754 positive drift, the radius of convergence of  $M(t)$  is the dominant pole of  $1/K(t, 1)$ .

755 **Theorem 6.7** (Asymptotics of meanders). *Assume that the model is generic. Let  $\rho$ ,  $\rho_K$ , and  $\tau$*   
 756 *be defined like in the previous theorems. We have one of the following three cases:*

- 757 • *If  $\tau = 1$  and  $\rho_K = \rho$ , then we are in the “zero drift” case.*
- 758 • *If  $\tau > 1$  and  $u_1(\rho_K) = 1$  and all large roots  $v$  satisfy  $v(t) \neq 1$  for  $\rho_K < t < \rho$ , then we*  
 759 *are in the “negative drift” case.*
- 760 • *If either  $\tau < 1$ , or  $\tau = 1$  but  $\rho_K < \rho$ , or  $\tau > 1$  but some large root  $v$  satisfies  $v(\rho_K) = 1$ ,*  
 761 *then we are in the “positive drift” case.*

762 *Then the asymptotics of the coefficients of the meander generating function*

$$763 \quad M(t) = \frac{(1 - u_1(t))Y(t)G(t, 1)}{K(t, 1)} \quad \text{with } Y(t) := \prod_{i=2}^c (1 - u_i(t))$$

764 *is given by*

$$765 \quad M_n \sim G(\rho, 1)Y(\rho) \sqrt{\frac{2}{\pi \rho K_t(\rho, 1) K_{uu}(\rho, 1)}} \cdot n^{-1/2} \rho^{-n} \quad (\text{“zero drift”}),$$

$$766 \quad M_n \sim -\frac{G(\rho, 1)Y(\rho)}{K(\rho, 1)} \sqrt{\frac{\rho K_t(\rho, \tau)}{2\pi K_{uu}(\rho, \tau)}} \cdot n^{-3/2} \rho^{-n} \quad (\text{“negative drift”}),$$

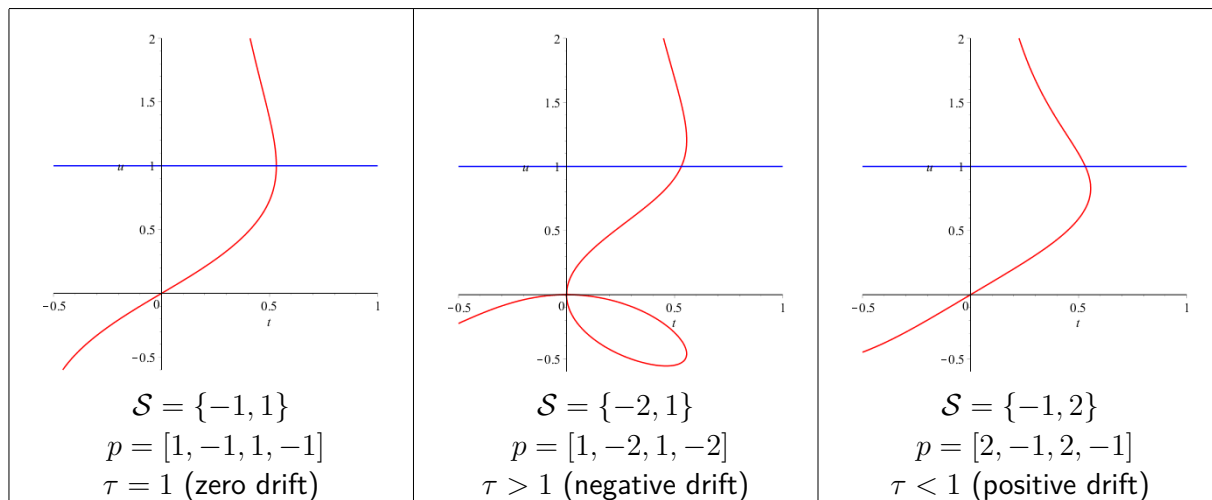
$$767 \quad M_n \sim -\frac{(1 - u_1(\rho_K))Y(\rho_K)G(\rho_K, 1)}{\rho_K K_t(\rho_K, 1)} \cdot \rho_K^{-n} \quad (\text{“positive drift”}).$$

768 *Proof.* Before we begin the case analysis, let us mention some preliminary facts. The line  $u = 1$   
 769 intersects the curve shaped by  $u_1 \cup v_1$  at some point  $(t_0, 1)$  (see Figure 5). Notice that the  
 770 right-most point of this curve is  $(\rho, \tau)$ . Thus  $t_0 \leq \rho$ , where equality holds if and only if  $\tau = 1$ .  
 771 Moreover, observe that  $K(t_0, 1) = 0$  and thus  $\rho_K \leq \rho$ , where equality can only hold if  $\tau = 1$ .  
 772 This last fact comes as no surprise, as the growth rate of all walks on  $\mathbb{Z}$  is larger or equal to the  
 773 growth rate of meanders, which are a subset of walks restricted on  $\mathbb{N}$ . It also tells us that the  
 774 three cases listed in the assertion cover all possibilities that may appear for generic walks.

775 *Zero drift case:* To prove the assertion, observe that the dominant singularity of the  
 776 generating function  $M(t) = (1 - u_1(t))Y(t)G(t, 1)/K(t, 1)$  is at  $\rho_K = \rho$  and it originates from  
 777 a simple zero in the denominator  $K(t, u)$  and from  $u_1$ . The singular expansion of  $u_1(t)$  at  $\rho$   
 778 (see Formula (41)) gives

$$779 \quad M(t) \sim \frac{G(\rho, 1)Y(\rho)}{\rho K_t(\rho, 1)} \sqrt{\frac{2\rho K_t(\rho, 1)}{K_{uu}(\rho, 1)}} \left(1 - \frac{t}{\rho}\right)^{-1/2} = \frac{G(\rho, 1)Y(\rho)\sqrt{2}}{\sqrt{\rho K_t(\rho, 1)K_{uu}(\rho, 1)}} \left(1 - \frac{t}{\rho}\right)^{-1/2}.$$

780 *Negative drift case:* We have  $\tau > 1$  and thus, by the preliminary facts listed in the first  
 781 paragraph,  $\rho_K < \rho$ . So, there is a simple zero of  $K(t, 1)$  at  $\rho_K$ , but this is cancelled, as



800 Figure 5: For the asymptotics of meanders, the key is to compare the location of the singularity  $\rho$   
 801 (the branching point of  $u_1, v_1$ ) with the zeroes of  $K(t, 1)$ , and the values  $t$  such that  $u_i(t) = 1$ .

782  $u_1(\rho_K) = 1$ . Now,  $u_1$  has a square-root type singularity at  $\rho$ , thus singularity analysis gives the  
 783 last claim of the theorem, via the following Puiseux expansion at the dominant singularity  $\rho$

$$784 \quad M(t) \sim M(\rho) + \frac{G(\rho, 1)Y(\rho)}{K(\rho, 1)} \sqrt{\frac{2\rho K_t(\rho, \tau)}{K_{uu}(\rho, \tau)}} \sqrt{1 - \frac{t}{\rho}}.$$

785 Note that there may be a second zero of  $K(t, 1)$ , say  $\rho_2$ , which is smaller than  $\rho$  (and larger  
 786 than  $\rho_K$ ). This means there is a small root  $u_2(t)$  (large roots are excluded in the negative drift  
 787 case) of the kernel satisfying  $u_2(\rho_2) = 1$ . As  $Y(t)$  contains the factor  $1 - u_2(t)$ , this zero is  
 788 cancelled. In case the root  $\rho_2$  is a multiple zero of  $K(t, 1)$ , say of order  $\omega$ , then there must be  
 789  $\omega$  roots  $u_2(t), \dots, u_{\omega+1}(t)$  which meet at  $t = \rho_2$ , causing a singularity of order  $\omega$ . But then  
 790 again the factors  $1 - u_i(t), i = 2, \dots, \omega$ , in  $Y(t)$  cancel the pole. The same happens if further  
 791 zeros of  $K(t, 1)$  appear before  $\rho$ . Hence we conclude that the dominant singularity of  $M(t)$  is  
 792 at  $\rho$  and originates from the dominant singularity of  $u_1(t)$ .

793 *Positive drift case:* Here we have several subcases. Assume first that  $\tau < 1$  or  $\tau = 1$  and  
 794  $\rho_K < \rho$ . In fact, by the preliminary facts from the first paragraph, we know that we have in  
 795 both cases  $\rho_K < \rho$ , and hence the generating function has the dominant singularity  $\rho_K$  which  
 796 comes from the kernel only and is a simple pole. This implies

$$M(t) \sim \frac{(1 - u_1(\rho_K))Y(\rho_K)G(\rho_K, 1)}{\rho_K K_t(\rho_K, 1)} \frac{1}{1 - t/\rho_K}.$$

797 If  $\tau > 1$ , like in the negative drift case, and some large root  $v$  satisfies  $v(\rho_K) = 1$ , then there is  
 798 no more a cancellation of the zero of  $K(t, 1)$  by one of the factors in  $Y(t)$ . Thus  $M(t)$  has a  
 799 simple pole at  $\rho_K$  and we get the same expression as in the other subcases.  $\square$

802 These asymptotics also allow us to get results on limit laws, as presented in the next section.

## 7 Limit law for the number of occurrences of a pattern

Our approach also allows us to count the number of occurrences of a pattern in paths. As usual, an occurrence of  $p$  in  $w$  is any substring of  $w$  that coincides with  $p$ , and when we count them we do not require that the occurrences will be disjoint. For example, the number of occurrences of 11 in 1111 is 3. One has

**Theorem 7.1** (Trivariate generating function for walks). *The generating function of the number of occurrences of the pattern  $p$  in walks on  $\mathbb{Z}$  is*

$$W(t, u, v) = \frac{1}{1 - tP(u) - t^\ell u^{\text{alt}(p)}(v-1)/(1 - (v-1)(R(t, u) - 1))}. \quad (42)$$

*Proof.* We give two proofs, each of them having its own interest. Both of them are of wider applicability, see [42, p. 60 and p. 212].

*First proof, via symbolic inclusion-exclusion.* Define a cluster as a sequence of repetitions of the pattern  $p$  (possibly overlapping), where each occurrence of  $p$  is marked by the variable  $v$ , so the set  $\mathcal{C}$  of clusters is given by  $\mathcal{C} = vp \text{Seq}(v(\bar{\mathcal{Q}} - \epsilon))$ , where  $\bar{\mathcal{Q}} - \epsilon$  is the set of nonempty complements of presuffixes of  $p$  (the generating function of which is  $R(t, u) - 1$ , see Formula (3)). Obviously,  $W(t, u, v+1) = \text{Seq}(\mathcal{S} + \mathcal{C})$ . This directly gives (42).

*Second proof, via a system/adding a jump approach.* Let  $W \equiv W(t, u, v)$  and  $W_p \equiv W_p(t, u, v)$  be the generating functions of all words and, respectively, of words ending with  $p$ , where  $v$  counts the number of occurrences of  $p$ . We show the following two identities:

$$1 + WtP = W - W_p + v^{-1}W_p, \quad (43)$$

$$Wt^\ell u^{\text{alt}(p)} = v^{-1}W_p R - (R - 1)W_p. \quad (44)$$

To show (43), take a word and add a letter to it. If the resulting word does not end with  $p$ , it is counted by  $W - W_p$ ; if it does, it is counted by  $v^{-1}W_p$ . To show (44), take a word  $w$  with  $i$  occurrences of  $p$  and consider the contribution of  $w.p$  to both sides of the equation. Adding the pattern  $p$  to  $w$  creates a number  $j$  of extra occurrences of  $p$ . The path  $w.p$  can be written in  $j$  ways as  $w'r$ , where  $w'$  ends with  $p$  and  $r$  is an autocorrelation factor, or  $j - 1$  ways if we impose that  $r \neq \epsilon$ . The word  $w.p$  therefore contributes with a factor  $v^i$  to  $Wt^\ell u^{\text{alt}(p)}$ , with a factor  $v^{i+1} + \dots + v^{i+j}$  to  $W_p R$  and with a factor  $v^{i+1} + \dots + v^{i+j-1}$  to  $W_p(R - 1)$ . The formula (42) follows.  $\square$

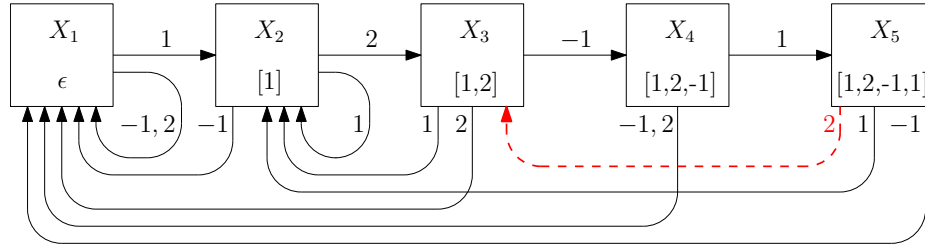
In order to get the formula for meanders, we reconsider the associated automaton and its adjacency matrix  $A$ : we add a mark  $v$  to the transition which would lead to an occurrence of  $p$  (see Figure 6). As

$$W(t, u, v) = \frac{(1 \ 0 \ \dots \ 0) \text{adj}(I - tA) \vec{\mathbf{1}}}{\det(I - tA)}, \quad (45)$$

the above theorem has the consequence that

$$\det(I - tA) = (1 - v)((1 - tP(u))R(t, u) + t^\ell u^{\text{alt}(p)}) + v(1 - tP(u)). \quad (46)$$

This last equality follows from the fact that the denominators of the rational functions (in  $\mathbb{Q}(t)$ ) in (45) and (42) are in fact the same irreducible polynomial of degree  $\ell$  in  $t$ .



$$A = \begin{bmatrix} u^{-1} + u^2 & u & \cdot & \cdot & \cdot \\ u^{-1} & u & u^2 & \cdot & \cdot \\ u^2 & u & \cdot & u^{-1} & \cdot \\ u^{-1} + u^2 & \cdot & \cdot & \cdot & u \\ u^{-1} & u & u^2 v & \cdot & \cdot \end{bmatrix}$$

839 Figure 6: Pushdown automaton and its associated adjacency matrix  $A$ . The automaton  
 840 generates walks with the set of steps  $\mathcal{S} = \{-1, 1, 2\}$ , and marks each occurrence of the pattern  
 841  $p = [1, 2, -1, 1, 2]$ . In dashed red we marked the arrow from the last state  $X_\ell$  labelled by  $a_\ell$ ,  
 842 the last letter of the pattern. Giving a weight  $v$  to this transition leads to enumerative formulas  
 843 involving  $\det(I - tA)$  as given in Equation (46).

844 Accordingly, the trivariate kernel is thus defined as  $K(t, u, v) := \det(I - tA)$ . Note that  
 845 for  $v = 0$  we get the kernel from the avoidance case (see equation (4)), and for  $v = 1$  we get  
 846  $1 - tP$  (which is, as expected, the kernel from [9]). The formulas for the trivariate generating  
 847 functions of bridges, meanders, and excursions are thus like in Theorems 3.1 and 3.2, where  
 848 the  $u_i$ 's are now the small roots of the trivariate kernel. This allows us to prove a universal  
 849 asymptotic behaviour, an instance of what Flajolet and Sedgewick pleasantly called *Borges's*  
 850 *theorem* (we comment more on it in the conclusion).

851 **Theorem 7.2** (“Borges’s theorem”: Gaussian limit laws for occurrences). *Let  $X_n$  be the random*  
 852 *variable which counts the number of occurrences of a pattern in a generic walk, bridge, meander,*  
 853 *excursion model. Then  $X_n$  has a Gaussian limiting distribution with  $\mathbb{E}[X_n] = \mu n + O(1)$  and*  
 854  *$\text{Var}[X_n] = \sigma^2 n + O(1)$  for some constants  $\mu > 0$  and  $\sigma^2 \geq 0$ :*

$$855 \quad \frac{1}{\sqrt{n}} (X_n - \mathbb{E}[X_n]) \rightarrow \mathcal{N}(0, \sigma^2).$$

856 *Proof.* All these combinatorial structures are generated by context-free grammars [7, 36, 56, 61].  
 857 Accordingly, their generating functions satisfy a positive algebraic systems. Thus, it leads to  
 858 Gaussian limit laws, as proven in [8, Theorem 9]: it comes from following the dependencies  
 859 in the graph associated with the system, and applying Hwang’s quasi-power theorem to each  
 860 component. To apply it, a positive variance condition has to be checked: this is done via  
 861 the trivariate version of the formulas of Theorems 3.1 and 3.2, with the additional variable  $v$   
 862 counting the number of occurrences of the pattern. □

## 8 Examples, pushdown automata

Directed lattice paths can be generated by context-free grammars [56], and it is well-known that context-free grammars and pushdown automata/counter automata are related (see the “Cinderella Book” [48, Sections 6.3, 6.4, and 8.5.4]). We want here to promote the idea that pushdown automata are a powerful approach for lattice path enumeration. Conversely, any pushdown automaton enumeration can be seen as a lattice path problem (the stack encodes the altitude of the path), this allows us to solve it via our vectorial kernel method. This is due to the fact that pushdown automata have mainly two notions of acceptance:

- acceptance by empty stack, this corresponds to the generation of excursions;
- acceptance by final state (whatever the value of the stack is), this corresponds to the generation of meanders.

Let us now illustrate our method by diverse examples.

**Example 8.1 (Forbidden patterns).** We start with our initial problem of lattice paths with forbidden patterns. Our theorems rediscover, in a uniform way, numerous results by different authors obtained in the last years by different methods. We present some of them in Table 3. ■

steps, pattern of length $\ell$ , model	generating function	OEIS reference <sup>8</sup>
$\mathcal{S} = \{-1, 0, 1\}$ $p = [1, 0, \dots, 0, -1]$ Bridges	$\frac{1}{\sqrt{1 - 2t - 3t^2 + 2t^\ell - 2t^{\ell+1} + t^{2\ell}}}$	$\ell = 2$ : <a href="#">A051286</a> (confirms a conjecture therein on Whitney numbers; see also [21])
$\mathcal{S} = \{-1, 0, 1\}$ $p = [1, 0, \dots, 0, -1]$ Meanders	$\frac{1 - 3t + t^\ell - \sqrt{1 - 2t - 3t^2 + 2t^\ell - 2t^{\ell+1} + t^{2\ell}}}{2t(1 - 3t + t^\ell)}$	$\ell = 2$ : <a href="#">A091964</a> (RNA folding, see [47])
$\mathcal{S} = \{-1, 0, 1\}$ $p = [1, 0, \dots, 0, -1]$ Excursions	$\frac{1 - t + t^\ell - \sqrt{1 - 2t - 3t^2 + 2t^\ell - 2t^{\ell+1} + t^{2\ell}}}{2t^2}$	$\ell = 2$ : <a href="#">A004148</a> [50] $\ell = 3$ : <a href="#">A114584</a> [58, 82]
$\mathcal{S} = \{1, -1\}$ $p = [1, -1, 1, -1, \dots, 1]$ Excursions	$\frac{1 - t^{\ell+1} - \sqrt{1 - 4t^2 + 2t^{\ell+1} + 4t^{\ell+3} - 3t^{2\ell+2}}}{2t^2(1 - t^{\ell-1})}$	$\ell = 3$ : $\approx$ <a href="#">A001006</a> (Motzkin numbers [73, 81])
$\mathcal{S} = \{1, -1\}$ $p = [1, -1, 1, -1, \dots, -1]$ Excursions	$\frac{1 - t^{\ell+2} - \sqrt{1 - 4t^2 + 6t^{\ell+2} - 4t^{2\ell+2} + t^{2\ell+4}}}{2t^2(1 - t^\ell)}$	$\ell = 4$ : <a href="#">A078481</a> (irreducible stack sortable permutations, [15, 59])

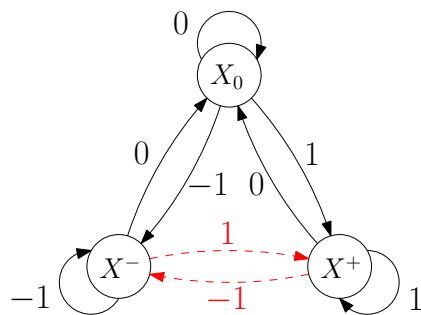
Table 3: A few examples in which our theorems produce sequences listed in the On-Line Encyclopedia of Integer Sequences. They uniformly rediscover results from several earlier works, often related to Dyck/Motzkin paths, and in certain cases also confirm conjectures.

Next we now give several examples of a wide variety of questions on lattice paths, that can also be tackled with our approach. In some cases, we omit the closed forms: they are just a direct application of our approach, once the adjacency matrix  $A$  is known. Some pushdown automata lead to a system for which an additional argument is needed to identify the factor  $G$  in Theorem 3.2; we study this situation in more detail in [3].

<sup>8</sup>Such references are links to the webpage dedicated to the corresponding sequence in the On-Line Encyclopedia of Integer Sequences (OEIS), <https://oeis.org>.



887 **Example 8.2 (Motzkin paths without peaks and valleys).**



The automaton on the left allows us to count Motzkin paths by marking each peak (i.e., the pattern  $[1, -1]$ ) and each valley (i.e., the pattern  $[-1, 1]$ ). Forbidding the transitions shown by red dashed lines, we get the automaton for Motzkin paths without peaks and valleys. Then the kernel is  $K(t, u) = \frac{t^3 u + t^2 u - t u^2 - t u - t + u}{u}$ , and it has one small root  $u_0(t) = \frac{1-t+t^2+t^3 - \sqrt{(1-t^4)(1-2t-t^2)}}{2t}$ . The degree of  $\Phi$  (as in Equation 25) is equal to the number of small roots of  $K$ , and taking care of the leading term (in  $u$ ) of  $\Phi$  we obtain

$$M(t, u) = \frac{1-t^2}{1-tu_0(t)} \frac{u-u_0(t)}{uK(t, u)}.$$

889 Setting  $u = 1$ , we get the generating function for meanders

$$M(t) = M(t, 1) = -\frac{(1+t)(1-2t-t^2) - \sqrt{(1-t^4)(1-2t-t^2)}}{2t^2(1-2t-t^2)} = \frac{\sqrt{\frac{1-t^4}{1-2t-t^2}} - t - 1}{2t^2},$$

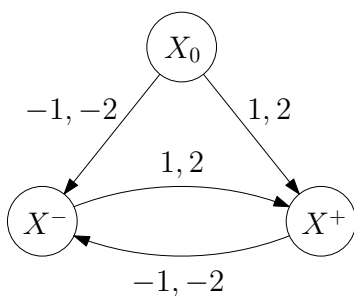
890 which corresponds to the sequence [A308435](#).

891 Setting  $u = 0$ , we get the generating function for excursions

$$E(t) = M(t, 0) = \frac{(1+t)(1-t)^2 - \sqrt{(1-t^4)(1-2t-t^2)}}{2t^2},$$

892 which corresponds to the sequence [A004149](#). ■

893 **Example 8.3 (Basketball with alternation of team scoring).**



In the pre-1980s basketball rules, each team could score 1 or 2 points at once. Zeilberger and other authors [4, 11, 23] therefore considered the model of (old-time) basketball excursions, where the step set is  $\mathcal{S} = \{2, 1, -1, -2\}$ . If we now add the constraint that each team cannot score twice in a row, this gives the automaton presented on the left. It yields the bivariate generating function

$$M(t, u) = (1 + tu + tu^2) \frac{u - (tC(t^2))^2}{u - (t(u+1))^2}, \quad \text{where } C(t) = \frac{1 - \sqrt{1-4t}}{2t}.$$

897 This yields  $M(t, 1) = C(t^2) \sqrt{\frac{1+2t}{1-2t}} = \frac{\sqrt{\frac{1+2t}{1-2t}} - 1 - 2t}{2t^2}$  and  $E(t) = M(t, 0) = C^2(t^2)$ ,

898 which correspond to the sequences [A001405](#) and [A126120](#). ■

899 **Example 8.4 (Walks/meanders/excursions of bounded height).**

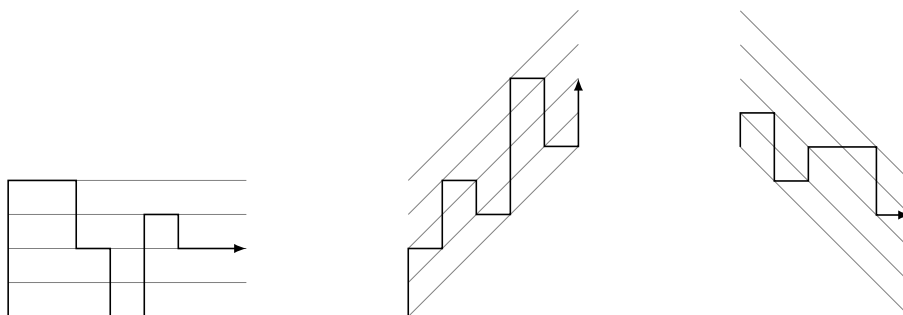
900 If one consider meanders of altitude bounded by  $h$ , they are easily generated by an automaton  
 901 with  $h$  states. In fact, the reader now familiar with our approach will realize that a pushdown  
 902 automaton with just one state is already enough: the kernel equation will encode the boundedness  
 903 by  $h$  and the positivity constraint! See [7, 12, 23] for further considerations on the closed forms  
 904 obtained. This bypasses the quickly intractable resolution of a system  $h \times h$ . ■

905 **Example 8.5 (Partially directed self-avoiding walks).**

906 We can link some self-avoiding walks and pattern avoiding lattice paths. In fact, the enumeration  
 907 and the asymptotics of self-avoiding walks in  $\mathbb{Z}^2$  is one of the famous open problems of  
 908 combinatorics and probability theory. As it is classical for intractable problems, many natural  
 909 subclasses have been introduced, and solved. E.g., partially directed self-avoiding walks with an  
 910 added constraint of living in a half-plane or a strip [6]: they have three kinds of steps, say  $n$ ,  $e$ ,  
 911 and  $s$ , and the self-avoiding condition means that factors  $ns$  and  $sn$  are disallowed. Consider the  
 912 following three models:

- 913 • in the first model, the half-plane is the one over the line  $x = 0$ ; the heights of the steps  
 914 are  $\text{alt}(n) = 1$ ,  $\text{alt}(e) = 0$  and  $\text{alt}(s) = -1$ ;
- 915 • in the second model, the half-plane is the one over the line  $x = y$ ; the heights are  
 916  $\text{alt}(n) = 1$ ,  $\text{alt}(e) = -1$  and  $\text{alt}(s) = -1$ ;
- 917 • in the third model, the half-plane is the one over the line  $x = -y$ ; the heights are  
 918  $\text{alt}(n) = 1$ ,  $\text{alt}(e) = 1$  and  $\text{alt}(s) = -1$ .

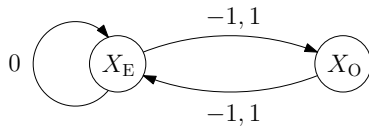
919 These models are illustrated in Figure 7. Each of them leads to an algebraic generating function,  
 920 expressible via our method as a closed form involving the roots of the kernel. ■



921 Figure 7: Some models of self-avoiding walks are encoded by partially directed lattice paths  
 922 avoiding a pattern (see [6]).

923 **Example 8.6 (Motzkin paths with horizontal steps only at even/odd altitude).**

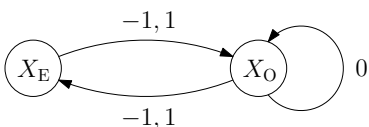
924 If one wants to allow some specific set of steps depending on the altitude modulo some period, the  
 925 vectorial kernel method will do the job! Let us illustrate this with Motzkin paths ( $\mathcal{S} = \{1, 0, -1\}$ )  
 926 with horizontal steps only at *even* altitude.



This pushdown automaton has two states, the first one corresponds to even altitudes, the second one to odd altitudes. By design, it is clear that it generates Motzkin walks with horizontal steps only at even altitude. Our vectorial kernel method then captures the additional constraint of the positivity of the stack, and thus gives the generating function for meanders (A307557) and excursions (A090344). *En passant*, let us mention that excursions are given by a pleasant continued fraction (see e.g. [41] for a nice survey):

927

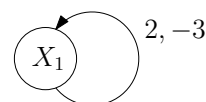
$$E(t) = \frac{1}{1 - t - \frac{t^2}{1 - \frac{t^2}{1 - t - \frac{t^2}{1 - \frac{t^2}{1 - \dots}}}}}$$



If we want to enumerate Motzkin paths with horizontal steps only at *odd* altitude, we adjust the automaton as shown on the left, and obtain the sequences A327421 (for meanders) and A090345 (for excursions), see e.g. [18, 32]. ■

929 **Example 8.7 (Duchon's club without lonely visitors).**

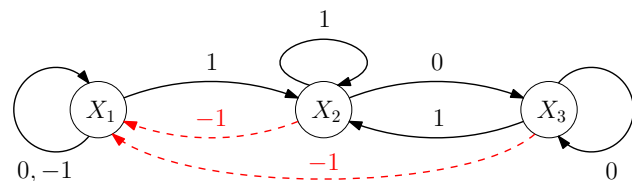
The terminology of *Duchon's club* was introduced in [9] as a playful reference to the nice article [36]. As some readers may wonder, we have to add that it is in no way an allusion to some fancy habit of our highly honourable and very respectable colleague Philippe Duchon! Here is the story. In the early evening, the club opens empty, and then, during the full night, people are entering in couples and leaving in groups of three people. Finally, the club closes empty at the end of the night. For sure, you do not want to be alone in such a club! What is the number of possible scenarios? These are excursions with step set  $\mathcal{S} = \{2, -3\}$  and never going to altitude 1. This enumeration is encoded by a single state automaton, and the kernel equation then fully handles all the constraints. This leads to the sequences A327422 (for meanders) and A327423 (for excursions). ■



931

**Example 8.8 (Counting/avoiding humps and peaks in Motzkin paths).**

Consider Motzkin walks:  $\mathcal{S} = \{-1, 0, 1\}$ . A peak is the pattern  $[1, -1]$ . A hump is an occurrence of the pattern  $[1, 0^*, -1]$ , that is, 1 followed by a (possibly empty) sequence of 0s, followed by  $-1$ . Humps were considered e.g. in [35, 60, 72]. We first consider the generating function of walks counting the number of humps, similarly to our approach from Section 7.



The transition which was forbidden in the pattern avoidance setting is here represented with a variable  $v$  in the adjacency matrix  $A$ . This yields the trivariate generating function with respect to length ( $t$ ), final altitude ( $u$ ), and number of occurrences of humps ( $v$ ). In particular, this leads to a Gaussian distribution, see Borges's Theorem 7.2.

Note that, upon a straightforward modification, the above automaton works not only for the Motzkin walks, but also for any set of steps  $\mathcal{S} \supseteq \{-1, 0, 1\}$ . Furthermore, this offers more multivariate statistics "for free". E.g., we can assign the weight  $v$  to  $X_3 \xrightarrow{-1} X_1$  and the weight  $vw$  to the transition  $X_2 \xrightarrow{-1} X_1$ : this will count humps by  $v$ , and peaks by  $w$ . Our approach then gives the following closed form for the corresponding quadrivariate generating function

$$F(t, u, v, w) = \frac{\sqrt{\Delta} + (1 - w)vt^3 + (vw - 2u - 2)t^2 + (2u + 2)t - 1}{2t(uv(w - 1)t^3 + (-uvw + u^2 + 2u + 1)t^2 - (1 + u)^2t + u)}$$

$$\Delta = ((w - 1)vt^3 - t^2vw + 1)((w - 1)vt^3 + (-vw + 4)t^2 - 4t + 1).$$

Now, the generating function of excursions with an even number of humps is

$$F(t, 0, v, w)/2 + F(t, 0, -v, w)/2.$$

If one additionally wants, say, an odd number of peaks, then the generating function is

$$\frac{1}{2} \left( \frac{F(t, 0, v, w) + F(t, 0, -v, w)}{2} - \frac{F(t, 0, v, -w) + F(t, 0, -v, -w)}{2} \right).$$

Similar manipulations for different combinations of avoidance and parity constraints for peaks and humps lead to results summarized in the following table (we give the OEIS entries for the corresponding sequences of meanders and excursions).

	even number of humps	odd number of humps	any number of humps
no peaks	M: <a href="#">A325917</a> E: <a href="#">A325918</a>	M: <a href="#">A325919</a> E: <a href="#">A325920</a>	M: <a href="#">A091964</a> E: <a href="#">A004148</a>
even number of peaks	M: <a href="#">A325921</a> E: <a href="#">A325922</a>	M: <a href="#">A325923</a> E: <a href="#">A325924</a>	M: <a href="#">A307575</a> E: <a href="#">A307576</a>
odd number of peaks	M: <a href="#">A325925</a> E: <a href="#">A325926</a>	M: <a href="#">A325927</a> E: <a href="#">A325928</a>	M: <a href="#">A307577</a> E: <a href="#">A307578</a>
any number of peaks	M: <a href="#">A307555</a> E: <a href="#">A307564</a>	M: <a href="#">A307572</a> E: <a href="#">A307573</a>	M: <a href="#">A005773</a> E: <a href="#">A001006</a>

We omit the case of walks with no hump, as it would lead to trivial rational functions. ■

In fact, there are several thousand sequences related to directed lattice paths in the *On-Line Encyclopedia of Integer Sequences*, and several hundreds of them can be generated and analysed via the method presented in this article. Let us give a few examples illustrating different types of set of steps and constraints:

- Dyck excursions with up steps colored in two colors (blue and red, denoted  $U_B$  and  $U_R$ ), and avoiding the pattern  $U_B U_R$  (A006013, which also enumerates many other objects!),
- Motzkin excursions avoiding UD, UU and DD (see A023432 and [31]),
- Schröder bridges (also called Delannoy paths) with  $k$  occurrences of DD that cross  $y = 0$  (see A110121 and [13, 80]),
- excursions from  $(0, 0)$  to  $(3n, 0)$  that use steps  $u = (2, 1)$ ,  $U = (1, 2)$ , and  $D = (1, -1)$ , and have  $k$  peaks  $uD$  or  $UD$  (see A108425 and [30]),
- Łukasiewicz paths with  $k$  U-steps that start at an even level (see A091894 and [16, 20]).

Moreover, there is a vast world of combinatorial models that can be encoded by pushdown automata and, therefore, fall under the scope of the vectorial kernel method, for example:

- patterns in permutations (see e.g. A105633, and [67, 70]),
- patterns in trees and lambda terms (see e.g. A114997, A105633, and [19, 29, 74]),
- patterns in RNA structures (see e.g. A103140 and [47, 50]),
- enumeration in Tamari-like posets (see e.g. A007477 and [5, 17]),
- a mathematical model for bobbin lace (see A291083 and [49]).

We also list 200 other examples at the webpage <https://lipn.fr/~cb/KernelMethod/>, where we maintain a list of OEIS references and publications focusing on objects which can be handled by the vectorial kernel method. Moreover, many of these OEIS entries contain conjectures on the recurrence satisfied by the sequence, or its asymptotics, or the algebraic equation satisfied by the corresponding generating function: in numerous cases such conjectures can directly be solved by the vectorial kernel method and basic computer algebra manipulations.

For some of these examples, it is possible to tackle the problem by using some ad hoc decompositions, bijections with trees, or some variation of a classical kernel method [28, 37, 43, 46, 54, 55, 63–65, 68, 71, 75, 79]. Some examples require to mark several patterns at once, or to take a diagonal of our multivariate generating function (e.g. when the number of occurrences of a given pattern is required to be proportional to the length  $n$ , or to the number of occurrences of another pattern), etc. The method is really flexible and allows playing with variants of patterns, step sets, colors, weights, values of the stack, constraints modulo 2, 3, ... an endless game!

We emphasize that the full approach is *automatable* (sic?), so all such studies can now morally be considered as computationally solved by our vectorial kernel method, with some subtleties which will be presented in the follow-up article [3]. Thus, in a style dear to the heart of Shalosh B. Ekhad and Philippe Flajolet, the vectorial kernel method can be considered as a method to automatically prove enumerative and asymptotic results for pattern occurrences in structures given by an algebraic system of functional equations.

## 9 Conclusion

In this article, we presented a unifying way which gives the generating functions and asymptotics of all families of lattice paths with a forbidden pattern, and we proved that the number of occurrences of a given pattern is normally distributed. The same techniques would, for instance, allow us to do the asymptotics of walks having exactly  $m$  occurrences of a given pattern, or to consider patterns which are no longer a word but a regular expression.

It is also nice that our approach gives a *method* to solve in an efficient way the question of the enumeration and asymptotics of words generated by a **pushdown automaton** (or words belonging to the intersection of an algebraic language and a rational language). What is more, it is possible to use our functional equation approach to analyse the intersection of two algebraic languages. Note that testing whether this intersection is empty is known to be an undecidable problem, even for deterministic context-free grammars (see e.g. [48]), so we cannot expect too much from a generic method in this case. However, we can specify the type of system of functional equations that we get: indeed, this problem is related to automata with two stacks, which, in turn, are known to have the same power as Turing machines; the evolution of these two stacks corresponds to lattice paths in the quarter plane (with steps of arbitrary length), the complexity of the problem is reflected by the fact that one can then get generating functions which are no more algebraic, D-finite, or differentially-algebraic, and we do not expect some nice universal *enumerative* closed-form results here but a wider zoo of formulas. On the other hand, some universal *asymptotic* behaviour is expected; it is highly believable that all these cases will be new instances of what Flajolet and Sedgewick called **Borges's Theorem**: *Any pattern which is not forbidden by design will appear a linear number of times in large enough structures, with Gaussian fluctuations.*

For sure, it is more a metatheorem, a natural credo, and the claim should be established rigorously in each case. Naturally, may it be with tools of probability theory or of analytic combinatorics, there is always some technical conditions to check to ensure this claim. In this article, our closed form expressions for the generating functions were one of the keys, together with the universal behaviour of the small roots. This allowed us to prove this Gaussian behaviour for the number of occurrences of any given pattern. Year after year, this claim is established for more and more combinatorial structures (it was done for patterns in Markov chains, trees, maps, permutations, context-free grammars, and now... lattice paths!).

Let us conclude with the passage of Flajolet and Sedgewick [42, p. 61] which explains where the concept of Borges's Theorem comes from:

This property is sometimes called "Borges's Theorem" as a tribute to the famous Argentinian writer Jorge Luis Borges (1899-1986) who, in his short story "The Library of Babel", describes a library so huge as to contain:

"Everything: the minutely detailed history of the future, the archangels' autobiographies, the faithful catalogues of the Library, thousands and thousands of false catalogues, the demonstration of the fallacy of those catalogues, the demonstration of the fallacy of the true catalogue, the Gnostic gospel of Basilides, the commentary on that gospel, the commentary on the commentary on that gospel, the true story of your death, the translation of every book in all languages, the interpolations of every book in all books."

1026 **Acknowledgements:** We thank the organizers and referees of the conferences AofA 2018  
1027 and LATA 2018, where we presented some aspects of this work (see [1, 2]). This work was  
1028 supported by the Austrian Science Fund (FWF), in the framework of the Special Research  
1029 Program *Algorithmic and Enumerative Combinatorics* (grant SFB F50-03) and of the project  
1030 *Analytic Combinatorics: Digits, Automata and Trees* (grant P 28466-N35).

## 1031 References

- 1032 [1] Andrei Asinowski, Axel Bacher, Cyril Banderier, and Bernhard Gittenberger. *Analytic*  
1033 *combinatorics of lattice paths with forbidden patterns: asymptotic aspects*. In *29th*  
1034 *International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the*  
1035 *Analysis of Algorithms (AofA 2018)*, volume 110 of *Leibniz International Proceedings in*  
1036 *Informatics (LIPIcs)*, pages 10.1–10.13, 2018.
- 1037 [2] Andrei Asinowski, Axel Bacher, Cyril Banderier, and Bernhard Gittenberger. *Analytic*  
1038 *combinatorics of lattice paths with forbidden patterns: enumerative aspects*. In *Language*  
1039 *and Automata Theory and Applications. LATA 2018*, volume 10782 of *Lecture Notes in*  
1040 *Computer Science*, pages 195–206. Springer, 2018.
- 1041 [3] Andrei Asinowski, Axel Bacher, Cyril Banderier, and Bernhard Gittenberger. Pushdown  
1042 automata, the vectorial kernel method, and underdetermined functional equations. *In*  
1043 *preparation*, 2019.
- 1044 [4] Arvind Ayyer and Doron Zeilberger. *The number of [old-time] basketball games with final*  
1045 *score  $n : n$  where the home team was never losing but also never ahead by more than  $w$*   
1046 *points*. *Electron. J. Combin.*, 14(1):Article R19, 8 pp., 2007.
- 1047 [5] Axel Bacher, Antonio Bernini, Luca Ferrari, Benjamin Gunby, Renzo Pinzani, and Julian  
1048 West. *The Dyck pattern poset*. *Discrete Math.*, 321:12–23, 2014.
- 1049 [6] Axel Bacher and Mireille Bousquet-Mélou. *Weakly directed self-avoiding walks*. *Journal of*  
1050 *Combinatorial Theory, Series A*, 118(8):2365–2391, 2011.
- 1051 [7] Cyril Banderier. *Combinatoire analytique des chemins et des cartes*. PhD thesis, Université  
1052 Paris VI, 2001.
- 1053 [8] Cyril Banderier and Michael Drmota. *Formulae and asymptotics for coefficients of algebraic*  
1054 *functions*. *Combinatorics, Probability and Computing*, 24(1):1–53, 2015.
- 1055 [9] Cyril Banderier and Philippe Flajolet. *Basic analytic combinatorics of directed lattice paths*.  
1056 *Theoretical Computer Science*, 281(1-2):37–80, 2002.

- 1057 [10] Cyril Banderier and Bernhard Gittenberger. [Analytic combinatorics of lattice paths: enumeration and asymptotics for the area](#). *Discrete Math. Theor. Comput. Sci. Proc.*, AG:345–355, 2006.
- 1058
- 1059
- 1060 [11] Cyril Banderier, Christian Krattenthaler, Alan Krinik, Dmitry Kruchinin, Vladimir Kruchinin, David Nguyen, and Michael Wallner. [Explicit formulas for enumeration of lattice paths: basketball and the kernel method](#). In *Lattice paths combinatorics and applications*, Developments in Mathematics Series, Vol. 58, pages 78–118. Springer, 2019.
- 1061
- 1062
- 1063
- 1064 [12] Cyril Banderier and Pierre Nicodème. [Bounded discrete walks](#). *Discrete Math. Theor. Comput. Sci.*, AM:35–48, 2010.
- 1065
- 1066 [13] Cyril Banderier and Sylviane Schwer. [Why Delannoy numbers?](#) *J. Statist. Plann. Inference*, 135(1):40–54, 2005.
- 1067
- 1068 [14] Cyril Banderier and Michael Wallner. [The kernel method for lattice paths below a rational slope](#). In *Lattice paths combinatorics and applications*, Developments in Mathematics Series, Vol. 58, pages 119–154. Springer, 2019.
- 1069
- 1070
- 1071 [15] Jean-Luc Baril. [Avoiding patterns in irreducible permutations](#). *Discrete Math. Theor. Comput. Sci.*, 17(3):13–30, 2016.
- 1072
- 1073 [16] Jean-Luc Baril, Sergey Kirgizov, and Armen Petrossian. [Enumeration of Łukasiewicz paths modulo some patterns](#). *Discrete Mathematics*, 342(4), 2019.
- 1074
- 1075 [17] Jean-Luc Baril and Jean Marcel Pallo. [Motzkin subposets and Motzkin geodesics in Tamari lattices](#). *Inf. Process. Lett.*, 114(1-2):31–37, 2014.
- 1076
- 1077 [18] Paul Barry. [Continued fractions and transformations of integer sequences](#). *J. Integer Seq.*, 12(7):Article 09.7.6, 37 pp., 2009.
- 1078
- 1079 [19] Maciej Bendkowski, Katarzyna Grygiel, Pierre Lescanne, and Marek Zaionc. [Combinatorics of  \$\lambda\$ -terms: a natural approach](#). *J. Logic Comput.*, 27(8):2611–2630, 2017.
- 1080
- 1081 [20] David Bevan. [Permutations avoiding 1324 and patterns in Łukasiewicz paths](#). *J. Lond. Math. Soc. (2)*, 92(1):105–122, 2015.
- 1082
- 1083 [21] Miklós Bóna and Arnold Knopfmacher. [On the probability that certain compositions have the same number of parts](#). *Ann. Comb.*, 14(3):291–306, 2010.
- 1084
- 1085 [22] Mireille Bousquet-Mélou. [Rational and algebraic series in combinatorial enumeration](#). In *International Congress of Mathematicians. Vol. III*, pages 789–826. EMS, 2006.
- 1086
- 1087 [23] Mireille Bousquet-Mélou. [Discrete excursions](#). *Sém. Lothar. Combin.*, 57:Article B57d, 23 pp., 2008.
- 1088



- 1089 [24] Mireille Bousquet-Mélou and Arnaud Jehanne. [Polynomial equations with one catalytic](#)  
1090 [variable, algebraic series and map enumeration](#). *Journal of Combinatorial Theory. Series B*,  
1091 96(5):623–672, 2006.
- 1092 [25] Charlotte Brennan and Simon Mavhungu. [Visits to level  \$r\$  by Dyck paths](#). *Fund. Inform.*,  
1093 117(1-4):127–145, 2012.
- 1094 [26] Fritz Carlson. [Über Potenzreihen mit ganzzahligen Koeffizienten](#). *Math. Z.*, 9(1-2):1–13,  
1095 1921.
- 1096 [27] Noam Chomsky and Marcel-Paul Schützenberger. [The algebraic theory of context-free](#)  
1097 [languages](#). In *Computer programming and formal systems*, pages 118–161. North-Holland,  
1098 Amsterdam, 1963.
- 1099 [28] Nachum Dershowitz. [Nonleaf patterns in trees: protected nodes and Fine numbers](#).  
1100 arXiv:1908.04329 [cs.DM], August 2019.
- 1101 [29] Nachum Dershowitz and Shmuel Zaks. [More patterns in trees: up and down, young and](#)  
1102 [old, odd and even](#). *SIAM J. Discrete Math.*, 23(1):447–465, 2009.
- 1103 [30] Emeric Deutsch. [Another type of lattice path](#). *Amer. Math. Monthly*, 107(4):368–370,  
1104 2000. Problem 10658, with solution by D. Callan, M. Beck, W. Bohm, R. F. McCoart, and  
1105 GCHQ Problems Group.
- 1106 [31] Emeric Deutsch and Sergi Elizalde. [Statistics on bargraphs viewed as cornerless Motzkin](#)  
1107 [paths](#). *Discrete Appl. Math.*, 221:54–66, 2017.
- 1108 [32] Emeric Deutsch, Emanuele Munarini, and Simone Rinaldi. [Skew Dyck paths, area, and](#)  
1109 [superdiagonal bargraphs](#). *J. Statist. Plann. Inference*, 140(6):1550–1562, 2010.
- 1110 [33] Emeric Deutsch and Louis W. Shapiro. [A bijection between ordered trees and 2-Motzkin](#)  
1111 [paths and its many consequences](#). *Discrete Math.*, 256(3):655–670, 2002.
- 1112 [34] Jean Dieudonné. *Calcul infinitésimal*. Hermann, Paris, 2 edition, 1980. 1st edition in 1968:  
1113 479pp., there is also an English translation of the 1st edition in 1971, 427pp.
- 1114 [35] Yun Ding and Rosena R. X. Du. [Counting humps in Motzkin paths](#). *Discrete Appl. Math.*,  
1115 160(1-2):187–191, 2012.
- 1116 [36] Philippe Duchon. [On the enumeration and generation of generalized Dyck words](#). *Discrete*  
1117 *Mathematics*, 225(1-3):121–135, 2000.
- 1118 [37] Maciej Dziemiańczuk. [On directed lattice paths with vertical steps](#). *Discrete Math.*,  
1119 339(3):1116–1139, 2016.
- 1120 [38] Sen-Peng Eu, Shu-Chung Liu, and Yeong-Nan Yeh. [Dyck paths with peaks avoiding or](#)  
1121 [restricted to a given set](#). *Stud. Appl. Math.*, 111(4):453–465, 2003.

- 1122 [39] Bertrand Eynard. *Counting Surfaces*, volume 70 of *Progress in Mathematical Physics*.  
1123 Springer/Birkhäuser, 2016. CRM Aisenstadt chair lectures.
- 1124 [40] Guy Fayolle, Roudolf Iasnogorodski, and Vadim Malyshev. *Random Walks in the Quarter-*  
1125 *plane*, volume 40 of *Applications of Mathematics*. Springer-Verlag, 1999.
- 1126 [41] Philippe Flajolet. *Combinatorial aspects of continued fractions*. *Discrete Math.*, 32(2):125–  
1127 161, 1980.
- 1128 [42] Philippe Flajolet and Robert Sedgewick. *Analytic Combinatorics*. Cambridge University  
1129 Press, 2009.
- 1130 [43] Evangelos Georgiadis, David Callan, and Qing-Hu Hou. *Circular digraph walks,  $k$ -balanced*  
1131 *strings, lattice paths and Chebychev polynomials*. *Electron. J. Combin.*, 15(1):Article R108,  
1132 11 pp., 2008.
- 1133 [44] Ira M. Gessel. *A factorization for formal Laurent series and lattice path enumeration*.  
1134 *Journal of Combinatorial Theory, Series A*, 28(3):321–337, 1980.
- 1135 [45] Leonidas J. Guibas and Andrew M. Odlyzko. *String overlaps, pattern matching, and*  
1136 *nontransitive games*. *J. Combin. Theory Ser. A*, 30(2):183–208, 1981.
- 1137 [46] Benjamin Hackl, Clemens Heuberger, and Helmut Prodinger. *Ascents in non-negative*  
1138 *lattice paths*. arXiv:1801.02996 [math.CO], January 2018. Long version of *Counting*  
1139 *Ascents in Generalized Dyck Paths*, in Proceedings of Analysis of Algorithms 2018.
- 1140 [47] Ivo L. Hofacker, Christian M. Reidys, and Peter F. Stadler. *Symmetric circular matchings*  
1141 *and RNA folding*. *Discrete Math.*, 312(1):100–112, 2012.
- 1142 [48] John E. Hopcroft, Raghavan Motwani, and Jeffrey D. Ullman. *Introduction to Automata*  
1143 *Theory, Languages, and Computation*. Addison-Wesley, 3rd edition, 2006. (First edition in  
1144 1979).
- 1145 [49] Veronika Irvine, Stephen Melczer, and Frank Ruskey. *Vertically constrained Motzkin-like*  
1146 *paths inspired by bobbin lace*. *Electron. J. Combin.*, 26(2):Article P2.16, 27 pp., 2019.
- 1147 [50] Emma Y. Jin, Jing Qin, and Christian M. Reidys. *Combinatorics of RNA structures with*  
1148 *pseudoknots*. *Bull. Math. Biol.*, 70(1):45–67, 2008.
- 1149 [51] Manuel Kauers and Peter Paule. *The Concrete Tetrahedron*. Texts and Monographs in  
1150 Symbolic Computation. Springer, 2011. Symbolic sums, recurrence equations, generating  
1151 functions, asymptotic estimates.
- 1152 [52] Donald Ervin Knuth. *The Art of Computer Programming. Vol. 1: Fundamental Algorithms*.  
1153 Addison-Wesley, 1968.

- 1154 [53] Christian Krattenthaler. [Lattice path enumeration](#). In *Handbook of enumerative combinatorics*, pages 589–678. Discrete Math. Appl. (Boca Raton), CRC Press, 2015.  
1155
- 1156 [54] Germain Kreweras and Paul Moszkowski. [A new enumerative property of the Narayana numbers](#). *J. Statist. Plann. Inference*, 14(1):63–67, 1986.  
1157
- 1158 [55] Germain Kreweras and Yves Poupard. [Subdivision des nombres de Narayana suivant deux paramètres supplémentaires](#). *European J. Combin.*, 7(2):141–149, 1986.  
1159
- 1160 [56] Jacques Labelle and Yeong-Nan Yeh. [Generalized Dyck paths](#). *Discrete Mathematics*, 82(1):1–6, 1990.  
1161
- 1162 [57] Kostas Manes, Aristidis Sapounakis, Ioannis Tasoulas, and Panagiotis Tsikouras. [Strings of length 3 in Grand-Dyck paths and the Chung–Feller property](#). *Electron. J. Combin.*, 19(2):Article P2, 10 pp., 2012.  
1163  
1164
- 1165 [58] Kostas Manes, Aristidis Sapounakis, Ioannis Tasoulas, and Panagiotis Tsikouras. [Equivalence classes of ballot paths modulo strings of length 2 and 3](#). *Discrete Mathematics*, 339(10):2557–2572, 2016.  
1166  
1167
- 1168 [59] Toufik Mansour. [Statistics on Dyck paths](#). *J. Integer Seq.*, 9(1):Article 06.1.5, 13 pp., 2006.  
1169
- 1170 [60] Toufik Mansour and Mark Shattuck. [Counting humps and peaks in generalized Motzkin paths](#). *Discrete Appl. Math.*, 161(13-14):2213–2216, 2013.  
1171
- 1172 [61] Donatella Merlini, Douglas G. Rogers, Renzo Sprugnoli, and Maria Cecilia Verri. [Under-diagonal lattice paths with unrestricted steps](#). *Discrete Appl. Math.*, 91(1-3):197–213, 1999.  
1173  
1174
- 1175 [62] Sri Gopal Mohanty. *Lattice Path Counting and Applications*. Academic Press, 1979.
- 1176 [63] Emanuele Munarini and Norma Zagaglia Salvi. [Binary strings without zigzags](#). *Sém. Lothar. Combin.*, 49:Article B49h, 15 pp., 2002.  
1177
- 1178 [64] Heinrich Niederhausen and Shaun Sullivan. [Ballot paths avoiding depth zero patterns](#). *J. Combin. Math. Combin. Comput.*, 74:181–192, 2010.  
1179
- 1180 [65] Heinrich Niederhausen and Shaun Sullivan. [Pattern avoiding ballot paths and finite operator calculus](#). *J. Statist. Plann. Inference*, 140(8):2312–2320, 2010.  
1181
- 1182 [66] Heinrich Niederhausen and Shaun Sullivan. [Counting depth zero patterns in ballot paths](#). *Integers*, 12(2):215–236, 2012.  
1183
- 1184 [67] Ran Pan, Dun Qiu, and Jeffrey Remmel. [Counting consecutive pattern matches in  \$\mathcal{S}\_n\(132\)\$  and  \$\mathcal{S}\_n\(123\)\$](#) . *Adv. in Appl. Math.*, 105:130–167, 2019.  
1185

- 1186 [68] Ran Pan and Jeffrey B. Remmel. Paired patterns in lattices paths. In *Lattice paths*  
1187 *combinatorics and applications*, Developments in Mathematics Series, Vol. 58, pages  
1188 382–418. Springer, 2019.
- 1189 [69] Youngja Park and Seung Kyung Park. Enumeration of generalized lattice paths by string  
1190 types, peaks, and ascents. *Discrete Math.*, 339(11):2652–2659, 2016.
- 1191 [70] Robert Parviainen. Lattice path enumeration of permutations with  $k$  occurrences of the  
1192 pattern 2-13. *J. Integer Seq.*, 9(3):Article 06.3.2, 8 pp., 2006.
- 1193 [71] Dun Qiu and Jeffrey Remmel. Quadrant marked mesh patterns in 123-avoiding permutations.  
1194 *Discrete Math. Theor. Comput. Sci.*, 19(2):Article 12, 49 pp., 2018.
- 1195 [72] Amitai Regev. Identities for the number of standard Young tableaux in some  $(k, \ell)$ -hooks.  
1196 *Sém. Lothar. Combin.*, 63:Article B63c, 8 pp., 2010.
- 1197 [73] Céline Righi. Number of “ $udu$ ”s of a Dyck path and  $ad$ -nilpotent ideals of parabolic  
1198 subalgebras of  $sl_{l+1}(\mathbb{C})$ . *Sém. Lothar. Combin.*, 59:Article B59c, 17 pp., 2008.
- 1199 [74] John Riordan. Enumeration of plane trees by branches and endpoints. *J. Comb. Theory,*  
1200 *Ser. A*, 19(2):214–222, 1975.
- 1201 [75] Aristidis Sapounakis, Ioannis Tasoulas, and Panagiotis Tsikouras. Counting strings in Dyck  
1202 paths. *Discrete Math.*, 307(23):2909–2924, 2007.
- 1203 [76] Marcel-Paul Schützenberger. On context-free languages and push-down automata. *Infor-*  
1204 *mation and Control*, 6:246–264, 1963.
- 1205 [77] Marcel-Paul Schützenberger. On the synchronizing properties of certain prefix codes.  
1206 *Information and Control*, 7:23–36, 1964.
- 1207 [78] Richard P. Stanley. *Enumerative Combinatorics. Vol. 1*, volume 49 of *Cambridge Studies*  
1208 *in Advanced Mathematics*. Cambridge University Press, 2nd edition, 2011.
- 1209 [79] Paul R. Stein and Michael S. Waterman. On some new sequences generalizing the Catalan  
1210 and Motzkin numbers. *Discrete Math.*, 26(3):261–272, 1979.
- 1211 [80] Robert A. Sulanke. Objects counted by the central Delannoy numbers. *J. Integer Seq.*,  
1212 6(1):Article 03.1.5, 19 pp., 2003.
- 1213 [81] Yidong Sun. The statistic “number of  $udu$ ’s” in Dyck paths. *Discrete Math.*, 287(1-3):177–  
1214 186, 2004.
- 1215 [82] Yan Zhuang. A generalized Goulden-Jackson cluster method and lattice path enumeration.  
1216 *Discrete Math.*, 341(2):358–379, 2018.