BASIC ANALYTIC COMBINATORICS OF DIRECTED LATTICE PATHS

CYRIL BANDERIER AND PHILIPPE FLAJOLET

ABSTRACT. This paper develops a unified enumerative and asymptotic theory of *directed 2-dimensional lattice paths* in half-planes and quarter-planes. The lattice paths are specified by a finite set of rules that are both time and space homogeneous, and have a privileged direction of increase. (They are then essentially 1-dimensional objects.) The theory relies on a specific "kernel method" that provides an important decomposition of the algebraic generating functions involved, as well as on a generic study of singularities of an associated algebraic curve. Consequences are precise computable estimates for the number of lattice paths of a given length under various constraints (bridges, excursions, meanders) as well as a characterization of the limit laws associated to several basic parameters of paths.

To Maurice Nivat, with many thanks for so many things!

INTRODUCTION

By a *lattice path* is meant in all generality a polygonal line of the discrete Cartesian plane $\mathbb{Z}\times\mathbb{Z}$. The lattice paths to be considered here are specified by a finite set of simple rules: typically, from each point, there is a finite set of allowable moves that are both "time independent" and "space independent". Throughout this study, we also assume the existence of some privileged *direction of increase* (the horizontal axis, say), so that paths become essentially similar to one-dimensional objects, namely, walks on the line. Such *directed* lattice paths intervene in many areas of mathematics and computer science. They play a rôle, for instance, in probability theory (sums of discrete random variables), statistics (non-parametric tests), formal language theory, random generation of planar diagrams (animals and polyominoes), the analysis of dynamic data structures, and queueing theory models.

In probability theory, lattice paths describe the evolution of sums of independent discrete random variables, for instance, the succession of your gains if a die is repeatedly cast and your capital is increased by j when face number j shows up. A typical question in this context is the following: *Determine the probability of a "lucky game" in the sense that, at* any *time t, the partial gain is at least as large as the "mean gain", $\frac{7}{2}t$.* Such questions are indeed addressed by classical probability theory, with Brownian motion entering the game. However, by design, stochastic processes only provide a first-order asymptotic theory, while some purely discrete phenomena remain out of reach of this theory.

Date: August 26, 2001.

Key words and phrases. Lattice path, analytic combinatorics, generating function, kernel method, algebraic function, singularity analysis, generalized ballot problem, Catalan numbers.

Statistics, though not our primary motivation in this paper, is historically another important source of problems regarding lattice paths. We may mention the Kolmogorov-Smirnov test in non-parametric statistics that aims at discerning whether two random variates have the same distribution (see, e.g., [47]). As a matter of fact, the early books on lattice path combinatorics and lattice path statistics by Mohanty and Narayana [57, 59] specifically draw some of their motivations from such questions.

In discrete mathematics, all sorts of constrained lattice paths serve to describe apparently complex objects. Two-sorted permutations are for instance equivalent to paths made of horizontal and vertical steps that connect the origin to a point lying on the main diagonal—such facts are directly relevant to the analysis of the merge-sort and shellsort algorithms [48, 69, 74]. Dyck paths that are closely related to diagonal paths describe traversal sequences of general and binary trees; they belong to what Riordan has named the "Catalan domain", that is, the orbit of structures counted by the Catalan numbers, $\frac{1}{n+1}\binom{2n}{n}$. The wealth of properties surrounding Dyck paths can be perceived when examining either Gould's monograph [41] that lists 243 references or from Exercise 6.19 in Stanley's book [72] whose statement alone spans more than ten full pages. More generally, trees constrained by degrees—e.g., term trees in free magmas, of interest in formal semantics [60]—are known to be bijectively equivalent to Łukasiewicz words, themselves isomorphic to lattice paths of a special form; Lothaire's book offers a good description within the framework of combinatorics on words [52, Chap. 11].

Lattice paths also intervene in the analysis of dynamically evolving structures, and, as such, they surface in the continuous as well as discrete parts of the theory. On the discrete side, we have Flajolet's combinatorial theory of continued fractions [29] motivated by Françon's theory of "histories" of dynamic data structures [32, 36] or Knuth's dynamic storage allocation model (see [46, 2.2.2–13] for the statement of the problem and [30, 75] for solutions). As regards continuous aspects, the Karlin-McGregor theory of birth-death processes (of which [33, 58] offer lattice-path perspectives), itself closely related to various queueing theory models, involves lattice paths that describe an interesting collection of events (the embedded Markov chain). The recent book by Fayolle *et al.* on random walks in the quarter-plane [26] is historically motivated by such queueing theory questions [25].

Word representations of lattice paths also provide many examples of context-free languages. This side of the coin is closely related to encodings of trees by words, so that Dyck paths (that are associated to general trees and binary trees) and Motzkin paths (that encode unary-binary trees) play an especially important rôle. The theory of context-free languages and pushdown automata then combines nicely with the Chomsky-Schützenberger theorems [10, 73], to the effect that many types of paths can be *a priori* recognized as admitting generating functions that are algebraic. Examples are provided by Labelle and Yeh [49, 50], Merlini *et al.* [56], and Duchon [22]. (In return, enumerative studies related to context-free languages can sometimes provide structural information on generation mechanisms and formal languages as is evidenced by the analytic theory of inherent ambiguity of [31].)

Finally, because of the rich combinatorics surrounding them, lattice paths intervene at many places in the random generation of structured objects. The problem there is to draw a combinatorial object from some class $\mathcal{C}$, and do so uniformly at random amongst all objects of size n in $\mathcal{C}$. Strong decomposability properties of paths usually make random generation possible in low polynomial time (usually with a complexity between $O(n)$ and $O(n^2)$). Consequently, any easily computable bijection between a class $\mathcal{C}$ and a class of simple enough lattice paths induces a random generation algorithm for $\mathcal{C}$. Known examples include the random generation of two dimensional diagrams like polyominoes and animals. For instance, the Delest-Viennot methodology of [18] allows us to generate parallelogram polyominoes in linear time; the rejection methods of the "Florence School" [8] make it possible to generate various types of directed lattice animals in a surprisingly efficient manner. The design of such algorithms is clearly dependent on the basic combinatorics of lattice paths while the corresponding performance analyses rely on fine probabilistic estimates of characteristic properties of paths; see Louchard's contribution [53] for a neat example and the paper [4] for algebraic techniques related to the present paper.

In this introduction, we cannot do more than scratch the surface of such rich combinatorial, probabilistic, and algorithmic aspects of lattice paths. Accordingly we cut short our discussion of motivations at this point.

Scope of the paper. This paper assembles combinatorics of words and paths, some algebra of formal power series, and complex analysis. Under this angle, we believe the enterprise to be original. Quite a lot is otherwise known regarding probabilistic properties of paths, as these represent sums of random variables. Accordingly, our treatment can be, to some extent, regarded as a parallel of probabilistic-analytic methods in the realm of enumerative combinatorics.

In Section 2, we show that the counting generating functions of paths of various sorts are invariably *algebraic functions*. This algebraic character is predictable since the word encodings of the object considered are clearly recognizable by deterministic pushdown automata, hence are deterministic context-free languages. However, for directed lattice paths, we demonstrate that a strong algebraic *decomposability* prevails that is obtained by a specific technique, the *"kernel method"* (historical remarks are given at the end of Section 2.2) and is not clearly visible on combinatorial and grammatical descriptions. Our purpose in this paper is to arrive eventually at a complete characterization of the singular structure of intervening generation functions (Section 3)—by virtue of the method of *singularity analysis*, this leads to very precise asymptotic information on the counting quantities involved. At this level also, the decomposability granted by the kernel method is central as it enables us to determine the location and nature of dominant singularities. Then, once the singular structure of counting generating functions has been extracted, tight estimates on probability distributions of parameters follow easily: see Section 4 for a sample of what can be done. Section 5 sketches extensions to the enumeration of certain types of planar objects provided they satisfy a strong directedness condition.

1. Lattice paths and generating functions

This section presents the varieties of lattice paths to be studied as well as their companion generating functions.

Definition 1. *Fix a finite set of vectors of $\mathbb{Z} \times \mathbb{Z}$, $\mathcal{S} = \{(a_1, b_1), \ldots, (a_m, b_m)\}$. A* lattice path *or* walk *relative to $\mathcal{S}$ is a sequence $v = (v_1, \ldots, v_n)$ such that each v_j is in $\mathcal{S}$. The geometric realization of a lattice path $v = (v_1, \ldots, v_n)$ is the sequence*

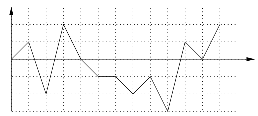
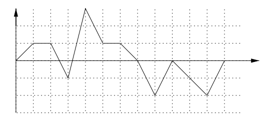
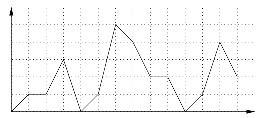
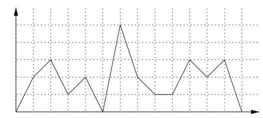
	ending anywhere	ending at 0
unconstr. (on $\mathbb{Z}$)	 walk/path ($\mathcal{W}$) $W(z,1) = \dfrac{1}{1 - zP(1)}$	 bridge ($\mathcal{V}$) $V(z) \equiv W_0(z) = z \sum_{i=1}^{c} \dfrac{u_i'(z)}{u_i(z)}$
constr. (on $\mathbb{Z}_{\geq 0}$)	 meander ($\mathcal{F}$) $F(z,1) = \dfrac{1}{1 - zP(1)} \prod_{i=1}^{c} (1 - u_i(z))$	 excursion ($\mathcal{E}$) $E(z) \equiv F_0(z) = \dfrac{(-1)^{c-1}}{p_{-c} z} \prod_{i=1}^{c} u_i(z)$

FIGURE 1. The four types of paths: walks, bridges, meanders, and excursions and the corresponding generating functions.

of points $(P_0, P_1, \ldots, P_n)$ *such that* $P_0 = (0,0)$ *and* $\overrightarrow{P_{j-1}P_j} = v_j$. *The quantity* n *is referred to as the* size *of the path.*

In the sequel, we shall identify a lattice path with the polygonal line admitting $P_0, \ldots, P_n$ as vertices. The elements of $\mathcal{S}$ are called *steps* or *jumps*, and we also refer to the vectors $\overrightarrow{P_{j-1}P_j} = v_j$ as the steps of a particular path.

Various constraints will be imposed on paths. In particular we restrict attention throughout this paper to *directed paths* defined by the fact that if (a,b) lies in $\mathcal{S}$, then necessarily one should have $a > 0$. In other words, a step always entails progress along the horizontal axis and the geometric realization of the path naturally lives in the half plane $\mathbb{Z}_{\geq 0} \times \mathbb{Z}$. (This constraint rules out paths like the ones occurring in Pólya's "drunkard problem" as described in the attractive booklet of Doyle and Snell [19]; it also implies that the paths studied can be treated essentially as 1-dimensional objects.) The following conditionings are to be considered (Figure 1).

Definition 2. *A* bridge *is a path whose end-point* P_n *lies on the* x*-axis. A* meander *is a path that lies in the quarter plane* $\mathbb{Z}_{\geq 0} \times \mathbb{Z}_{\geq 0}$. *An* excursion *is a path that is at the same time a meander and a bridge; it thus connects the origin to a point lying on the* x*-axis and involves no point with negative* y*–coordinate.*

A family of paths is said to be simple *if each allowed step in* $\mathcal{S}$ *(Definition 1) is of the form* $(1,b)$ *with* $b \in \mathbb{Z}$. *In this case, we also abbreviate* $\mathcal{S}$ *as* $\mathcal{S} = \{b_1, \ldots, b_m\}$.

In the simple case the size of a path coincides with its span along the horizontal direction, that is, its *length*. The terminology of bridges, meanders, and excursions is chosen to be consistent with the standard one adopted in Brownian motion theory; see, e.g., [62].

The main objective of this paper is to enumerate exactly as well as asymptotically paths, bridges, and meanders, this with special attention to simple families. Once the set of steps is fixed, we let $\mathcal{W}$ and $\mathcal{V}$ denote the set of paths and bridges respectively ($\mathcal{W}$ being reminiscent of "walk"); we denote by $\mathcal{F}$ and $\mathcal{E}$ the set of meanders and excursions ($\mathcal{E}$ being reminiscent of "excursion").

Given a class $\mathcal{C}$ of paths, we let $\mathcal{C}_n$ denote the subclass of paths that have size n, and, whenever appropriate, $\mathcal{C}_{n,k} \subset \mathcal{C}_n$ those that have final vertical abscissa (also known as "final altitude") equal to k. With the convention of using standard fonts to denote cardinalities of the corresponding sets (themselves in calligraphic style), $C_n = \mathrm{card}(\mathcal{C}_n)$ and $C_n = \mathrm{card}(\mathcal{C}_{n,k})$, the corresponding (ordinary) *generating functions* (GF's) are then

$$C(z) := \sum_n C_n z^n, \qquad C(z,u) = \sum_{n,k} C_{n,k} u^k z^n.$$

This paper is entirely devoted to characterizing these generating functions: they are either rational functions (W) or algebraic functions (V, E, F). As we shall see, a strong algebraic decomposition prevails which, as opposed to other approaches, renders the calculation of the GF's effective. Even more importantly, the decomposability of GF's makes it possible to extract their singular structure, and in turn solve the corresponding asymptotic enumeration problems in a wholly satisfactory fashion.

Weighted paths. For several applications, it is useful to associate *weights* to single steps. In this case, the set of steps $\mathcal{S}$ is coupled with a system of weights $\Pi = \{w_1, \dots, w_m\}$, with $w_j > 0$ the weight associated to $(a_j, b_j) \in \mathcal{S}$; the weight of a path is then defined as the *product* of the weights of its individual steps. Then the quantity C_n, still referred to as *number of paths* (of size n), represents the total weight of all paths of size n. Such weighted paths cover several situations of interest: (i) combinatorial paths in the standard sense above when each $w_j = 1$; (ii) paths with coloured steps, e.g., $w_j = 2$ means that the corresponding step (a_j, b_j) has two possible coloured incarnations (say blue and yellow); (iii) $\sum w_j = 1$ corresponds to a probabilistic model of paths where, at each stage, step (a_j, b_j) is chosen with probability w_j.

2. Algebraic structures and the kernel method

In this section, we characterize the generating functions of the four types of directed paths (unconstrained, bridges, meanders, and excursions). For ease of exposition, we restrict attention to simple families of paths till Section 5, where we briefly discuss the more general directed models. It will be seen that a specific algebraic curve, the "characteristic curve" plays a central rôle. In this section, a modicum of analysis is introduced for convenience, but it is limited to the vicinity of $z = 0$, and consequently, it is largely equivalent to formal series manipulations[1].

[1]Following a remark by a referee, we note that analyticity considerations in this section could be logically dispensed with; see Gessel's paper [38] for a proper framework. However, the authors' feeling is that purely algebraic proofs, though feasible, tend to be less transparent. More importantly, analyticity considerations developed here serve as a useful preparation for our "nonlocal" treatment of singularities in the next section.

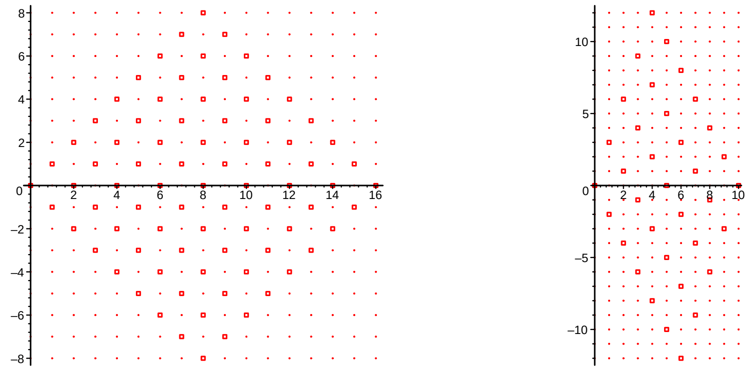


FIGURE 2. Fragments of the sublattices accessible from the origin by the Dyck walk ($\mathcal{S} = \{-1, +1\}$) and Duchon's clubs ($\mathcal{S} = -2, +3\}$). The periods are 2 and 5 respectively.

Definition 3. *Let $\mathcal{S} = \{b_1, \dots b_m\}$ be a simple set of jumps, with $\Pi = \{w_1, \dots, w_j\}$ the corresponding system of weights ($w_j \equiv 1$ in the unweighted case). The* characteristic polynomial *of $\mathcal{S}$ is defined as the polynomial in u, u^{-1} (a Laurent polynomial)*

$$P(u) := \sum_{j=1}^{m} w_j\, u^{b_j}.$$

Let $c = -\min_j b_j$ and $d = \max_j b_j$ be the two extreme vertical amplitudes of any jump, and assume throughout $c, d > 0$. The characteristic curve *of the lattice paths determined by $\mathcal{S}$ is the plane algebraic curve defined by the equation*

$$1 - zP(u) = 0, \qquad \text{or equivalently} \quad u^c - z(u^c P(u)) = 0. \tag{1}$$

The quantity $K(z, u) := u^c - zu^c P(u)$ is also referred to as the kernel *and Equation* (1) *as the* kernel equation..

As we shall see the characteristic equation plays a central rôle, the second form being the entire version (that is, a form without negative powers).

We also need to introduce technical conditions on periodicities. In a coin-tossing game ($\mathcal{S} = \{-1, +1\}$) for instance, a bridge or an excursion only exists for even lengths; consequently, what is observed of a random path at time n depends on the residue class of n modulo 2 (Figure 2).

Definition 4. *A Laurent series $h(z) = \sum_{n \geq -a} h_n z^n$ is said to admit* period p *if there exists a Laurent series H and an integer b such that*

$$h(z) = z^b H(z^p); \tag{2}$$

the largest p such that a decomposition (2) holds is called the period *of h and is denoted by* $\operatorname{per}(h)$. *The series h is called* aperiodic *if* $\operatorname{per}(h) = 1$.

A simple walk defined by the set of jumps $\mathcal{S}$ is said have period p *if the characteristic polynomial $P(u)$ has period p.*

A simple walk is said to be reduced *if the gcd of the jumps is equal to 1.*

In what follows, we systematically restrict attention to *reduced walks* since, up to a linear change of abscissa, any walk can be reduced. For instance, the walks

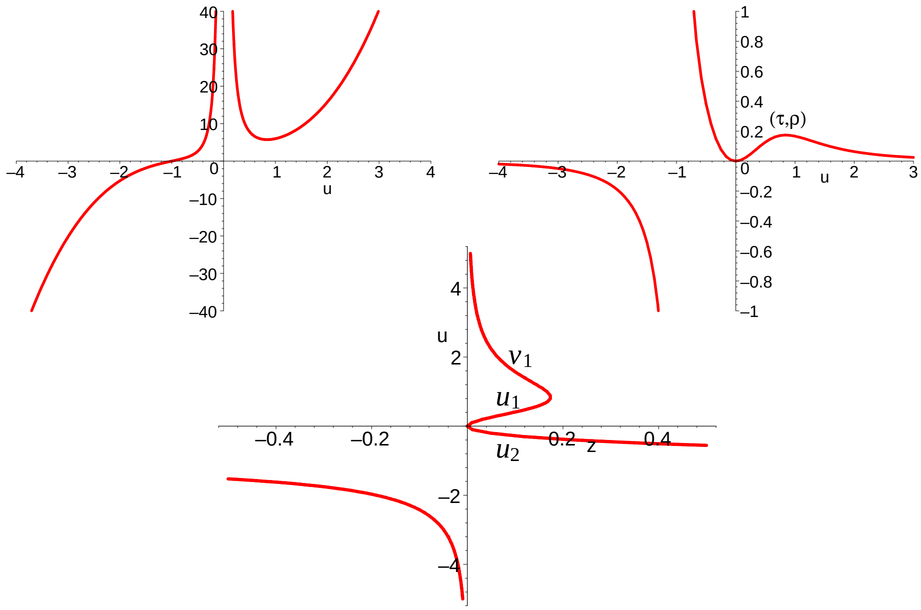


FIGURE 3. Graphs associated to the set of jumps $\mathcal{S} = \{-2,-1,0,1,2,3\}$, with characteristic polynomial $P(u) = u^{-2} + u^{-1} + 1 + u + u^2 + u^3$. Top: the graphs of $P(u)$ and $1/P(u)$ for real u. Bottom: the three real branches of the characteristic curve, one large of order $z^{-1/3}$, and two small of order $\pm z^{1/2}$ (two complex branches of order $e^{\pm 2i\pi/3} z^{-1/3}$ are not shown).

corresponding to $\mathcal{U} = \{-3,+3\}$ are transformed (upon shrinking the vertical axis by a factor of $\frac{1}{3}$) into the reduced form $\mathcal{S} = \{-1,+1\}$. (Aperiodic walks are from their definition automatically reduced.) *Periodic walks* live on sublattices: the walks associated to $\mathcal{S} = \{-1,+1\}$ (Dyck walks) and $\mathcal{T} = \{-1,0,+1\}$ (Motzkin walks) are naturally reduced, but Dyck walks are periodic with $p = 2$ (since $uP(u) = 1 + u^2$), while Motzkin walks are aperiodic; "Duchon's clubs" studied below and defined by $\mathcal{S} = \{-2,+3\}$ have period $p = 5$ (since $u^2 P(u) = 1 + u^5$), etc.

Notice that, if we write

$$P(u) = \sum_{j=1}^{m} w_j u^{b_j}, \qquad w_j \neq 0,\ b_j \in \mathbb{Z}, \tag{3}$$

the period of P (and of the set of jumps $\mathcal{S}$) is

$$p = \operatorname{per}(P) = \gcd(b_2 - b_1, \ldots, b_m - b_1).$$

Also, by the strong form of the triangle inequality, for an aperiodic $P(u)$, the *strict* inequality holds in

$$|P(u)| < P(|u|) \quad \text{for all } u \in \mathbb{C} \setminus \mathbb{R}_{>0}. \tag{4}$$

It proves convenient to rewrite

$$P(u) = \sum_{k=-c}^{d} p_k u^k.$$

Examination of the asymptotic regimes consistent with the characteristic equation near $z = 0$ shows that the equation can only be satisfied if one of the two relations,

$$p_d z u^d \sim 1 \qquad \text{or} \qquad p_{-c} z u^{-c} \sim 1 \qquad (z \to 0), \tag{5}$$

is satisfied. The characteristic equation being of degree $c + d$ in u is known to have generically $c + d$ roots; these constitute the *branches* of a single algebraic curve defined by (1) and called the *characteristic curve*. Then, as suggested by (5), one expects, in the complex domain (for z near 0), c "small branches" that we write as $u_1, \ldots, u_c$ and d "large branches" $v_1 \equiv u_{c+1}, \ldots, v_d \equiv u_{c+d}$ satisfying (Figure 3)

$$u_j(z) \sim e^{2i(j-1)\pi/c}(p_{-c})^{1/c} z^{1/c}, \qquad v_k(z) \sim e^{2i(1-k)\pi/d}(p_d)^{-1/d} z^{-1/d}. \tag{6}$$

For determinacy, *one restricts attention to the complex plane slit along the negative real axis*, which allows us to talk freely of the individual branches in the sequel.

The informal discussion summarized by (6) is vindicated by the classical theory of Newton-Puiseux expansions—the fundamental result in the elementary theory of algebraic curves that determines constructively all the possible behaviours of solutions of polynomial equations. For an exposition, we refer to one of the many excellent books on the basic theory of algebraic curves, e.g., [1, 45]. Precisely, the general theory teaches us that the small branches are conjugate of each other at 0, and similarly for the large branches at ∞. This means that there exist functions A and B analytic at 0 and nonzero there, such that, in a neighbourhood of 0, one has

$$\begin{array}{lllll} u_j(z) & = & \omega^{j-1} z^{1/c} A(\omega^{j-1} z^{1/c}) & = & u_1(e^{2i(j-1)\pi} z), \qquad \omega = e^{2i\pi/c} \\ v_k(z) & = & \varpi^{1-k} z^{-1/d} B(\varpi^{k-1} z^{1/d}) & = & v_1(e^{2i(k-1)\pi} z), \qquad \varpi = e^{2i\pi/d}. \end{array} \tag{7}$$

In summary, the u_j and v_ℓ organize themselves into two "cycles" of c and d elements respectively; for analytic details, we refer to Hille's crisp presentation based on monodromy and analytic continuation in [44].

The branch u_1 defined near 0 by (6) is real positive and is called the *principal* (small) branch. The graph of branches is obtained by interchanging the axes in the graph of $1/P(u)$, with u_1 appearing as the real positive branch near the origin; see Figure 3 for an example. We shall prove in Section 3 that in a proper sense u_1 "dominates" all the other small branches.

2.1. **Walks and bridges.** We start with the easy case of unconstrained walks and bridges. This already makes use of the characteristic curve and some of its branches.

Theorem 1. *The bivariate generating function (BGF) of paths (with z marking size and u marking final altitude) relative to a simple set of steps $\mathcal{S}$ with characteristic polynomial $P(u)$ is a rational function. It is given by*

$$W(z,u) = \frac{1}{1 - zP(u)}. \tag{8}$$

The GF of bridges is an algebraic function given by

$$V(z) = z \sum_{j=1}^{c} \frac{u_j'(z)}{u_j(z)} = z \frac{d}{dz} \log\left(u_1(z) \cdots u_c(z)\right), \tag{9}$$

where the expressions involve all the small branches $u_1, \ldots, u_c$ *of the characteristic curve (1). Generally, the GF* W_k *of paths terminating at altitude* k *is, for* $-\infty < k < c$,

$$W_k(z) = z \sum_{j=1}^{c} \frac{u_j'(z)}{u_j(z)^{k+1}} = -\frac{z}{k}\frac{d}{dz}\left(\sum_{j=1}^{c} u_j(z)^{-k}\right), \tag{10}$$

and for $-d < k < +\infty$,

$$W_k(z) = -z \sum_{j=1}^{d} \frac{v_j'(z)}{v_j(z)^{k+1}} = \frac{z}{k}\frac{d}{dz}\left(\sum_{j=1}^{d} v_j(z)^{-k}\right), \tag{11}$$

where $v_1, \ldots, v_d$ *are the large branches.*

(For W_0, the second form is to be taken in the limit sense $k \to 0$.)

Proof. Set $w_n(u) = [z^n]W(z,u)$, the Laurent polynomial that describes the possible altitudes and the number of ways to reach them in n steps. We have $w_0(z) = 1$, $w_1(z) = P(u)$, and $w_{n+1}(z) = P(u)w_n(z)$, so that $w_n(z) = P(u)^n$ for all n. The determination of $W(z,u)$ in (8) follows from

$$\sum_{n \geq 0} P(u)^n z^n = \frac{1}{1 - zP(u)},$$

where the sum converges and represents an analytic function of both arguments for $|z| < 1/P(|u|)$. Observe that the resulting series is entire in z but of the Laurent type in u (it involves arbitrary negative powers of u).

For positive u, the radius of convergence of $W(z,u)$ viewed as a function of z is exactly $1/P(u)$. Also, by dominance of coefficients (one has $V_n \leq P(1)^n$), the radius of convergence of $V(z)$ as a function of z is at least $1/P(1)$. Consider now $|z| < r$, where $r := \frac{1}{2}P(1)^{-1}$. Then, since $1/P(u)$ is continuous and unimodal for $u \in (0, +\infty)$ (where $P''(u) > 0$, so that P is convex) and $1/P(0) = 1/P(\infty) = 0$, there exists an interval (α, β) such that for $\alpha \leq u \leq \beta$, one has $1/P(u) > r$. More generally, by positivity of the coefficients, the function $W(z,u)$ is seen to be analytic in the product domain

$$(z,u) \in \{z \mid |z| < r\} \times \{u \mid \alpha < |u| < \beta\}.$$

Thus, by Cauchy's formula applied to the function $W(z,u)$ (viewed now as a function of u analytic in a crown), one has[2]

$$V(z) = [u^0]W(z,u) = \frac{1}{2i\pi}\int_{|u|=(\alpha+\beta)/2} W(z,u)\,\frac{du}{u}.$$

Take z small enough, so that all the large branches that escape to infinity lie outside of $|u| \leq (\alpha+\beta)/2$ *and* the small branches are all distinct. Then, only the small branches remain inside, and, since there are only simple poles, one has

$$\operatorname*{Res}_{u=u_j}\left(\frac{1}{u(1-zP(u))}\right) = -\frac{1}{zu_jP'(u_j)}. \tag{12}$$

The integration contour is shrunk to 0, which is legitimate since $W(z,u)$ remains $O(1)$, and residues are taken into account. The residue theorem then gives $V(z)$ as

[2] We make use of the conventional notation for coefficients of entire and Laurent series: $[z^n]\sum_n f_n z^n := f_n$.

a sum of residues of the form (12) over all small branches. The formula simplifies to (9) since differentiation of the characteristic equation shows that $P'(u)^{-1} = -z^2u'$ for any branch u.

The same procedure is applicable to

$$W_k(z) \equiv [u^k]W(z,u) = \frac{1}{2i\pi}\int_{|u|=(\alpha+\beta)/2} W(z,u)\,\frac{du}{u^{k+1}}.$$

The integration contour can be shrunk to zero provided the integrand (which is of order u^{c-k-1}) remains bounded as $u \to 0$, which necessitates $k \leq (c-1)$. The result of (10) follows again from a residue calculation involving small branches. (The proof shows the formulæ to be valid in a small enough neighbourhood of the origin. The identities are then *a posteriori* valid as identities between formal (fractional) power series.)

When $k > -d$, which covers the case (11) of an arbitrary positive k, the residue calculation is completed by extending the contour to a large circle at ∞; in this case, the large branches contribute.

The algebraic character of $V(z)$ and the $W_k(z)$ finally results from the well-known fact that algebraic functions are closed under sums, products, and multiplicative inverses. □

The quantity $V(z) \equiv W_0(z)$ is equivalently given as the diagonal of a bivariate rational function,

$$V(z) = \sum_n \left([z^n u^{cn}] \frac{1}{1 - zu^cP(u)} \right) z^n,$$

and as such it must be algebraic: see Pólya's paper [63] of 1921 and [37] for developments regarding diagonals of rational functions.

EXAMPLE 1. *Central binomial and trinomial numbers.* These are perhaps the most famous examples, associated to the sets $\mathcal{S} = \{-1, +1\}$ and $\mathcal{T} = \{-1, 0, +1\}$. The corresponding polynomials are $P^{\mathcal{S}}(u) = u^{-1} + u$ and $P^{\mathcal{T}}(u) = u^{-1} + 1 + u$. In this case, the characteristic curve is of degree 2 and there is only one small branch, namely

$$u_1^{\mathcal{S}}(z) = \frac{1 - \sqrt{1-4z^2}}{2z}, \quad u_1^{\mathcal{T}}(z) = \frac{1 - z - \sqrt{1-2z-3z^2}}{2z}.$$

The algebraic generating functions of bridges are then

$$\begin{aligned} V^{\mathcal{S}}(z) &= \frac{1}{\sqrt{1-4z^2}} = 1 + 2\,z^2 + 6\,z^4 + 20\,z^6 + 70\,z^8 + 252\,z^{10} + \cdots \\ V^{\mathcal{T}}(z) &= \frac{1}{\sqrt{1-2z-3z^2}} = 1 + z + 3\,z^2 + 7\,z^3 + 19\,z^4 + 51\,z^5 + \cdots, \end{aligned}$$

the coefficients being[3] *EIS* **A000984** and *EIS* **A002426**

$$[z^n]V^{\mathcal{S}}(z) = [t^n](1+t^2)^n \equiv \binom{2n}{n}, \qquad [z^n]V^{\mathcal{T}}(z) = [t^n](1+t+t^2)^n.$$

The names of central binomial and trinomial numbers are suggested by the usual expansions of $(1+t^2)^n$ and $(1+t+t^2)^n$:

[3] References to EIS point to Sloane's *Encyclopedia of Integer Sequences* [70], of which a version also exists in print [71].

$$\begin{array}{c} \mathbf{1} \\ 1 \;+\; t^2 \\ 1 \;+\; \mathbf{2t^2} \;+\; t^4 \\ 1 \;+\; 3t^2 \;+\; 3t^4 \;+\; t^6 \\ 1 \;+\; 4t^2 \;+\; \mathbf{6t^4} \;+\; 4t^6 \;+\; t^8 \end{array} \qquad \begin{array}{c} \mathbf{1} \\ 1 \;+\mathbf{1t} \;+t^2 \\ 1 \;+2t \;+\mathbf{3t^2} \;+2t^3 \;+t^4 \\ 1 \;+3t \;+6t^2 \;+\mathbf{7t^3} \;+6t^4 \;+3t^5 \;+t^6 \\ 1 \;+4t \;+10t^2 \;+16t^3 \;+\mathbf{19t^4} \;+16t^5 \;10t^6 \;+4t^7 \;+t^8 \end{array}$$

It is notable that these cases were already considered by Euler [24] who also gave linear recurrences (with polynomial coefficients) satisfied by $V_n^{\mathcal{T}}$. □

2.2. **Meanders and excursions.** In this section, we consider meanders, that is paths that never go below the horizontal axis. The meanders whose final altitude is 0 are called excursions, in accordance with Definition 2, and they turn out to be the objects with the richest combinatorial properties.

We continue with a simple system of paths defined by the set of jumps $\mathcal{S}$, possibly endowed with weights. The new generating functions will again involve the characteristic curve together with its small and large branches. Let now $F_{n,k}$ be the number of meanders of size (i.e., length) n that end at altitude k. The corresponding BGF is

$$F(z,u) := \sum_{n,k} F_{n,k} u^k z^n,$$

which is now an entire series in both z and u. By the combinatorial origin of the problem, $F(z,u)$ is bivariate analytic for $|u| \leq 1$ and $|z| < 1/P(1)$. We also make use of the polynomials $f_n(u)$ that describe the possible positions after n steps and write

$$F(z,u) = \sum_{n\geq 0} f_n(u) z^n = \sum_{k\geq 0} F_k(z) u^k. \tag{13}$$

Combinatorially, the natural decomposition is the one based on the last step added. For the $f_n(u)$, "adding a slice" is translated by the recurrence,

$$f_0(u) = 1, \qquad f_{n+1}(u) = P(u) f_n(u) - \{u^{<0}\} P(u) f_n(u). \tag{14}$$

There, the notation $\{u^{<r}\}g(u)$ means the sum of all the monomials with exponent less than r that appear in the Laurent series $g(u)$:

$$\{u^{<r}\}\left(\sum_{j=-a}^{+\infty} g_j u^j\right) := \sum_{j=-a}^{r-1} g_j u^j. \tag{15}$$

Then, multiplying the terms of the recurrence by z^n and summing yields

$$F(z,u) = 1 + zP(u)F(z,u) - z\{u^{<0}\}\left(P(u)F(z,u)\right), \tag{16}$$

where $\{u^{<0}\}$ is to be understood as applied to the u-expansion of $F(z,u)$ in (13). The relation (16) is the *fundamental functional equation* defining meanders. It reads as follows: "A path is either the empty path or it consists of a step ($zP(u)$ describes the possibilities) added to a path except that the steps that would take the walk below level 0 (the operator $\{u^{<0}\}$) are to be taken out". Now, P involves only a finite number of negative powers, so that

$$F(z,u)(1 - zP(u)) = 1 - z\sum_{k=0}^{c-1} r_k(u) F_k(z), \tag{17}$$

for some Laurent polynomials $r_k(u)$ that are immediately computable from P via (16):

$$r_k(u) := \{u^{<0}\}\left(P(u)u^k\right) \equiv \sum_{j=-c}^{-k-1} p_j u^{j+k}. \tag{18}$$

Theorem 2. *For a simple set of steps, the BGF of meanders (with z marking size and u marking final altitude) relative to a simple set of path $\mathcal{S}$ is algebraic. It is given in terms of the small and large branches of the characteristic curve of $\mathcal{S}$ by*

$$F(z,u) = \frac{\prod_{j=1}^{c}(u-u_j(z))}{u^c(1-zP(u))} = -\frac{1}{p_d z}\prod_{\ell=1}^{d}\frac{1}{(u-v_\ell(z))}. \tag{19}$$

In particular the GF of excursions, $E(z) = F(z,0)$, satisfies

$$E(z) = \frac{(-1)^{c-1}}{p_{-c}z}\prod_{j=1}^{c} u_j(z) = \frac{(-1)^{d-1}}{p_d z}\prod_{\ell=1}^{d}\frac{1}{v_\ell(z)}. \tag{20}$$

Proof. The point is that the fundamental equation in its form (17) looks grossly underdetermined as it involves $(c+1)$ unknown functions; to wit, the bivariate $F(z,u)$ and the univariate $\{F_k(z)\}_{k=0}^{c-1}$. The main idea of a method known as the "*kernel method*" (see also historical notes below) consists in binding z and u in such a way that the left hand side vanishes.

Indeed, substitute in (17) any small branch of the characteristic equation. Take $|z| < 1/P(1)$ and restrict z to a small neighbourhood of the origin in such a way that: (i) all the small branches are distinct; (ii) all the small branches satisfy $|u_j(z)| < 1$. Then the substitution is analytically legitimate and, taking all small branches into account, it provides a system of c equations in the unknown functions $F_0, \ldots, F_{c-1}$:

$$\left\{\begin{array}{lcl} u_1^c - z\sum_{k=0}^{c-1} u_1^c r_k(u_1)F_k & = & 0 \\ \vdots & & \\ u_c^c - z\sum_{k=0}^{c-1} u_c^c r_k(u_c)F_k & = & 0. \end{array}\right. \tag{21}$$

This system is nonsingular for the reason that its determinant is a variant of the Vandermonde determinant and the small branches are clearly all distinct. This observation is enough to justify that each of the F_k is an algebraic function expressible rationally in terms of the algebraic branches u_j.

Instead of pursuing in the direction of determinantal calculations, we make use here of a cute observation of Mireille Bousquet-Mélou (introduced in [13] and employed in the parallel paper [4]). The quantity

$$N(z,u) := u^c - z\sum_{k=0}^{c-1} u^c r_k(u)F_k \tag{22}$$

is by (21) a polynomial in u whose roots are precisely all the u_j. The leading monomial of this polynomial is u^c, so that the polynomial factorizes as

$$N(z,u) = \prod_{j=1}^{c}(u - u_j(z)). \tag{23}$$

Then, the constant term is at the same time the product $(-1)^c u_1 \cdots u_c$ and the quantity $-zp_{-c}F_0$, as is apparent from the definition (22) and the form (18) of the coefficients. The form of F_0 follows.

Finally, the result for the BGF $F(z,u)$ derives from (17) made entire,

$$F(z,u) = \frac{N(z,u)}{u^c(1 - zP(u))},$$

and from the factorization (23). □

An immediate corollary of Theorems 1 and 2 is the generating function of all paths and meanders irrespective of their final altitude.

Corollary 1. *The generating functions of all paths and all meanders are*

$$W(z,1) = \frac{1}{1 - zP(1)}, \quad F(z,1) = \frac{1}{1 - zP(1)} \prod_{j=1}^{c}(1 - u_j(z)) = -\frac{1}{p_d z} \prod_{\ell=1}^{d} \frac{1}{1 - v_\ell(z)}.$$

A somewhat deeper consequence is a direct relation between the GF's of excursions and bridges that obtains by comparing Equations (9) and (20).

Corollary 2. *The generating functions of bridges (V) and excursions (E) are related by*

$$\begin{aligned} V(z) &= 1 + z\frac{d}{dz}(\log E(z)) = 1 + z\frac{E'(z)}{E(z)} \\ E(z) &= \exp\left(\int_0^z (V(t) - 1)\,\frac{dt}{t}\right). \end{aligned}$$

In the same vein, consider paths whose intermediate steps may be negative, but with a final altitude that is ≥ 0. Their BGF is

$$W^+(z,u) := \sum_{k=0}^{\infty} W_k(z) u^k.$$

Then, comparison of the forms involving large branches for W_k and $F(z,u)$ and a trite calculation shows that

$$\begin{aligned} W^+(z,u) &= 1 + z\frac{d}{dz}(\log F(z,u)) \\ F(z,u) &= \exp\left(\int_0^z (W^+(t,u) - 1)\,\frac{dt}{t}\right). \end{aligned}$$

Finally, with $F_k(z)$ being the generating function of meanders that end at altitude k, one has $F_k(z) = [u^k]F(z,u)$. Since $F(z,u)$ is a rational function of u with a simple product expression in terms of the large branches, its expansion with respect to u is easily accessible via a partial fraction decomposition, and one finds:

Corollary 3. *The generating function of meanders terminating at altitude k is*

$$F_k(z) = \frac{1}{p_d z} \sum_{\ell=1}^{d} \xi_\ell v_\ell^{-k-1}, \qquad \xi_\ell := \prod_{j \neq \ell} \frac{1}{v_j - v_\ell}.$$

Some of these relations admit of combinatorial interpretations succinctly discussed in Section 4.1.

EXAMPLE 2. *Ballot problem, Dyck paths, and Motzkin paths.* These are the most famous problems in the area, and they are closely related to Example 1. The ballot problem asks for the probability, in a two candidate election between A and B that eventually results in a tie, of A dominating B throughout the poll. Recording the difference between the scores of A and B as time evolves, we model the problem as the counting of excursions associated with $\mathcal{S} = \{-1, +1\}$. The characteristic curve is the one examined in Example 1 in connection with central binomial coefficients and the GF of excursions is

$$E^{\mathcal{S}}(z) = \frac{1-\sqrt{1-4z^2}}{2z^2} = \sum_{n\geq 0} \frac{1}{n+1}\binom{2n}{n} z^{2n},$$

where the coefficients $\frac{1}{n+1}\binom{2n}{n}$ are the Catalan numbers (*EIS* **A000108**). For $\mathcal{T} = \{-1, 0, +1\}$, one finds similarly

$$E^{\mathcal{T}}(z) = \frac{1-z-\sqrt{1-2z-3z^2}}{2z^2} = \sum_{n\geq 0} M_n z^n,$$

where the coefficients are the Motzkin numbers (*EIS* **A001006**). □

EXAMPLE 3. *Łukasiewicz paths and tree codes.* Consider generally a finite set Ω that contains -1 as single negative value. The corresponding paths are known as Łukasiewicz paths. Set $\phi(u) := uP(u)$, which is a polynomial. There is only one small branch satisfying

$$u_1(z) = z\phi(u_1(z)), \tag{24}$$

and the GF of excursions is $\frac{1}{zp_{-1}}u_1(z)$. Łukasiewicz paths of type Ω encode trees whose node degrees are constrained to lie in $1+\Omega$, this by virtue of a well-known correspondence [52, Chap. 11]. (Traverse the tree in preorder and output a step of $d-1$ when a node of outdegree d is encountered.) In this way, it is seen that Equation (24) gives the GF of trees counted according to the number of their nodes, an otherwise classical result [55]. By Lagrange inversion, the number of trees comprised of n nodes is

$$T_n = \frac{1}{n}[w^{n-1}]\phi(w)^n,$$

where ϕ can be directly interpreted as the characteristic polynomial of the allowed node (out)degrees. □

EXAMPLE 4. *Walks with steps in* $\{-2, -1, 0, +1, +2\}$. This is our first example involving inherently more than one branch. The characteristic equation is

$$u^2 - z(1+u+u^2+u^3+u^4) = 0.$$

The two small branches are conjugate and given by

$$\begin{aligned} u_1(z) &= +z^{1/2} + \tfrac{1}{2}z + \tfrac{5}{8}z^{3/2} + z^2 + \tfrac{231}{128}z^{5/2} + 3z^6 + \cdots \\ u_2(z) &= -z^{1/2} + \tfrac{1}{2}z - \tfrac{5}{8}z^{3/2} + z^2 - \tfrac{231}{128}z^{5/2} + 3z^6 + \cdots . \end{aligned}$$

Then, by (20), the first few terms of $E(z)$ are easily determined as

$$E(z) = -\frac{1}{z}u_1(z)u_2(z) = 1 + z + 3z^2 + 9z^3 + 32z^4 + 120z^5 + 473z^6 + 1925z^7 + \cdots .$$

Similarly, for meanders, one has

$$F(z,1) = \frac{(1-u_1(z))(1-u_2(z))}{1-5z} = 1+3z+12z^2+51z^3+226z^4+1025z^5+4724z^6+\cdots .$$

It is then a natural question to ask for an equation satisfied directly by $E(z)$ or $F(z,1)$. Regarding excursions, an equation may be obtained by elimination of u_1, u_2 from the system

$$zE + u_1u_2 = 0, \quad u_1^2 - z(1 + u_1^2 + u_1^3 + u_1^4) = 0, \quad u_2^2 - z(1 + u_2u_2^2 + u_2^3 + u_2^4) = 0.$$

Either resultants or Gröbner bases do the job. For instance, resultants give a polynomial equation of degree 12 satisfied by $E(z)$. The polynomial factorizes (this is expected as we did not impose conditions like $u_1 \neq u_2$ in the process). Eventually, it is found that $E(z)$ satisfies a polynomial equation of degree 4:

$$z^4y^4 - z^2(1+z)y^3 + z(2+z)y^2 - (1+z)y + 1 = 0. \tag{25}$$

We shall examine shortly a much better way to perform such computations. □

EXAMPLE 5. *Duchon's clubs and underdiagonal paths.* The following problem[4] was considered by Duchon [22] (under a different formulation): *A club opens in the evening and closes in the morning. People arrive by pairs and leave in threesomes. What is the possible number of scenarios from dusk to dawn as seen from the club's entry?* For instance, an event may be +2 (two enter), +2 (two more enter), −3 (three leave), +2 (two, again arrive), −3 (and the club closes). Naturally the population inside the club is never negative and a business night starts with the empty club and ends with the empty club. The generalized problem then calls for the number of excursions with step set $\{-c, d\}$ (where Duchon's case is $\widehat{\mathcal{S}} = \{-3, +2\}$ or, equivalently by time reversal, $\mathcal{S} = \{-2, +3\}$). We assume here without loss of generality that c and d are coprime integers, so that the system of paths is reduced.

The characteristic polynomial is $P(u) = u^{-c} + u^d$ and the kernel equation is equivalent to

$$u^c = z(1 + u^e) \quad \text{with} \quad e = c + d.$$

Thus, the period is $e = c + d$ and the horizontal axis is only touched at places that are a multiple of e. Set $z = t^c$, where t is a local uniformizing parameter at 0. Then, the quantity $y(t) := u_1(t^c)$ satisfies the equation $y = t(1 + y^e)^{1/c}$, which is Lagrangean. By Lagrange inversion [42], one finds

$$y(t) = \sum_{n\geq 1} \frac{1}{n}\binom{n/c}{(n-1)/e} t^n. \tag{26}$$

[4]After this paper had been submitted, Christian Krattenthaler pointed us to Ref. [68] by Masako Sato, dating from 1989. In that paper, Sato derives directly our equation (27) by matrix generating function methods and provides valuable additional results regarding underdiagonal paths in a strip.

(By convention, $\binom{a}{b} = 0$ if b is nonintegral.) Let ω be a primitive cth root of unity; then all the branches admit an expansion similar to $y(z)$. Indeed, by conjugacy, one has

$$u_{j+1}(t^c) = y(\omega^j t) = \sum_{n\geq 1} y_n \omega^{nj} t^n,$$

where $y_n = [t^n]y(t)$ is given by (26). Then, the number of excursions is a convolution:

$$(-1)^{c-1} E_n = \sum_{n_1+\cdots+n_c=c(n+1)} y_{n_1} y_{n_2} \cdots y_{n_c} \omega^{0n_1+1n_2+\cdots+(c-1)n_c}.$$

It can be checked that E_n is automatically zero unless $n \equiv 0 \pmod e$ (see also the discussion on periodicities in Section 3.3 below). In summary, taking ω any primitive cth root of unity, and setting $n_j = 1 + e\nu_j$, $n = e\nu$, we find

(27)
$$E_{e\nu} = \sum_{\nu_1+\cdots\nu_c=c\nu} \frac{1}{1+\nu_1 e}\binom{(1+\nu_1 e)/c}{\nu_1} \cdots \frac{1}{1+\nu_c e}\binom{(1+\nu_c e)/c}{\nu_c} \omega^{0\nu_1+1\nu_2+(c-1)\nu_c}.$$

In particular, for $c = 1$, no summation is needed and

$$\frac{1}{1+ne}\binom{1+ne}{n}$$

gives the number of excursions of length n and type $\{-1, e-1\}$, which is also the number of e-ary trees having n internal nodes (Example 3). If $c = 2$ the formula (27) yields a single convolution. For $\mathcal{S} = \{-2, 3\}$, the result is

$$E_{5n} = \sum_{\nu=0}^{2n} \frac{(-1)^\nu}{1+5\nu}\binom{(1+5\nu)/2}{\nu} \frac{1}{1+5(2n-\nu)}\binom{(1+5(2n-\nu))/2}{2n-\nu},$$

to be compared to

$$E_{5n} = \sum_{i=0}^{n} \frac{1}{5n+i+1}\binom{5n+1}{n-i}\binom{5n+2i}{i}, \tag{28}$$

which Duchon obtained from quite specific series manipulations. In general if the jump in the negative direction is $-c$, formula (27) is a $(c-1)$-fold convolution of binomial coefficients.

Duchon's clubs can also be interpreted as *underdiagonal paths*. Consider paths in the $\mathbb{Z}_{\geq 0} \times \mathbb{Z}_{\geq 0}$ lattice whose allowed steps are of type either *East* (horizontal) or *North* (vertical), with a straight line barrier Δ. It is assumed that Δ passes through the origin and has a rational slope, $\frac{p}{q} \leq 1$. The number of ways $N_{m,n}$ of reaching point (m, n) by North and East steps then satisfies a recurrence of the same type as Pascal's triangle but with boundary conditions. For instance, the case of slope 1 gives rise to the original formulation [54] of the *ballot problem* (Example 2).

If one measures at each step of a path the vertical distance to Δ, then, this distance can only evolve by $+\frac{p}{q}$ for a horizontal step and -1 for a vertical step. Thus, up to rescaling, such an underdiagonal path is equivalent to a Duchon path of type $\{-q, +p\}$. The numbers $N_{m,n}$ are then amenable to the analysis of the paper since their determination is equivalent to counting meanders and excursions. For instance, here is a table of values for slope $\frac{2}{3}$:

									377	1144
								136	377	767
						23	66	136	241	390
					9	23	43	70	105	149
			2	5	9	14	20	27	35	44
		1	2	3	4	5	6	7	8	9
1	1	1	1	1	1	1	1	1	1	1

The sequence of numbers in this array that correspond to the number of ways of touching the boundary line is (*EIS* **A060941**)

$$1,\ 2,\ 23,\ 377,\ 7229,\ 151491,\ 3361598,\ 77635093,\ 1846620581,\ \cdots$$

which precisely coincides with the sequence of Duchon numbers, $\{E_{5n}\}_{n\geq 0}$, in (28).

Related enumerative results have been obtained by Durand [23] in the context of the "klam" recurrence that arises in complexity theory. Mohanty [57, p. 22] even quotes results of Takács relative to underdiagonal paths under a line of arbitrary slope. □

As the last example shows, the decomposability afforded by the kernel method provides a grasp on the structural complexity of summatory formulæ expressing the number of walks, excursions, etc. Following Comtet [15, p. 216], we observe that the "rank" (defined as the minimal number of summations) of the excursion formula in the general case is at most $c(q-1)-1$ if $P(u)$ comprises q terms. For instance, Catalan numbers ($(c,q)=(1,2)$) are of rank 0, Motzkin numbers ($(c,q)=(1,3)$) and the Duchon numbers E_n of (28) (having $(c,q)=(2,2)$) are of rank 1, etc.

Some origins of the kernel method. What we named here the "kernel method" has been part of the folklore of combinatorialists for some time. Earlier references usually deal with the case of a functional equation of the form

$$K(z,u)F(z,u)=A(z,u)+B(z,u)G(z)$$

(with F,G the unknown functions), when there is only one small branch, u_1, such that $K(z,u_1(z))=0$. In that case, a single substitution does the job, and $G(z)=-A(z,u_1)/B(z,u_1)$. One clear source of this is the exercise section of the first edition (in 1968) of Knuth's book [46]: the detailed solution to Exercise 2.2.1–4 (see [46, p.536–537] and also Ex. 2.2.1.11) presents a "new method for solving the ballot problem", for which the characteristic equation is quadratic. See also Odlyzko's splendid survey [61, Sec. 15.4] for a discussion of a pebbling game and Prodinger's recent note [64] for an original application to a quadratic problem arising fom queueing theory.

The kernel method in its more general version was used recently in a few unpublished works by the authors, including a systematization to directed lattice paths by Banderier in his memoir [2]. Independent combinatorial developments at the end of the last century are due to Bousquet-Mélou and Petkošek whose recent paper offers a penetrating perspective on the subject of multidimensional walks, recurrences, and kernels [13]. In fact, as indicated earlier, a remark of Bousquet-Mélou has been used to simplify our proof of Theorem 2 (see also [4] for another application).

That probabilists had known a lot since the early 1950's regarding related questions is manifest upon reading Chapter XII of Fellers' book [28]. It appears that our presentation parallels in some ways what is obtained by the famous Wiener-Hopf approach: refer in particular to the example on bounded arithmetic distributions in [28, p. 407–408]. Such techniques prove in turn valuable in the theory of queueing systems: see, e.g., Robert's book [66] for an account. The synthesis by Fayolle, Iasnogorodski, and Malyshev [26] exposes the deep ramifications of the theory in the harder case of walks in a quarter plane *not* satisfying directedness restrictiction (thus, a "pure" 2-dimensional problem), but their methods only apply to nearest-neighbour moves. The book [26] itself draws some of its inspiration from the early paper [25] where a sophisticated use of the kernel method already plays a central rôle (amongst other techniques like conjugacy and Riemann–Hilbert problems); see also the references to Flatto and Malyshev's works in [61, p. 1208] and the historical comments in [26, p. VII-XI].

2.3. **Computational aspects.** We discuss now a way to determine directly the equations satisfied by the algebraic functions encountered so far. Because of Corollary 2, we know that bridges and excursions are tightly coupled, and the case of excursions will be detailed here.

It is assumed that the characteristic polynomial $P(u)$ is fixed. Then, what is needed in view of Theorem 2 is the equation satisfied by the product $Y = u_1 \cdots u_c$ of c distinct roots of a polynomial of degree $c+d$. As roots are in general "undistinguishable", we expect a polynomial of degree $\binom{c+d}{c}$ to cancel Y.

Take a polynomial $Q(u)$ of degree e in $\mathbb{C}(z)[u]$ normalized by $Q(0) = 1$ and assume it has distinct roots $u_1, \ldots, u_e$. For us, $e = c + d$, and

$$Q(u) = -\frac{1}{zp_{-c}}\left(u^c - zu^c P(u)\right),$$

yet another reformulation of the kernel. We first develop the computational process when $c = 2$, so that the equation for $Y = u_1 u_2$ with u_1, u_2 two distinct roots of Q is sought. Write α, α' for generic roots of Q. Then, since $Q(0) = 1$, one has

$$Q(u) = \prod_{\alpha}\left(1 - \frac{u}{\alpha}\right),$$

while what we need to determine is

$$R(u) = \prod_{\{\alpha,\alpha'\}}\left(1 - \frac{u}{\alpha\alpha'}\right).$$

(A sum or product over $\{\alpha, \alpha'\}$ means a sum or product over all unordered pairs of *distinct* elements.) Now, take logarithms. One has

$$\begin{aligned} \log\left(\frac{1}{Q(u)}\right) &= \sum_{n\geq 1} S_n \frac{u^n}{n} \quad \text{with} \quad S_n := \sum_{\alpha} \frac{1}{\alpha^n} \\ \log\left(\frac{1}{R(u)}\right) &= \sum_{n\geq 1} S_n^{(2)} \frac{u^n}{n} \quad \text{with} \quad S_n^{(2)} := \sum_{\{\alpha,\alpha'\}} \frac{1}{\alpha^n \alpha'^n}. \end{aligned}$$

Then, a simple combinatorial reasoning shows that

$$\sum_{\{\alpha,\alpha'\}} \frac{1}{\alpha^n\alpha'^n} = \frac{1}{2}\sum_{(\alpha,\alpha')} \frac{1}{\alpha^n\alpha'^n} - \frac{1}{2}\sum_{\alpha}\frac{1}{\alpha^{2n}},$$

so that

$$S_n^{(2)} = \frac{1}{2}S_n^2 - \frac{1}{2}S_{2n}. \tag{29}$$

The degree of R is $\delta := \binom{e}{2}$ *a priori*, and R can be recovered from the formula ("I am always the exponential of my logarithm!")

$$R(u) := \{u^{\leq\delta}\}\left[\exp\left(-\sum_{n=1}^{\delta}\frac{1}{2}(S_n^2 - S_{2n})\frac{u^n}{n}\right)\right], \tag{30}$$

where $\{u^{\leq\delta}\} f$ means the truncation of the series expansion of f with all terms of degree $\leq \delta$ included (see the analogous notation (15)).

The general formulæ for $c > 2$ are easily found from the usual relations between elementary and power sum symmetric functions. Set $x_j = \alpha_j^{-n}$. What is sought is plainly a formula expressing the sum Φ_c of all products $x_{j_1}\cdots x_{j_c}$ taken over all distinct *subsets* $\{j_1, \ldots, j_c\}$ when the power sums $s_k := \sum_j x_j^k$ are known. Then, one has (by exponentials of logarithms again)

$$\Phi_c = [t^c]\prod_j(1 + tx_j) = [t^c]\exp\left(\sum_{k\geq 1}(-1)^{k-1}s_k\frac{t^k}{k}\right). \tag{31}$$

Thus, Φ_c is a computable polynomial in $s_1, \ldots, s_c$, obtained from extracting the coefficient $[t^c]$ in the exponential form of (31) that we write as $\Phi_c(s_1, \ldots, s_c)$. Define finally

$$S_n^{(c)} := \sum_{\{j_1,\ldots,j_c\}} u_{j_1}^{-n}\cdots u_{j_c}^{-n},$$

the sum being on all subsets of c elements. Then we have

$$S_n^{(c)} = \Phi_c(S_n, S_{2n}, \cdots, S_{cn}).$$

For instance, the formulæ analogous to (29) for $c = 3, 4$ are found to be

$$\begin{aligned} S_n^{(3)} &= \tfrac{1}{6}S_n^3 - \tfrac{1}{2}S_nS_{2n} + \tfrac{1}{3}S_{3n} \\ S_n^{(4)} &= \tfrac{1}{24}S_n^4 - \tfrac{1}{4}S_n^2S_{2n} + \tfrac{1}{3}S_nS_{3n} + \tfrac{1}{8}S_{2n}^2 - \tfrac{1}{4}S_{4n}. \end{aligned} \tag{32}$$

These considerations give rise to a simple algorithm for computing the polynomial cancelled by the product of all small branches.

Algorithm R. *Computes the polynomial $R(u) \in \mathbb{C}(z)[u]$ of degree $\delta = \binom{e}{c}$ such that $R(Y) = 0$, where $Y = u_1\cdots u_c = (-1)^{c-1}zp_cE(z)$ is the product of all small branches of the characteristic curve.* The input is the characteristic polynomial of steps, $P(u)$.

1. Set up the symbolic formulæ of type (29) and (32) appropriate for the given value of c. To this effect, perform the symbolic expansion of (31) with $\Phi_c(s_1, \ldots, s_c)$ denoting the coefficient of t^c in the exponential form.
2. Take the normalized kernel $Q(u) = (-zp_c)^{-1}(u^c - zu^cP(u))$. Set $\delta = \binom{e}{c}$ and determine the expansion

$$\log\left(\frac{1}{Q(u)}\right) = \sum_{n=1}^{c\delta} S_n\frac{u^n}{n} + O(u^{c\delta+1}).$$

3. Recover $R(u)$ from the truncated series

$$R(u) := \{u^{\leq\delta}\}\left[\exp\left(-\sum_{n=1}^{\delta}\Phi_c(S_n, S_{2n}, \ldots, S_{cn})\frac{u^n}{n}\right)\right].$$

Half a dozen instructions in a symbolic manipulation language are sufficient to translate the algorithm. In contrast to Gröbner basis or resultant calculations, the process is efficient, whenever the degree of the result remains reasonable. For instance, we could successfully determine polynomials R of degree $45 = \binom{10}{2}$ in a matter of seconds on a machine with a 500MHz clock.

On coefficients of algebraic functions. As it is well known [14], any algebraic function $f(z)$ satisfies a linear differential equation $L(f) = 0$ with coefficients that are rational functions of the variable. This in turn translates into a linear recurrence with polynomial coefficients in n for the quantities $[z^n]f$. Thus, the coefficient of index n of any algebraic function is computable in a number of operations that is linear in n. (The procedure is implemented in Salvy and Zimmermann's `Gfun` package [67].) This remark applies to all the generating functions considered in this paper. For instance, the excursion generating function $E(z)$ corresponding to the set of jumps $\{-2, -1, 0, +1, +2\}$ (Example 4) satisfies an inhomogeneous differential equation of order 3

$$(33) \quad z^3(5z+4)(5z+1)(z-1)^2(5z-1)^2\frac{d^3E}{dz^3} + \cdots + (-100z^2+56z-4) = 0,$$

and its coefficients can be obtained from a recurrence of order 6,

$$(34) \quad 2(n+7)(n+8)(2n+13)E_{n+6} + \cdots + 625(n+1)(n+2)(n+3)E_n = 0.$$

3. Singular structures

We now examine paths, bridges, meanders and excursions under the angle of asymptotics. As is well known, the asymptotic behaviour of counts is closely related to the singular structure of the corresponding generating functions [34, 61]. Thanks to the factorizations afforded by the kernel method, the singular forms of intervening generating functions become manageable. This part of the analysis makes use of global properties of branches followed by local analysis in the vicinity of a quantity called the "structural radius" ρ.

Lemma 1. *Let $P(u)$ be the polynomial associated to the steps of a simple walk. Then, there exists a unique number τ, called the* structural constant, *such that*

$$P'(\tau) = 0, \quad \tau > 0.$$

The structural radius *is by definition the quantity*

$$\rho := \frac{1}{P(\tau)}.$$

Proof. Differentiating twice P as given in (3), we see that $P''(x) > 0$ for all $x > 0$. Thus, the real function $x \mapsto P(x)$ is strictly convex. Since it satisfies $P(0) = P(+\infty) = +\infty$, it must have a unique positive minimum attained at some τ, and $P'(\tau) = 0$. □

Structural constants *a priori* live in a field of degree $e := c + d$ over the base field of weights. However, for symmetric walks ($P(u) = P(u^{-1})$), they automatically reduce to the value $\tau = 1$ and ρ becomes automatically a member of the field of coefficients of P.

In Section 2, we have defined the principal branch $u_1(z)$ near the origin by means of its expansion at 0. We show here that this branch satisfies a useful domination property for $0 \le z \le \rho$. Cf. Figure 4 for an illustration.

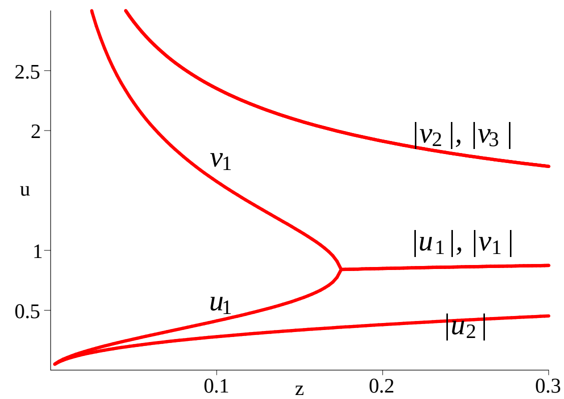


FIGURE 4. A rendering of the modulus of the five branches of the characteristic curve in the example of Figure 3 illustrates the domination properties of the principal small and large branches.

Lemma 2. *For an aperiodic walk, the principal small branch $u_1(z)$ is analytic on the open interval $z \in (0, \rho)$. It dominates strictly in modulus all the other small branches, $u_2(z), \ldots, u_c(z)$, throughout the half-closed interval $z \in (0, \rho]$.*

Proof. By the discussion of Lemma 1, the function $1/P(z)$ is continuously increasing for $z \in [0, \tau]$. Hence the equation (in u) $z = 1/P(u)$ admits a unique positive solution, say $u^+(z)$, that is less than τ when $z \in [0, \rho]$. This positive solution $u^+(z)$ must coincide with the branch u_1 at 0^+ (since the expansions at 0^+ are the same). Also, the analytic version of the implicit function theorem guarantees that the positive solution $u^+(z)$ remains analytic all along $z \in (0, \rho)$, so that the principal small branch u_1 and the positive solution u^+ must coincide throughout this interval. Consequently, u_1 (originally only defined near 0^+) increases from 0 to τ as ρ increases from 0 to ρ.

Next, a general fact about polynomials with positive coefficients enters the game: if $P(u)$ is aperiodic, then one has for positive r

$$|P(re^{i\theta})| < P(r) \quad \text{for all } \theta \not\equiv 0 \pmod{2\pi}, \tag{35}$$

as seen from the strong form of the triangle inequality. Fix $z = x$, with x real positive and $x < \rho$, and let w be an arbitrary solution of the kernel equation $1 - xP(w) = 0$ that is at most τ in modulus and *not* equal to $u_1(x)$ (i.e., not real and positive). Then, one has by the strict inequality in (35)

$$x = \frac{1}{P(u_1(x))} = \frac{1}{P(w)} > \frac{1}{P(|w|)},$$

which implies $|w| < u_1(x)$ since $1/P$ is increasing in the region considered, $[0, \tau]$. Thus, near 0^+ and since the nonprincipal small branches $u_2, \ldots, u_c$ are majorized by τ in modulus (they tend to 0), they must satisfy $|u_j(x)| < u_1(x)$. Additionally, the domination property cannot cease to hold on $(0, \rho)$: by continuity of the modulus of any branch, this would imply that $u_1(x)$ itself reaches the value τ for some $x < \rho$, yielding a clear contradiction. Domination must finally continue to hold at ρ, since otherwise, there would be a contradiction with the strong triangle inequality (35). □

Stronger domination properties are in fact derivable from similar uses of the strong triangle inequality, under the aperiodicity condition (see also [3] for details). For $|z| \le \rho$, one has: $|u_j(z)| < u_1(|z|)$ for $j = 2, \ldots, c$; also, $|u_1(z)| < |v_1(z)|$ safe at $z = \rho$. Simply put, the principal small branch u_1 is the "largest" of all the small branches.

In Section 4, it will also prove handy to have available the corresponding properties of large branches. For instance, the principal large branch, v_1, is in a similar sense the smallest of all large branches. Generally, the domination properties of large branches are counterparts of those of small branches, as can be seen by mimicking the arguments. Alternatively, one can introduce duality: If $P(u)$ is a Laurent polynomial, then $\widetilde{P}(u) = P(u^{-1})$ is called its dual. It is then easy to see that the small and large branches, $\widetilde{u}_j$ and $\widetilde{v}_\ell$ of the dual are respectively the inverses of the large and small branches of the primal: $\widetilde{u}_j v_j = 1$ and $\widetilde{v}_\ell u_\ell = 1$. Duality thus exchanges small and large branches. (Combinatorially, duality may be realized either as a symmetry along the horizontal axis applied to steps, or by the time-reversal transformation that changes a path into another path obtained by reading steps backwards.)

3.1. **Bridges and excursions.** We first address the important problem of estimating the numbers of bridges and excursions. The discussion makes use of the assumption that the walk is reduced and aperiodic.

Theorem 3. *Consider a simple system of walks that is aperiodic. Let τ be the structural constant determined by $P'(\tau) = 0$, $\tau > 0$. The number of bridges of size n admits a complete asymptotic expansion*

$$V_n \sim \lambda_0 \frac{P(\tau)^n}{\sqrt{2\pi n}} \left(1 + \frac{a_1}{n} + \frac{a_2}{n^2} + \cdots\right), \quad \lambda_0 = \frac{1}{\tau}\sqrt{\frac{P(\tau)}{P''(\tau)}}. \tag{36}$$

The number of excursions of size n satisfies

$$E_n \sim \mu_0 \frac{P(\tau)^n}{2\sqrt{\pi n^3}} \left(1 + \frac{b_1}{n} + \frac{b_2}{n^2} + \cdots\right), \tag{37}$$

where (the u_j are the small branches, with u_1 the principal branch)

$$\mu_0 = \frac{(-1)^{c-1}}{p_{-c}} \sqrt{\frac{2P(\tau)^3}{P''(\tau)}} Y_1(\rho), \quad Y_1(z) := \prod_{j=2}^{c} u_j(z), \quad \rho := \frac{1}{P(\tau)}. \tag{38}$$

By Lemma 2, the constant $Y_1(\rho)$ is equivalently characterized as

$$Y_1(\rho) = \prod_{|v|<\tau,\ P(v)=\rho^{-1}} v.$$

Proof. The result for bridges is known as it is equivalent to the local limit theorem for sums of discrete random variables [40, Chapter 9], of which the first proof goes back to Laplace[5] in [51]. For completeness, we briefly sketch the argument here.

[5]Quite remarkably, in his *Théorie analytique des probabilités*, in 1812. Laplace expresses the problem as a Cauchy coefficient formula presented by its Fourier series counterpart (analytic functions are not yet invented by Cauchy!) and proceeds with a saddle point argument expressed

Start from the fact that the number of bridges of length n is $[u^0]P(u)^n$. By Cauchy's coefficient formula, one has

$$V_n = \frac{1}{2i\pi}\int_\gamma P(u)^n \frac{du}{u},$$

where the contour γ is any positively oriented loop about the origin. The positive real point τ is a simple saddle point of $P(u)$ (hence of $P(u)^n$), so that the choice of the circle $|u| = \tau$ as integration contour suggests itself by the saddle-point method [16]. By the aperiodicity condition, $P(u)$ is uniquely maximal in modulus along the contour at $u = \tau$; see (4). Therefore, the following saddle-point approximations are justified:

$$\begin{aligned} V_n &= \frac{1}{2i\pi}\int_{|u|=\tau} P(u)^n \frac{du}{u} \\ &\sim \frac{1}{2i\pi}\int_{\tau e^{-i\epsilon}}^{\tau e^{+i\epsilon}} \exp\left(n\left(\log P(\tau) + \frac{1}{2}\frac{P''(\tau)}{P(\tau)}(u-\tau)^2 + O((u-\tau)^3)\right)\right)\frac{du}{u} \\ &\sim \frac{P(\tau)^n}{2\pi\tau}\int_{-\infty}^{+\infty} e^{-nht^2/2}\,dt \;=\; \frac{P(\tau)^n}{\tau\sqrt{2\pi n h}}, \qquad h = \frac{P''(\tau)}{P(\tau)}. \end{aligned}$$

By the usual process, the contribution is first localized near τ, taking for instance $\epsilon = (\log n)/\sqrt{n}$, and local expansions are applied; then the contour is extended back to yield a complete Gaussian integral. This streamlined version of the method is then extended to a full asymptotic expansion in the usual way [43, p. 419], so that (36) results.

The saddle point method thus provides an easy access to the enumeration of bridges. This gives indirectly valuable information on the small branches that can be translated into the singular structure of the GF $V(z)$. First, the relation that determines the branches of the characteristic curve can be put under the form

$$z = \frac{1}{P(u)}. \tag{39}$$

This shows that a branch can become infinite only at $z = 0$; in fact the corresponding solutions give rise precisely to the large branches $v_1, \ldots, v_d$. By general principles (the inverse of an analytic function at a point where the derivative is nonzero is analytic), the relation (39) is invertible analytically in a the neighbourhood of any point v such that $P'(v) \neq 0$. Accordingly, a singularity (in the sense of analytic functions) *must* occur at any value ζ such that $P'(\zeta) = 0$.

At $u = \tau$, with τ the structural constant, one has $P'(\tau) = 0$ by construction, while $P''(\tau) > 0$. Then, the local form of (39), reads

$$z = \rho - \frac{1}{2}P''(\tau)(u-\tau)^2 + O((u-\tau)^3). \quad \rho := \frac{1}{P(\tau)}. \tag{40}$$

This is readily inverted, yielding two local solutions

$$u(z) = \tau \pm \sqrt{2\frac{P(\tau)}{P''(\tau)}}\sqrt{1 - z/\rho} + \cdots \qquad (z \to \rho^-). \tag{41}$$

---

as an application of the "Laplace method" that was specifically developed for that occasion (saddle point integrals will only emerge half-a-century later!).

In particular, the principal branch $u_1(z)$ has a square root singularity; it takes as value the structural constant τ at the place

$$\rho = \frac{1}{P(\tau)}.$$

and the $-\sqrt{}$ determination must be adopted in (41) since $u_1(z)$ increases as $z \to \rho^-$:

$$u_1(z) = \tau - \sqrt{2\frac{P(\tau)}{P''(\tau)}}\sqrt{1 - z/\rho} + \cdots \qquad (z \to \rho^-). \tag{42}$$

Next, for $z \neq 0$, all singularities of the solutions of (39), since they correspond to finite values of u, can only be finite branch points ζ with a local expansion of the form $a_0 + b_0(z - \zeta)^{1/r}$ for some ramification index $r > 1$. (This is easily seen directly by a suitable generalization of (40) and(41) upon taking into accont the first nonzero derivative of $1/P$).

We can now confront the result of (42) with the the saddle point estimation (36), remembering that one has by (9)

$$V(z) = z\frac{d}{dz}\log Y(z), \qquad Y(z) := (u_1(z) \cdots u_c(z)).$$

First, $Y(z)$ that is analytic near 0 must remain analytic throughout the disk $|z| < \rho$, since otherwise $V(z)$ would be singular for some value inside the disk and this would contradict the asymptotic growth (36) that is of type $P(\tau)^n$ for V_n. Next, $Y(z)$ cannot have any (algebraic) singularity other than $z = \rho$ on the circle $|z| = \rho$, since, by singularity analysis[6], this would entail the presence of oscillating terms in the asymptotic expansion of V_n, again contradicting (36). Also, $Y(z)$ can only have a branch point of ramification index $r = 2$ at $z = \rho$, since otherwise some term of the form $n^{-1+1/r}$ would have been present in the expansion of V_n. Finally, the deflated product $Y_1(z) = u_2(z) \cdots u_c(z)$ must be analytic at ρ since otherwise, being capable only of having a branch point with ramification index 2, one would reach a contradiction regarding the leading coefficient of V_n (as checked from comparing (36) against the consequences of (42) on coefficients).

In other words, this sequence of indirect arguments shows the following[7]: *The product of all the nonprincipal small branches*

$$Y_1(z) = u_2(z) \cdots u_c(z) \tag{43}$$

is analytic at all points of the closed disk $|z| \leq \rho$.

It is now an easy matter to complete the estimate of the number of excursions by singularity analysis applied to (20) in Theorem 2. The unique dominant singularity of $E(z)$ must be at $z = \rho$ where the local expansion (42) gives

$$E(z) \sim E(\rho) - \mu_0\sqrt{1 - z/\rho}, \qquad \mu_0 = \frac{(-1)^{c-1}}{p_{-c}\rho}Y_1(\rho)\sqrt{2\frac{P(\tau)}{P''(\tau)}},$$

---

[6]Singularity analysis [34, 61] allows us to transfer a singular element of the form $(1 - z/\alpha)^\kappa$in the expansion of a function $f(z)$ at a singularity α into a corresponding asymptotic element of the form $\alpha^{-n}n^{-\kappa-1}/\Gamma(-\kappa)$ in the expansion of the coefficient $[z^n]f(z)$ at infinity. It is applicable unconditionally to algebraic functions.

[7]An alternative argument based on the refinement of domination relations evoked after the proof of Lemma 2 is possible; see Banderier's thesis [3] for details.

with Y_1 given by (43). A full expansion of $u_1(z)$ in powers of $(1-z/\rho)^{1/2}$ being available, and $Y_1(z)$ being analytic on the whole of $|z| \le \rho$, the proof of (37) is at last completed. □

EXAMPLE 6. *Asymptotics of tree codes.* The case of walks with only one type of descending step equal to -1 corresponds to tree codes, as discussed in Example 3. In this very special case, there is only one small branch, and the GF of excursions is $E(z) = u_1(z)/(p_{1-}z)$. For aperiodic walks, the result (37) of Theorem 3, or plainly the estimate (41), gives us

$$\tag{44} \begin{aligned} \tau &: \quad P'(\tau) = 0 \\ E_n &\sim \frac{1}{p_{-1}} \frac{1}{\sqrt{2\pi n^3}} \sqrt{\frac{P(\tau)^3}{P''(\tau)}}\, P(\tau)^n. \end{aligned}$$

In terms of trees, the principal branch $u_1(z)$ is precisely the GF of trees corresponding to the degree set $1+\mathcal{S}$ with generating polynomial $\phi(u) := uP(u)$ and one has $T(z) = p_{-1}zE(z) = u_1(z)$. The estimate (44) then coincides with the well-known asymptotic estimate of the number T_n of trees of size n,

$$\tag{45} \begin{aligned} \tau &: \quad \phi(\tau) - \tau\phi'(\tau) = 0 \\ T_n &\sim \frac{1}{\sqrt{2\pi n^3}} \sqrt{\frac{\phi(\tau)}{\phi''(\tau)}} \left(\frac{\phi(\tau)}{\tau}\right)^n, \end{aligned}$$

which was first discovered by Meir and Moon [55]. □

As soon as $c > 1$, there are several small branches, and, in this case, the algebraic constant $Y_1(\rho)$ intervenes. Numerically, this constant can be determined easily as it only involves the product of the small solutions to the kernel equation taken at $z = \rho$. Algebraically, since $Y_1(\rho)$ is the product of $c-1$ solutions to an algebraic equation of degree $c+d$, it is an algebraic number of degree at most $\binom{c+d}{c-1}$ over $\mathbb{Q}(\rho) \equiv \mathbb{Q}(\tau)$ that is computable by the techniques of Section 2.3 (upon changing c to $c-1$ in Algorithm R). However, since τ is a double root of the kernel equation instantiated at $z = \rho$, further simplifications accrue. This explains that constants involving radicals are often to be observed when analysing problems of relatively low "complexity". The next example is typical of this state of affairs.

EXAMPLE 7. *Asymptotics of the* $\{-2,-1,0,1,2\}$*-excursions.* The walk introduced in Example 4 is symmetric, and like for any symmetric walk system, the structural constant is equal to 1 while the structural radius is the rational number, $\rho = 1/P(1) = \frac{1}{5}$. The product of the nonprincipal small branches at ρ reduces to $u_2(\rho)$. This quantity is *a priori* one of the roots of an equation of degree 4 (Equation (25) instantiated at $z = \rho$), but since this equation has already $\tau = 1$ as a double root, the equation satisfied by $u_2(\rho)$ is in fact of degree 2 (it is $u^2 + 3u + 1 = 0$) so that

$$u_2(\rho) = -\frac{3}{2} + \frac{1}{2}\sqrt{5},$$

and this quantity is precisely $Y_1(\rho)$ of (38). Thus, we can conclude and get easily

$$E_n = \frac{5}{4}(3-\sqrt{5})\frac{5^n}{\sqrt{\pi n^3}}\left(1 + O(\frac{1}{n})\right).$$

The quality of the asymptotic approximation provided by the first term is 11% when $n = 10$ and 1.2% when $n = 100$, where the E_n are conveniently determined by (34). The estimate is also consistent with the nature of the singularity at $\rho = \frac{1}{5}$ of the differential equation (33). □

3.2. **Paths and meanders.** Now that the bulk of the work is done, asymptotic estimates of the basic counts of paths and meanders fall as a ripe fruit. The result for unconstrained paths is trivial, since the number of possibilities for size n is $P(1)^n$, a fact consistent with the simple pole of $W(z,1) = (1 - zP(1))^{-1}$. For meanders, three cases are to be distinguished depending upon the value of a quantity called the drift.

Definition 5. *Given a simple walk with characteristic polynomial $P(u)$, the* drift *is by definition the quantity*

$$\delta = P'(1).$$

In the unweighted case, the drift is thus the sum of all the possible values of the jumps, which constitutes an indicator of the "tendency" for the walk to go up or down. In the probabilistic case ($P(1) = 1$), the drift represents exactly the expected movement in the y-direction of any single step. For a symmetric walk, the drift is $\delta = 0$, while $\tau = 1$.

Theorem 4. *Consider a simple aperiodic walk. The number of paths of length n, $[z^n]W(z,1)$, is $P(1)^n$ exactly. Set*

$$\overline{Y}_1(z) := \prod_{j=2}^{c} (1 - u_j(z)).$$

The asymptotic number of meanders depends on the sign of the drift $\delta = P'(1)$ as follows:

$$\begin{aligned}
\delta = 0: \quad [z^n]F(z,1) \quad &\sim \quad \nu_0 \frac{P(1)^n}{\sqrt{\pi n}} \left(1 + \frac{c_1}{n} + \frac{c_2}{n^2} + \cdots\right) \\
&\qquad \nu_0 := \sqrt{2\frac{P(1)}{P''(1)}}\,\overline{Y}_1(\rho), \quad \rho = P(\tau)^{-1} = P(1)^{-1}; \\
\delta < 0: \quad [z^n]F(z,1) \quad &\sim \quad \nu_0^{\pm} \frac{P(\tau)^n}{2\sqrt{\pi n^3}} \left(1 + \frac{c_1'}{n} + \frac{c_2'}{n^2} + \cdots\right) \\
&\qquad \nu_0^{\pm} := -\sqrt{2\frac{P(\tau)^3}{P''(\tau)}}\,\frac{\overline{Y}_1(\rho)}{P(\tau) - P(1)}, \quad \rho = P(\tau)^{-1}; \\
\delta > 0: \quad [z^n]F(z,1) \quad &\sim \quad \xi_0 P(1)^n + \nu_0^{\pm} \frac{P(\tau)^n}{2\sqrt{\pi n^3}} \left(1 + \frac{c_1''}{n} + \frac{c_2''}{n^2} + \cdots\right) \\
&\qquad \xi_0 := (1 - u_1(\rho_1))\overline{Y}_1(\rho_1), \quad \rho_1 := P(1)^{-1}.
\end{aligned}$$

The formulæ have an intuitive meaning. In the case of a positive drift, a fraction close to ξ_0 of all the (unconstrained) walks is a meander, in accordance for the walks to have a natural tendency to go up. For negative drift, most paths tend to go down and the proportion of meanders is exponentially small, roughly like $(P(\tau)/P(1))^n$. For zero drift, the proportion becomes as large as $1/\sqrt{n}$, while the walks tend to oscillate not too far from the horizontal axis.

Proof. The discussion is based on the formula of Corollary 1 rewritten as

$$F(z,1) = \frac{1-u_1(z)}{1-zP(1)}\overline{Y}_1(z), \qquad \overline{Y}_1(z) := \prod_{j=2}^{c}(1-u_j(z)).$$

It suffices to examine the position of the zeros and the dominant singularity of the numerator in relation to $1/P(1)$ that is always a zero of the denominator. By proof arguments similar to Lemma 2, the quantity $\overline{Y}_1(z)$, being a symmetric function of small branches each of which is dominated by u_1, must remain analytic throughout $|z| \le \rho$.

In the case $\delta = 0$, one has $P'(1) = 0$, $\tau = 1$, and $\rho = 1/P(\tau) = 1/P(1)$. Thus, $(1-u_1)$ contributes a term of the form $(1-z/\rho)^{1/2}$ at $z = \rho$ while the denominator $(1-zP(1))$ has a simple zero there. Globally, the singularity of $F(z,1)$ is thus of type $1/\sqrt{}$, and the result follows.

For a negative drift, meaning $P'(1) < 0$, one must have $\tau > 1$, since $P'(u)$ increases from $-\infty$ to $+\infty$ when u ranges from 0^+ to $+\infty$. With $\rho = 1/P(\tau)$ (the structural radius) and $\rho_1 := 1/P(1)$, one then has $\rho_1 < \rho$. In this case, the prefactor $(1-zP(1))^{-1}$ has a pole at ρ_1; this pole is however cancelled by a zero in the numerator induced by the numerator $(1-u_1(z))$ (since $u_1(\rho_1) = 1$), so that ρ_1 is a removable singularity of $F(z,1)$. Consequently, the dominant singularity of $F(z,1)$ is at ρ, where $F(z,1)$ is of the square-root type.

For a positive drift, one must have $\tau < 1$, so that the prefactor induces a pole at $\rho_1 := 1/P(1)$ before $\overline{Y}_1$ or $1-u_1$ become singular. The argument concludes by "subtracting singularities", since the function,

$$F(z,1) - \frac{\overline{Y}_1(\rho_1)(1-u_1(\rho_1))}{1-zP(1)}, \quad \rho_1 := \frac{1}{P(1)},$$

now has a dominant singularity of the square-root type at ρ. □

The earlier discussion about the algebraic character of asymptotic constants applies: quantities like $\overline{Y}_1(\rho_1)$ and $\overline{Y}_1(\rho)$ can be determined by adapting Algorithm R of Section 2.3. Should the degrees of the algebraic numbers involved become fairly large, one can always resort to numerical analysis as the next example illustrates.

EXAMPLE 8. *Lucky periods in die casting.* In [63, p. 45], Pólya introduces the following problem: *"En jetant 2n dés à la fois, on peut obtenir différentes sommes de points de 2n à 12n. Le cas le plus probable est celui de 7n points. Désignons par A_n le nombre de combinaisons où se produit cet événement."* Imagine that at each of n rounds two dice are cast and the score of the round is the sum of the two dice's values. Pólya thus considers the number of ways A_n (and probability $A_n/36^n$) of reaching the balanced score $7n$ at the end of a game of dice consisting of n rounds. Pólya proceeds by an integral representation (precisely of the type used in the proof of Theorem 1) from which he concludes that the GF $A(z)$ has the character of an algebraic function, but does not make the calculation explicit.

By centring around the mean score of a round, which equals 7, it is easily realized that the problem is equivalent to a walk whose characteristic polynomial is

$$P(u) = u^{-5}\left(1+u+u^2+u^3+u^4+u^5\right)^2.$$

Let V_n be the number of bridges. (The quantity V_n is exactly Pólya's A_n.) Here, $c = -5$, $d = +5$; also $\tau = 1$ as the walk is symmetric, and $\rho = 1/36$. The asymptotic number of bridges is simply

$$V_n \sim \frac{6 \cdot 36^n}{\sqrt{2^2 \cdot 3 \cdot 5 \cdot 7\,\pi n}},$$

which is nothing but an avatar of the local limit gaussian law.

Consider next the modification of Pólya's problem where we ask for the number of "lucky" games, in the sense that at any time t the score is at least $7t$. This is equivalent to finding the number of meanders. Excursions surface if we further impose the final score to be $7n$ exactly. We have $\tau = 1$ and $\rho = \frac{1}{36}$. One should then examine the kernel equation at $z = \rho$,

$$u^5 - \frac{1}{36} u^5 P(u) = 0,$$

as this gives all the values of the small branches there. We find that there are 10 roots, amongst which $\tau = 1$ is a double root. The eight other go by pairs of complex conjugates, with

$$\zeta \doteq -0.36381 + 0.22924\, i, \quad \zeta' \doteq 0.06208 + 0.47622\, i,$$
$$\zeta'' \doteq -1.96746 + 1.23976\, i, \quad \zeta''' \doteq 0.26919 + 2.06476\, i.$$

Then, the quantity $Y_1(\rho)$ is determined numerically as the product of the roots of modulus less than $\tau = 1$, namely, $\zeta\,\overline{\zeta}\,\zeta'\,\overline{\zeta'}$. We find $Y_1(\rho) \doteq 0.42648$, so that the constant in the asymptotic formula for excursions can be determined to great accuracy:

$$E_n \sim C \cdot \frac{36^n}{\sqrt{n^3}}, \quad C \doteq 0.35865\,42111\,34518\,86172. \tag{46}$$

In the same vein, we determine $\overline{Y}_1(\rho) = (1-\zeta)(1-\overline{\zeta})(1-\zeta')(1-\overline{\zeta'})$ to be $\overline{Y}_1(\rho) \doteq 2.11615$, and

$$\frac{1}{36^n}[z^n]F(z,1) \sim \frac{C'}{\sqrt{n}}, \quad C' \doteq 0.93071\,59694\,87799\,20216$$

gives the probability of a lucky game (a meander). □

Pólya's example is interesting structurally. For instance, the excursion constant C in (46) involves $Y_1(\rho)$ that is a root of a self-reciprocal polynomial $\Xi(y)$ of degree 16 (found by Algorithm R and factorization), itself equivalent to a resolvent of degree 8 that turns out to be irreducible,

$$\Xi(y) = y^8 \widehat{\Xi}(y + y^{-1})$$
$$\widehat{\Xi}(v) = v^8 - 17v^7 - 152v^6 + 34v^5 - 551v^4 - 12053v^3 + 8038v^2 + 38692v + 12664,$$

but algebra stops there. In contrast, analyis based on the decomposability devolving from the kernel method provides fully satisfactory numerical answers.

3.3. **Periodicities.** The discussion above has been conducted under the assumption of aperiodicity. As we explain now, similar results hold for *periodic* walks provided suitable congruence conditions are imposed on the indices of coefficients of generating function. For reasons explained after Definition 4, we freely assume the set of jumps to be at least reduced, as this implies no loss in generality.

Take a set $\mathcal{S}$ corresponding to period p. We sketch the discussion in the case of excursions, with $E(z)$ the corresponding GF. Then, $E(z)$ is periodic with period p, meaning that it is of the form $E(z) = \widehat{E}(z^p)$ for some $\widehat{E}(z)$ that is analytic at 0.

The foregoing discussion of small branches continues to apply as long as $|z|$ stays inside the disk $|z| < \rho$, and the local analysis (42) of u_1 continues to hold as $z \to \rho$. However, it appears now that there are p conjugate dominant singularities at the points

$$\rho_j := \rho\eta^j, \qquad \eta = e^{2i\pi/p}.$$

Indeed, $E(z)$ satisfies $E(z) = E(\eta z)$, while Equation (42) describes the behaviour of $u_1(z)$ at ρ_j upon changing z into z/η^j. Then, each of the p singular elements cumulate and contribute jointly to $[z^n]E(z)$ provided $n \equiv 0 \mod p$. One finds in this way

$$E_n \sim p\mu_0 \frac{P(\tau)^n}{2\sqrt{\pi n^3}}, \qquad n = p\nu,\ \nu \in \mathbb{Z}_{\geq 0}$$

where μ_0 is (still) given by (38).

The analysis easily adapts to the other types of paths considered, and is summarized by a simple rule: *For a system of jumps of period p, the asymptotic form of the count of index n must be restricted to a suitable congruence class of n* mod *p in order for objects to exists; then the corresponding asymptotic formula is obtained from the estimate of the aperiodic case through multiplication by a factor of p.*

EXAMPLE 9. *Asymptotics of generalized Duchon's clubs.* We return to Example 5. The kernel equation is $1 - z(u^{-c} + u^d) = 0$, which gives the structural constant

$$\tau = \left(\frac{c}{d}\right)^{1/e}, \quad e = c + d.$$

The period is equal to e. The number of excursions of length n is nonzero only if $n \equiv 0 \pmod e$ and it satisfies (with $r = \rho^e$)

$$E_{e\nu} \sim D_{c,d}\, r_{c,d}^{-\nu}\, \nu^{-3/2}, \qquad r_{c,d} = \frac{c^c d^d}{e^e},$$

for some computable constant $D_{c,d}$. This generalizes the estimate of Duchon [22] who determined $D_{2,3}$ by a particular grammar construction followed by a specific algebraic elimination. □

4. Basic parameters and limit laws

The singular structure of basic generating functions of paths, bridges, meanders, and excursions is well established by Section 3. On the other hand, many parameters "decompose" combinatorially, so that their GF's are expressible in terms of the basic generating functions, or equivalently, they lie in $\mathbb{Q}(z, X; u_1, \ldots, u_c)$ for some set X of markers. In this paper, we only exhibit few sample cases of application of this methodology. As pointed by Philippe Robert (private communication), the whole combinatorial-analytic apparatus largely parallels what probabilists do by means of Wiener-Hopf decompositions (this is analogous to the separation between small and large branches) and Tauberian theorems (instead of singularity analysis that affords greater asymptotic accuracy through complete asymptotic expansions).

4.1. **Arches and contacts.** Define an arch as an excursion of size > 0 whose only contact with the horizontal axis is at its end points and let $\mathcal{A}$ be the set of arches. The set $\mathcal{E}$ of excursions satisfies the combinatorial equation

$$\mathcal{E} \cong \mathfrak{S}\{\mathcal{A}\},$$

where $\mathfrak{S}$ denotes the combinatorial construction that freely forms sequences. By well known mechanisms this translates directly into the GF equation

$$E(z) = \frac{1}{1-A(z)}, \quad \text{or, equivalently,} \quad A(z) = 1 - \frac{1}{E(z)}. \tag{47}$$

The singular form of $A(z)$ then reads immediately:

$$E(z) \sim E(\rho) - \mu_0\sqrt{1-z/\rho}, \quad \text{implying} \quad A(z) \sim \left(1 - \frac{1}{E(\rho)}\right) - \frac{\mu_0}{E(\rho)^2}\sqrt{1-z/\rho}.$$

Thus, the number of arches A_n is asymptotically proportional to $\rho^{-n}n^{-3/2}$, hence also to the number of excursions E_n.

Define a vertex of an excursion not equal to one of the end points to be a *contact* if its altitude is 0. Then, $A(z)^{k+1}$ is the GF of excursions having k contacts. For any fixed k, the function A^{k+1} has again a singularity of the square root type that is amenable to singularity analysis. An easy calculation then gives:

Theorem 5. *The probability that a random excursion of size n has k contacts is for any fixed k of the form*

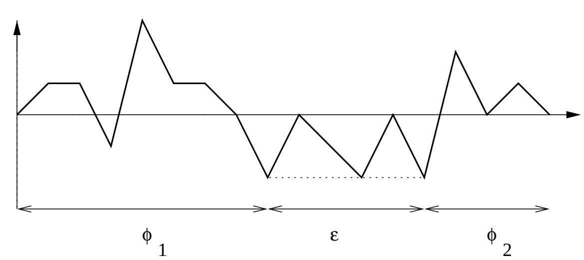
$$\frac{1}{E(\rho)^2}(k+1)\left(1-\frac{1}{E(\rho)}\right)^k + O\left(\frac{1}{n}\right).$$

The number of contacts is thus asymptotically distributed like the sum of two independent geometric random variables with parameter $1 - E(\rho)^{-1}$. In particular,

$$A_n \sim \frac{1}{E(\rho)^2}E_n.$$

The constant $E(\rho)$ is expressible in terms of the quantity $Y_1(\rho)$ and is thus a close relative of λ_0 introduced in Theorem 3.

On the relation between bridges and excursions. We briefly discuss here a construction that relates excursions to arches. Consider a bridge and let m (with $m \le 0$) be the minimal altitude of any vertex. Any nonempty bridge β decomposes uniquely into a walk φ_1 of size ≥ 1 from 0 to m that only reaches level m at its right end, followed by an excursion ε (this is the part where one wanders around but above level m), followed by a path φ_2 of size ≥ 0 from m to 0 that only touches level m at its beginning. By rearrangement, one can write $\beta = \varepsilon \cdot (\varphi_2|\varphi_1)$, where the glueing of $\varphi_2\varphi_1$ is an arch and the bar keeps track of where the splitting should occur. This construction is illustrated by the following diagram:



In other words, the set of nonempty bridges is combinatorially isomorphic to the product of the set of excursions by the set of arches with a split step that is distinguished. This construction is then nothing but the combinatorial reflex of the identity

$$\overbrace{V(z)-1}^{\text{bridges}} = \overbrace{E(z)}^{\text{excursions}} \cdot \overbrace{\left(z\frac{d}{dz}A(z)\right)}^{\text{split arches}}, \tag{48}$$

which, in view of (47) is equivalent to

$$V(z)-1 = E(z)\cdot z\frac{d}{dz}\left(1-\frac{1}{E(z)}\right) = z\frac{E'(z)}{E(z)}.$$

(Thus, combinatorics of arches gives back Corollary 2.) Such relations are ubiquitous in the theory of paths, the most famous ones being known by the names of Spitzer and Sparre Andersen: see Kittel's appendix to [35] and Lothaire's book [52, Sec. 5.3] for a summary. Raney's classic [65] and Gessel's papers [38, 39] make use of similar ideas (*inter alia*, the "cycle lemma") in combinatorial proofs of the Lagrange inversion formula. One of the many consequences of this orbit of ideas, is for instance the possibility of analysing the number of times a bridge attains its minimum value by adapting the decomposition (48) and closely mimicking the proof of Theorem 5. Louchard's analyses in [53] provide many striking illustrations of such an interplay between probabilistic and combinatorial properties.

4.2. **Final altitude of a meander.** The *final altitude* of a path is the abscissa of its end point. For unconstrained paths, the usual local and central limit theorems for discrete random variables apply [40, Chapter 9], so that the limit law, after normalization, is Gaussian, the underlying technology being plainly the saddle point method. We consider now meanders. The random variable associated to finite altitude when taken over the set of all meanders of length n is denoted by X_n, and it satisfies

$$\Pr(X_n = k) = \frac{[z^n u^k]F(z,u)}{[z^n]F(z,1)}.$$

We state:

Theorem 6. *The final altitude of a random meander of size n admits a limit distribution, with the limit law being dictated by the value of the drift δ.*

(*i*) *For a negative drift, $\delta < 0$, the limit distribution is a discrete one characterized in terms of the large branches:*

$$\lim_{n\to\infty} \Pr(X_n = k) = [u^k]\varpi(u), \quad \textit{where} \quad \varpi(u) = \frac{(1-\tau)^2}{(u-\tau)^2}\prod_{\ell\geq 2}\frac{1-v_\ell(\rho)}{u-v_\ell(\rho)}.$$

(*ii*) *In the case of zero drift, $\delta = 0$, the normalized random variable*

$$\frac{X_n}{\vartheta\sqrt{n}}, \qquad \vartheta = \sqrt{\frac{P''(1)}{P(1)}},$$

converges in law to a Rayleigh distribution defined by the density $xe^{-x^2/2}$:

$$\lim_{n\to\infty}\Pr\left(\frac{X_n}{\vartheta\sqrt{n}} \leq x\right) = 1 - e^{-x^2/2}.$$

(iii) In the case of a positive drift, $\delta > 0$, the standardized version of X_n,

$$\frac{X_n - \mu n}{\sigma\sqrt{n}}, \quad \mu = \frac{P'(1)}{P(1)}, \quad \sigma^2 = \left(\frac{P''(1)}{P(1)} + \frac{P'(1)}{P(1)} - \left(\frac{P'(1)}{P(1)}\right)^2\right),$$

converges in law to a Gaussian variable $\mathcal{N}(0,1)$:

$$\lim_{n\to\infty} \Pr\left(\frac{X_n - \mu n}{\sigma\sqrt{n}} \le x\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-y^2/2}\, dy.$$

In the case of a negative drift, the limiting distribution admits an explicit form

$$[u^k]\varpi(u) = \tau^{-k}(c_0 + c_1 k) + \sum_{\ell\ge 2} c_\ell v_\ell(\rho)^{-k},$$

for a set of constants c_j that can be made explicit by a partial fraction expansion of $\varpi(u)$.

Proof. (i) For a negative drift, one directly shows that the probability generating function of X_n at u converges pointwise to a limit that precisely equals $\varpi(u)$, the convergence holding for $u \in (0,1)$. By the fundamental continuity theorem [27, p. 280] for probability generating functions (PGF's), this entails convergence in law of the corresponding discrete distributions.

We now fix a value of u taken arbitrarily in $(0,1)$ and treated as a parameter. The PGF of X_n is

$$\frac{[z^n]F(z,u)}{[z^n]F(z,1)},$$

where $F(z,u)$ is given by Theorem 2. In the case of a negative drift we know from the proof of Theorem 4 that $\tau = v_1(\rho)$ satisfies $\tau > 1$ while the radius of convergence of $F(z,1)$ coincides with the structural radius ρ. Then, the quantity

$$\overline{Y_1}(z,u) = \prod_{\ell\ge 2}^{d} \frac{1}{u - v_\ell(z)}$$

is analytic in the closed disk $|z| \le \rho$: being a symmetric function of the nonprincipal large branches, it has no algebraic singularity there; given the already known domination relations between the large branches (Lemma 2), the denominators cannot vanish.

It then suffices to analyse the factor containing the principal large branch v_1. This factor has a branch point at ρ, where

$$\frac{1}{u - v_1(z)} \sim \frac{1}{u-\tau} + \frac{1}{(u-\tau)^2}\sqrt{2\frac{P(\tau)}{P''(\tau)}}\sqrt{1 - z/\rho},$$

as follows directly from (42) and the fact that v_1 is conjugate to u_1 at $z = \rho$. Singularity analysis then gives instantly the fact that, for some nonzero constant C,

$$[z^n]F(z,u) \sim C\rho^{-n} n^{-3/2}\Omega(u), \quad \text{where} \quad \Omega(u) = \frac{1}{(u-\tau)^2}\overline{Y_1}(\rho,u),$$

and the result follows after normalization by $[z^n]F(z,1)$.

For the remaining two cases, it will prove convenient first to estimate the mean value (expectation $\mathrm{E}(\cdot)$) of X_n,

$$\mathrm{E}(X_n) = \frac{[z^n]F'_u(z,1)}{[z^n]F(z,1)}, \tag{49}$$

where F'_u indicates differentiation with respect to u. Logarithmic differentiation gives

$$F'_u(z,1) = F(z,1)\sum_{\ell=1}^{d}\frac{1}{1-v_\ell(z)} \tag{50}$$

from which one attains singularities easily.

(*ii*) In the case of a zero drift, the value of the structural constant is $\tau=1$ and the radius of convergence of $F(z,1)$ is $\rho = 1/P(\tau) = 1/P(1)$. Then, the singularity at ρ of $F'_u(z,1)$ combines a factor $1/\sqrt{1-z/\rho}$ that arises from $F(z,1)$ and another similar factor that arises from the term $(1-v_1(z))^{-1}$. This singularity is thus, to first order asymptotics, similar to a simple pole. A computation based again on (42) reveals that the mean value of X_n is of the order of $\sqrt{n}$. Precisely, one finds

$$E(X_n) \sim \vartheta\sqrt{\frac{\pi n}{2}}, \qquad \vartheta = \sqrt{\frac{P''(1)}{P(1)}}.$$

(Note that $\sqrt{\pi/2}$ is the mean of the standard Rayleigh distribution.)

The formula of Corollary 3 then suggests that $F_k(z)$ should behave very much like v_1^k, implying that the coefficients should resemble, up to scaling, the coefficients in the large power $[z^n](1-\sqrt{1-z})^k$. Such a situation is known to be conducive to Rayleigh laws: it is covered extensively in Drmota and Soria's study [21] and revisited in the paper [5]; see also [20]. In particular Theorem 1 of [21] gives us the convergence in distribution to the Rayleigh law, while a simple adaptation of the results of Appendix B in [5] provides corresponding density estimates (a "local" limit law). We omit the tedious but routine details.

(*iii*) For a positive drift, probabilistic intuition indicates that there are relatively few chances for a walk to ever come under the negative axis, and when this happens, it only tends to do so early in the history of the walk. Consequently, the final altitude should be only marginally affected by the meander conditioning.

In this case, one has $\tau < 1$ and the radius of convergence of $F(z,1)$ is $\rho_1 = 1/P(1)$ while the structural radius satisfies $\rho > \rho_1$. By definition, one has $v_1(\rho_1) = 1$. Consequently, the function $F'_u(z,1)$ in (50) admits a double pole at ρ_1, with

$$F'_u(z,1) \sim F(z,1)\frac{1}{v'_1(\rho_1)(z-\rho_1)}.$$

so that (one has $v'_1(\rho_1) = -(\rho_1^2 P'(1))^{-1}$),

$$\mathrm{E}(X_n) = \frac{[z^n]F'_u(z,1)}{[z^n]F(z,1)} = n\frac{P'(1)}{P(1)} + O(1).$$

In the probabilistic case, the coefficient of n in the estimate reduces to the drift, and this estimate does agree with the probabilistic argument sketched above. Similarly, the variance is found to satisfy

$$\mathrm{Var}\, X_n = \left(\frac{P''(1)}{P(1)} + \frac{P'(1)}{P(1)} - \left(\frac{P'(1)}{P(1)}\right)^2\right) n + O(1).$$

Finally, the Gaussian law is established from the power-sum form of Corollary 3 upon applying Cauchy's coefficient formula. One has

$$[z^n]F_k(z) = \frac{1}{2i\pi}\int_{|z|=\rho_1} \xi_1(z)v_1(z)^{-k-1}\,\frac{dz}{z^{n+1}} + R_{n,k},$$

The error term $R_{n,k}$ that arises from all the nonprincipal branches is exponentially smaller than ρ_1^{-n} because of the domination properties of $1/v_1(z)$ (see the proof of Lemma 2, once more). The main integral is then treated by the saddle point method in the range considered, $k = \mu n + O(\sqrt{n})$ with $\mu := P'(1)/P(1)$. The saddle point of the integrand is at ρ_1, very nearly. The Gaussian density then comes out from a standard saddle point perturbation analysis. □

5. Directed two-dimensional models

The kernel method is generally well suited to problems where all the jumps are of the form (a_j, b_j) with $a_j \geq 0$. In this case, each choice of a step implies progression along the horizontal axis. One considers the trivariate GF

$$F(z; x, y) := \sum_{n,p,q} F_{n,p,q} z^n x^p y^q,$$

where $F_{n,p,q}$ is the number of meander paths in $\mathbb{Z}_{\geq 0} \times \mathbb{Z}_{\geq 0}$ with size (number of steps) equal to n that connect the origin to the point of coordinates (p, q). The walk is thus directed in the sense of Section 1. As we now explain, such enumeration problems, though formulated in two-dimensional space, are in fact fake 1-dimensional problems amenable to the kernel method.

In the directed case, the method of "adding a slice" encountered in Equations (14) and (16) gives rise to the fundamental equation

$$F(z; x, y)(1 - zP(x, y)) = 1 - z\{y^{<0}\}\left(P(x, y)F(z; x, y)\right), \tag{51}$$

where the characteristic polynomial is now

$$P(x, y) := \sum_j x^{a_j} y^{b_j},$$

which is entire in x but of Laurent type with respect to y. The parameters of size (marked by z) and horizontal displacement (marked by x) are bound by linear inequalities, and one of them can be treated as the basic variable, the other as an auxiliary parameter or even the constant 1. Then, the adaptation of the kernel method consists in *binding* the Laurent variable, here y, to the basic variable chosen (x or z) by

$$1 - zP(x, y) = 0. \tag{52}$$

Newton's polygon then shows that, for the bound equation, the number of "small" roots of the kernel equation coincides with the maximum negative vertical span, namely, $c := |\min_j b_j|$, and this number is precisely the number of unknown functions in the right side of (51). We let u_j represent these small branches. The treatment of walks and bridges adapts easily from what has been done earlier. Regarding excursions and meanders, substitution of the u_j then shows the following: *The GF of excursions (defined by final altitude 0) and the BGF of meanders (defined by final altitude ≥ 0) depend rationally on the variables z, x and the set of small branches $\{u_j\}$ of the associated "kernel equation"* (52).

Example 10. *Chess moves of Labelle and Yeh.* In two papers [49, 50], Labelle and Yeh develop an interesting set of decompositions for generalized knight moves on a chessboard. The standard version of the problem is: *Consider the $\mathbb{Z}_{\geq 0} \times \mathbb{Z}_{\geq 0}$ chessboard. How many sequences of Eastbound knight moves ($\mathcal{S} =$*

$\{(1,2),(1,-2),(2,1),(2,-1)\})$ *are there from* $(0,0)$ *to* $(n,0)$*?* By definition, the moves are not allowed to involve points with negative coordinates.

As size is not needed, we take x as the independent variable and set $z=1$. The kernel equation is then

$$1-(xy^2+xy^{-2}+x^2y+x^2y^{-1})=0.$$

so that the characteristic curve is a quartic. The vertical symmetry of the moves implies that the kernel equation can be rewritten as a combination of two quadratic equations,

$$1-x(W^2+xW-2)=0, \qquad W:=y+\frac{1}{y}.$$

There results that the four branches of the characteristic equation are given by

$$y_\pm(W)=\frac{1}{2}\left(W\pm\sqrt{W^2-4}\right), \quad W_\pm(x)=\frac{1}{2x}\left(-x^2\pm\sqrt{x^4+8x^2+4x}\right).$$

It appears that the two small branches u_1,u_2 correspond to taking opposite signs in the determinations of $y(W)$ and $W(x)$, and one finds for the GF of excursions (i.e., paths terminating at altitude 0), in complete analogy to the simple walk,

$$\begin{aligned} E(x) &= -\frac{1}{x}(u_1(x)u_2(x)) = -\frac{1}{x}y_-(W_+(x))\cdot y_+(W_-(x)) \\ &= 1+x^2+3\,x^4+2\,x^5+12\,x^6+14\,x^7+54\,x^8+86\,x^9+\cdots. \end{aligned}$$

This is the sequence (a_n) of [49] and also *EIS* **A005220**. Decomposability renders especially easy the asymptotic analysis of the number of excursions and of corresponding parameters. More general knight moves can be treated similarly by the kernel method. In particular, the equation satisfied by the excursion generating functions tends to be of a degree exponential in c; see [49, 50]. Here, the kernel method yields a reduction to an equation of degree $2c$, which even reduces to a resolvent of degree c when symmetry is taken into account via the W-parameterization. This illustrates a sharp contrast between the exponential blow-up in combinatorial complexity and the linear character of the analytic complexity. □

6. Conclusion

In this paper, we have aimed at illustrating the analytic tractability of many 1-dimensional path problems, a boon of the kernel method. The reduction in the asymptotic-analytic complexity of the problem is often spectacular, as exemplified by Duchon's clubs or the Labelle-Yeh knight moves. Parameters that are easily readable on paths lead to generating functions whose singularities arise simply from the branches of a characteristic curve of low degree. The method applies to all 1-dimensional problems as well as to 2-dimensional problems provided they remain directed. For a thorough discussion of the algebraic power of the kernel method, we refer once more to the study by Bousquet-Mélou and Petkovšek [13]. (The kernel technique is also reminiscent of Tutte's quadratic method much of use in the enumerative theory of planar maps [42]; see Bousquet-Mélou's paper [11] for a perspective.)

The case of undirected 2-dimensional problems, where one can go back and forth in all four cardinal directions, is appreciably harder. Even in the case of movement of amplitude ≤ 1, Fayolle *et al.* show in [26] that *stationary* solutions involve elliptic functions and integrals. Some directed path problems in dimension higher than 2 can however still be successfully treated by specific combinatorial decompositions; see [12] for an example.

A tribute to Maurice Nivat. As is apparent from the bibliography of this paper, many papers directly relevant to our study have been published in the journal *Theoretical Computer Science* along the years. We owe much for this to the Editor-in-Chief, Maurice Nivat. His openness of mind has been a constant help in the emergence and shaping up of sub-communities within theoretical computer science. Examples are the GASCOM (Generation of Random Combinatorial Objects) and AofA (Analysis of Algorithms) communities which have greatly benefitted from special issues of TCS, this at the invariably encouraging initiative of Maurice. In view of this and of Maurice's long-standing interest in similar discrete geometrical objects (see, e.g., [6, 7, 9, 17]), we kindly dedicate this study to him.

Acknowledgements. This work was supported in part by the IST Programme of the EU under contract number IST-1999-14186 (ALCOM-FT). The authors are grateful to Philippe Robert for many insightful discussions on probabilistic aspects of random walks as well as to Mireille Bousquet-Mélou and an anonymous referee for a careful scrutiny of the paper that greatly helped us improve our presentation. Thanks finally to Christian Krattenthaler for valuable bibliographical remarks.

References

1. S.-S. Abhyankar, *Algebraic geometry for scientists and engineers*, American Mathematical Society, 1990.
2. Cyril Banderier, *Combinatoire analytique: application aux marches aléatoires*, D.E.A. memoir, Université Paris VI, July 1998.
3. ———, *Combinatoire analytique des chemins et des cartes*, Ph.D. thesis, Université Paris VI, June 2001.
4. Cyril Banderier, Mireille Bousquet-Mélou, Alain Denise, Philippe Flajolet, Danièle Gardy, and Dominique Gouyou-Beauchamps, *Generating functions of generating trees*, Technical Report ALCOM FT-TR-01-17, Alcom-FT Project, February 2001, 26 pages. Accepted for publication in *Discrete Mathematics*.
5. Cyril Banderier, Philippe Flajolet, Gilles Schaeffer, and Michèle Soria, *Random maps, coalescing saddles, singularity analysis, and Airy phenomena*, Preprint, March 2001, 47 pages. Accepted for publication in *Random Structures & Algorithms*.
6. Elena Barcucci, Sara Brunetti, Alberto Del Lungo, and Maurice Nivat, *Reconstruction of discrete sets from three or more X-rays*, Algorithms and complexity (Rome, 2000), Springer, Berlin, 2000, pp. 199–210.
7. Elena Barcucci, Alberto Del Lungo, Maurice Nivat, and Renzo Pinzani, *Reconstructing convex polyominoes from horizontal and vertical projections*, Theoretical Computer Science **155** (1996), no. 2, 321–347.
8. Elena Barcucci, Renzo Pinzani, and Renzo Sprugnoli, *The random generation of directed animals*, Theoretical Computer Science **127** (1994), no. 2, 333–350.
9. Danièle Beauquier, Maurice Nivat, Éric Remila, and Mike Robson, *Tiling figures of the plane with two bars*, Computational Geometry, Theory and Applications **5** (1995), no. 1, 1–25.
10. Jean Berstel (ed.), *Séries formelles*, LITP, University of Paris, 1978, (Proceedings of a School, Vieux–Boucau, France, 1977).
11. Mireille Bousquet-Mélou, *On (some) functional equations arising in enumerative combinatorics*, Preprint, 2001.
12. Mireille Bousquet-Mélou and Anthony J. Guttmann, *Three-dimensional self-avoiding convex polygons*, Physical Review E. Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics. Third Series **55** (1997), no. 6, part A, R6323–R6326.
13. Mireille Bousquet-Mélou and Marko Petkovšek, *Linear recurrences with constant coefficients: the multivariate case*, Discrete Mathematics **225** (2000), no. 1-3, 51–75.

14. Louis Comtet, *Calcul pratique des coefficients de Taylor d'une fonction algébrique*, Enseignement Mathématique. **10** (1964), 267–270.
15. ———, *Advanced combinatorics*, Reidel, Dordrecht, 1974.
16. N. G. de Bruijn, *Asymptotic methods in analysis*, Dover, 1981, A reprint of the third North Holland edition, 1970 (first edition, 1958).
17. Alberto Del Lungo, Maurice Nivat, and Renzo Pinzani, *The number of convex polyominoes reconstructible from their orthogonal projections*, Proceedings of the 6th Conference on Formal Power Series and Algebraic Combinatorics (New Brunswick, NJ, 1994), vol. 157, 1996, pp. 65–78.
18. Marie-Pierre Delest and Gérard Viennot, *Algebraic languages and polyominoes enumeration*, Theoretical Computer Science **34** (1984), 169–206.
19. Peter G. Doyle and J. Laurie Snell, *Random walks and electric networks*, Mathematical Association of America, Washington, DC, 1984.
20. Michael Drmota, *Asymptotic distributions and a multivariate Darboux method in enumeration problems*, Journal of Combinatorial Theory, Series A **67** (1994), 169–184.
21. Michael Drmota and Michèle Soria, *Images and preimages in random mappings*, SIAM Journal on Discrete Mathematics **10** (1997), no. 2, 246–269.
22. Philippe Duchon, *On the enumeration and generation of generalized Dyck words*, Discrete Math. **225** (2000), no. 1-3, 121–135, Formal power series and algebraic combinatorics (Toronto, ON, 1998).
23. Marianne Durand, *Asymptotics of the "klam" recurrence*, Preprint, 2001.
24. L. Euler, *Observationes analyticae*, Novi Commentarii Acad. Sci. Imper. Petropolitanae **11** (1765), 124–143.
25. Guy Fayolle and Roudolf Iasnogorodski, *Two coupled processors: the reduction to a Riemann-Hilbert problem*, Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete **47** (1979), no. 3, 325–351.
26. Guy Fayolle, Roudolf Iasnogorodski, and Vadim Malyshev, *Random walks in the quarter-plane*, Springer-Verlag, Berlin, 1999.
27. W. Feller, *An introduction to probability theory and its applications*, third ed., vol. 1, John Wiley, 1968.
28. ———, *An introduction to probability theory and its applications*, vol. 2, John Wiley, 1971.
29. Philippe Flajolet, *Combinatorial aspects of continued fractions*, Discrete Mathematics **32** (1980), 125–161.
30. ———, *The evolution of two stacks in bounded space and random walks in a triangle*, Mathematical Foundations of Computer Science (J. Gruska, B. Rovan, and J. Wiedermann, eds.), Lecture Notes in Computer Science, vol. 233, Springer Verlag, 1986, Proceedings of the 12th MFCS Symposium, Bratislava, August 1986, pp. 325–340.
31. ———, *Analytic models and ambiguity of context–free languages*, Theoretical Computer Science **49** (1987), 283–309.
32. Philippe Flajolet, Jean Françon, and Jean Vuillemin, *Sequence of operations analysis for dynamic data structures*, Journal of Algorithms **1** (1980), 111–141.
33. Philippe Flajolet and Fabrice Guillemin, *The formal theory of birth-and-death processes, lattice path combinatorics, and continued fractions*, Advances in Applied Probability **32** (2000), 750–778.
34. Philippe Flajolet and Andrew M. Odlyzko, *Singularity analysis of generating functions*, SIAM Journal on Algebraic and Discrete Methods **3** (1990), no. 2, 216–240.
35. D. Foata, *La série génératrice exponentielle dans les problèmes d'énumération*, S.M.S, Montreal University Press, 1974.
36. Jean Françon, *Histoires de fichiers*, RAIRO Informat. Théor. **12** (1978), no. 1, 49–62.
37. Harry Furstenberg, *Algebraic functions over finite fields*, Journal of Algebra **7** (1967), 271–277.
38. Ira M. Gessel, *A factorization for formal Laurent series and lattice path enumeration*, J. Combin. Theory Ser. A **28** (1980), no. 3, 321–337.
39. ———, *A noncommutative generalization and q–analog of the Lagrange inversion formula*, Transactions of the American Mathematical Society **257** (1980), no. 2, 455–482.
40. B. V. Gnedenko and A. N. Kolmogorov, *Limit distributions for sums of independent random variables*, Addison-Wesley, 1968.

41. H. W. Gould, *Research bibliography on two number sequences*, In *Mathematica Monongaliae*, 1971, (A comprehensive bibliography on Bell and Catalan numbers).
42. Ian P. Goulden and David M. Jackson, *Combinatorial enumeration*, John Wiley, New York, 1983.
43. Peter Henrici, *Applied and computational complex analysis*, vol. 2, John Wiley, New York, 1974.
44. E. Hille, *Analytic function theory*, Blaisdell Publishing Company, Waltham, 1962, 2 Volumes.
45. Frances Kirwan, *Complex algebraic curves*, London Mathematical Society Student Texts, no. 23, Cambridge University Press, 1992.
46. Donald E. Knuth, *The art of computer programming*, 3rd ed., vol. 1: Fundamental Algorithms, Addison-Wesley, 1997.
47. ———, *The art of computer programming*, 3rd ed., vol. 2: Seminumerical Algorithms, Addison-Wesley, 1998.
48. ———, *The art of computer programming*, 2nd ed., vol. 3: Sorting and Searching, Addison-Wesley, 1998.
49. Jacques Labelle and Yeong Nan Yeh, *Dyck paths of knight moves*, Discrete Appl. Math. **24** (1989), no. 1-3, 213–221, First Montreal Conference on Combinatorics and Computer Science, 1987.
50. ———, *Generalized Dyck paths*, Discrete Mathematics **82** (1990), 1–6.
51. Pierre-Simon Laplace, *Théorie analytique des probabilités. Vol. I, II*, Éditions Jacques Gabay, Paris, 1995, Reprint of the 1819 and 1820 editions.
52. M. Lothaire, *Combinatorics on words*, Encyclopedia of Mathematics and its Applications, vol. 17, Addison–Wesley, 1983.
53. Guy Louchard, *Asymptotic properties of some underdiagonal walks generation algorithms*, Theoretical Computer Science **218** (1999), no. 2, 249–262.
54. E. Lucas, *Théorie des Nombres*, Gauthier-Villard, Paris, 1891, Reprinted by A. Blanchard, Paris 1961.
55. A. Meir and J. W. Moon, *On the altitude of nodes in random trees*, Canadian Journal of Mathematics **30** (1978), 997–1015.
56. Donatella Merlini, D. G. Rogers, Renzo Sprugnoli, and M. Cecilia Verri, *Underdiagonal lattice paths with unrestricted steps*, Discrete Applied Mathematics. Combinatorial Algorithms, Optimization and Computer Science **91** (1999), no. 1-3, 197–213.
57. Sri Gopal Mohanty, *Lattice path counting and applications*, Academic Press [Harcourt Brace Jovanovich Publishers], New York, 1979, Probability and Mathematical Statistics.
58. ———, *Combinatorial aspects of some random walks*, Random walks (Budapest, 1998), János Bolyai Math. Soc., Budapest, 1999, pp. 259–273.
59. T. V. Narayana, *Lattice path combinatorics with statistical applications*, University of Toronto Press, Toronto, Ont., 1979.
60. Maurice Nivat, *Langages algébriques sur le magma libre et sémantique des schémas de programme*, Automata, languages and programming (Proc. Sympos., Rocquencourt, 1972), North Holland, Amsterdam, 1973, pp. 293–308.
61. A. M. Odlyzko, *Asymptotic enumeration methods*, Handbook of Combinatorics (R. Graham, M. Grötschel, and L. Lovász, eds.), vol. II, Elsevier, Amsterdam, 1995, pp. 1063–1229.
62. Jim Pitman, *Brownian motion, bridge, excursion, and meander characterized by sampling at independent uniform times*, Electron. J. Probab. **4** (1999), no. 11, 33 pp. (electronic).
63. G. Pólya, *Sur les séries entières dont la somme est une fonction algébrique*, Enseignement mathématique **1–2** (1921–1922), 38–47.
64. Helmut Prodinger, *On a functional-difference equation of Runyon, Morrison, Carlitz, and Riordan*, Séminaire Lotharingien de Combinatoire **46** (2001), paper B46a, 4 pages (electronic).
65. G. N. Raney, *Functional composition patterns and power series reversion*, Transactions of the American Mathematical Society **94** (1960), 441–451.
66. Philippe Robert, *Réseaux et files d'attente: méthodes probabilistes*, Mathématiques & Applications, vol. 35, Springer, Paris, 2000.
67. Bruno Salvy and Paul Zimmermann, *GFUN: a Maple package for the manipulation of generating and holonomic functions in one variable*, ACM Transactions on Mathematical Software **20** (1994), no. 2, 163–167.
68. Masako Sato, *Generating funtions for the number of lattice paths between two parallel lines with a rational incline*, Mathematica Japonica **34** (1989), no. 1, 123–137.

69. Robert Sedgewick and Philippe Flajolet, *An introduction to the analysis of algorithms*, Addison-Wesley Publishing Company, 1996.
70. N. J. A. Sloane, *The on-line encyclopedia of integer sequences*, 2000, Published electronically at `http://www.research.att.com/~njas/sequences/`.
71. N. J. A. Sloane and Simon Plouffe, *The encyclopedia of integer sequences*, Academic Press, 1995.
72. R. P. Stanley, *Enumerative combinatorics*, vol. II, Cambridge University Press, 1998.
73. J. van Leeuwen (ed.), *Handbook of theoretical computer science*, vol. B: Formal Models and Semantics, North Holland, 1990.
74. Andrew Chi Chih Yao, *An analysis of* $(h, k, 1)$*-Shellsort*, J. Algorithms **1** (1980), no. 1, 14–50.
75. ———, *An analysis of a memory allocation scheme for implementing stacks*, SIAM Journal on Computing **10** (1981), no. 2, 398–403.

Cyril Banderier, Algorithms Project, INRIA, Rocquencourt, 78150 Le Chesnay (France).
E-mail address: `Cyril.Banderier@inria.fr`, `http://algo.inria.fr/banderier`

Philippe Flajolet, Algorithms Project, INRIA, Rocquencourt, 78150 Le Chesnay (France).
E-mail address: `Philippe.Flajolet@inria.fr`, `http://algo.inria.fr/flajolet`