Fourth Colloquium on Mathematics and Computer Science DMTCS proc. **AG**, 2006, 345–356

unidimensional lattice paths from an enumerative and asymptotic point of view. The authors developed a generating function approach to have exact enumeration of lattice paths (via the kernel method), and then used singularity analysis to study some basic parameters (like number of returns to zero, final altitude) according to some constraints (drift/reflecting conditions). For random walks there are in fact numerous works using an approach with a flavour of "generating function" and "analysis of singularity", either with probabilistic or combinatorial methods, e.g. for interaction of random walks, for random walks on groups or for stationary distributions of 2-dimensional models in queueing theory.

In our simpler model of unidimensional lattice paths, for simple parameters which are in one sense "exactly solvable", one can expect more than for the above difficult problems: not only one can get here critical exponents, but also we get fast computation schemes, for exact enumeration and for full asymptotics (to any order). It is then natural to ask what can be obtained for less trivial parameters, like the area or the height. We will investigate the height in a future article, and we concentrate here on the area, as already investigated in combinatorics, mainly for Dyck paths or with Riordan arrays for internal path lengths of some generating trees [23]. As lattice paths are algebraic objects, as easily proven with context free grammars [19], some techniques from language theory (Q-grammars [11]) can be used to solve the simplest cases. We extend most of these results in this article. The probabilistic corresponding object was analysed by G. Louchard [20], who proved that the limiting distribution of the area below the Brownian excursion was related to the Airy function, as further investigated in [24], and also by other authors in different contexts [14, 7, 6]. The area is also naturally related to queueing theory, polyominoes in statistical physics [27, 28], cumulative cost of some algorithms.

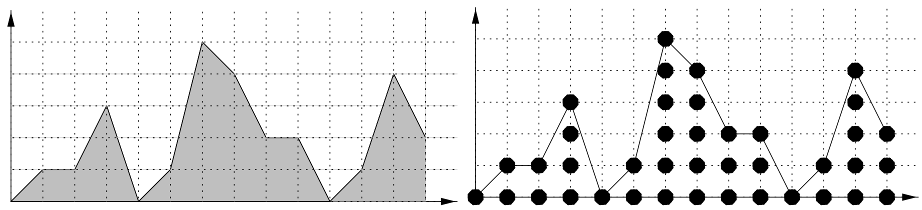


Fig. 1: The two kind of areas we consider in this article: To the left, a path of "continuous" area 25. Enumeration are given in Section 3 and asymptotics for simple families of walks are given in Section 4. To the right, a path of "discrete" area 40 (i.e., there are 40 points below the path). Section 5 is devoted to the analysis of this parameter.

In our approach, one of the key trick is the so-called "kernel method", which is a way of solving functional equations of the type $K(z,u)F(z,u) = A(z,u) + B(z,u)G(z)$ where F and G are the unknowns one wishes to determinate. The kernel method consists in getting additional equations by plugging the roots of the "kernel" K in the initial equation, which in general is enough to solve the system. The kernel method shares the spirit of the "quadratic method" of Tutte and Brown (for enumeration of maps). In combinatorics, only the simplest case of the kernel method (namely, when there is only one root) was used for 30 years, see Knuth [16] for sorting with stacks, Chung et al. [9] for a pebbling game, or [25] for generation of binary trees. During the same time period, and independently, difficult 2-dimensional generalisations of this trick were well studied in queueing theory; the classification of the different cases for the nearest neighbour walks in $\mathbb{N}^2$ was already quite a challenge, see the book [13] or [18]. This last decade, there has been a revival in combinatorics for functional equations, and the full power of the kernel method was better put into evidence, both for enumeration and for asymptotics. There are indeed nearly 20 (sic!) articles by M. Bousquet-Mélou [5], which e.g. showed how the kernel method can be, once bootstrapped, also be used in higher dimensions or for algebraic (non linear) equations. Solving equations is not the only miracle that the kernel method offers, it also gives compact formulae [5, 1, 3] thus giving access to asymptotics [4, 2, 29]. Our article thus adds a new stone to the "kernel method" edifice, and gives more results on complete asymptotics for the average area of directed lattice paths.

2 Summary of results for directed lattice paths

To each directed lattice path, we associate a Laurent polynomial which encodes all the possible jumps $P(u) := \sum_{i=-c}^{d} p_i u^i$ (where c is the size of the largest backward jump and d is the size of largest ahead jump, and where the p_i's are some "weights", "multiplicities", or "probabilities").

Figure 2 shows four drawings (for four different constraints) of lattice paths with jumps in

$$\mathcal{S} = \{(1,-3),(1,-1),(1,0),(1,1),(1,5)\}\,;$$

the associated Laurent polynomial is therefore $P(u) := p_{-3}u^{-3} + p_{-1}u^{-1} + p_0 + p_1 u + p_5 u^5$.

In [4], Banderier and Flajolet showed that the kernel method was the key to get enumeration and asymptotics of directed lattice paths. The main results are summarised in Figure 2. The proofs rely on the following facts:
Fact 1: There are c distinct roots, $u_1, \ldots, u_c$ of the "kernel" $1 - zP(u) = 0$ which are analytical in zero.
Fact 2: There is a nice trick (the "kernel method") which allows to write all the GF's with these u_i's.
Fact 3: There is a unique positive real number τ such that $P\ (\tau) = 0$, and the radius of convergence of the GF's is $\rho := 1/P(\tau)$.
Fact 4: Asymptotics are coming from the real root u_1, which is singular at ρ, the other roots are analytical at ρ and therefore, they only affect the multiplicative constant. (Some easy modifications have to be made here if the walk is "periodic").
Fact 5: u_1 has a square root behaviour near its singularity: $u_1 \sim \tau + K\sqrt{1 - z/\rho}$.
Fact 6: The drift $\delta := P\ (1)$ of the walk plays a rôle for asymptotics of meanders ($\delta \geq 0$ when $\tau \geq 1$, $\delta \leq 0$ when $\tau \leq 1$).

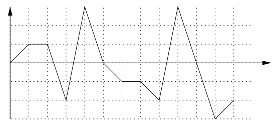
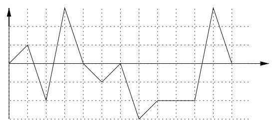
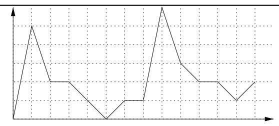
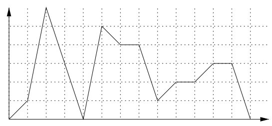
	ending anywhere	ending at 0
unconstrained (on $\mathbb{Z}$)	 walk/path ($\mathcal{W}$) $W(z) = \dfrac{1}{1 - zP(1)}$ $W_n = P(1)^n$	 bridge ($\mathcal{B}$) $B(z) = z\sum_{i=1}^{c} \dfrac{u_i'(z)}{u_i(z)}$ $B_n \sim \beta_0 \frac{P(\tau)^n}{\sqrt{\pi n}}$
constrained (on $\mathbb{Z}_{\geq 0}$)	 meander ($\mathcal{M}$) $M(z) = \dfrac{1}{1 - zP(1)} \prod_{i=1}^{c} (1 - u_i(z))$ $M_n \sim \mu_0 \frac{P(1)^n}{\sqrt{\pi n}}$ (zero drift) $M_n \sim \mu_0^- \frac{P(\tau)^n}{\sqrt{\pi n^3}}$ (negative drift) $M_n \sim \mu_0^+ P(1)^n + \mu_0^- \frac{P(\tau)^n}{\sqrt{\pi n^3}}$ (positive drift)	 excursion ($\mathcal{E}$) $E(z) = \dfrac{(-1)^{c-1}}{p_{-c}z} \prod_{i=1}^{c} u_i(z)$ $E_n \sim {}_0 \frac{P(\tau)^n}{\sqrt{\pi n^3}}$

Fig. 2: The four types of paths: walks, bridges, meanders, and excursions and the corresponding generating functions. The u_i's are such that $1 - zP(u_i(z)) = 0$ and the constants ${}_0$, β_0, and the μ_0's are algebraic numbers which can be made explicit. In the rest of this paper, we want to analyse the area enclosed between a constrained path and the $y = 0$ line. For the meander drawn here, the area is 27, and the area of the excursion is 34.

3 Generating function for the area

Taking the unit square of $\mathbb{Z}^2$ as unit of area, it is convenient to consider the generating function of the "doubled" area (i.e., area multiplied by 2): it precludes a mathematically equivalent but practically boring use of Puiseux series in our context.

Definition 1 (Area Generating Function) *Let f_{nkm} denote the number of walks of length n with final altitude k and area $m/2$ and*

$$F(z,u,q) = \sum_{n,k,m} f_{nmk} z^n u^k q^m = \sum_{k \geq 0} F_k(z,q) u^k .$$

Theorem 1 (Fundamental Functional Equation for $F(z,u,q)$) *The area generating function satisfies the following recursive definition:*

$$F(z,u,q) = 1 + zP(uq)F(z,uq^2,q) - z\sum_{k=0}^{c-1} r_k(uq)F_k(z,q)q^k , \qquad (1)$$

where the r_k's are Laurent polynomial defined by $r_k(u) := \{u^{<0}\}\left(P(u)u^k\right) \equiv \sum_{j=-c}^{-k-1} p_j u^{j+k}$.

Proof: If at time n, one is at altitude k with doubled-area r, and if one makes a jump j, then at time $n+1$ one is at altitude and doubled-area encoded by $u^{k+j}q^{r+2k+j}$. This transformation can be encoded with some linear operators: $f_{n+1}(u,q) = \{u^{\geq 0}\}P(uq)f_n(uq^2,q)$. This leads to

$$F(z,u,q) = 1 + zP(uq)F(z,uq^2,q) - \{u^{<0}\}zP(uq)F(z,uq^2,q)\,.$$

We get the theorem by noting that

$$\{u^{<0}\}P(uq)F(z,uq^2,q) = \sum_{k=0}^{c-1} F_k(z,q)q^{2k}\{u^{<0}\}P(uq)u^k\,. \qquad \square$$

In some cases (mainly for walks with jumps +1, 0, and -1, which are related to the combinatorics of continued fractions), it is possible to get q-analog expressions for $F_0(z,q)$. It is also related to the fact that one then has a simpler functional equation, which is possible to iterate, and from which it is also possible to pump moments, and a recurrence for them which leads to an Airy distribution for the area. For the more general case that we consider here, there is no hope to get such expressions, and the mixture of addition/subtractions, prevents us from using the same "pumping moment" approach. However, we will see that we can get nice results for the moments. For this, we will apply ∂_q^n on both sides of our fundamental functional equation, and this is why we need now the following theorem.

Theorem 2 (Bivariate Faà di Bruno Formula (Most, 1870)) *Consider $f(u,q)$ (and note ∂_u for $\frac{d}{du}$), then*

$$\partial_x^n f(g_1(x),g_2(x)) = \sum_{\lambda_1=0}^{n}\sum_{\lambda_2=0}^{n-\lambda_1}(\partial_u^{\lambda_1}\partial_q^{\lambda_2}f)(g_1(x),g_2(x))\sum_{\mathcal{K}} n!\prod_{j=1}^{n}\frac{(\partial_x^j g_1)^{k_{1j}}(\partial_x^j g_2)^{k_{2j}}}{k_{1j}!k_{2j}!j!^{k_{1j}+k_{2j}}}\,,$$

where the last summation runs over the set $\mathcal{K}$ defined by

$$\mathcal{K} := \{k_{11},\dots,k_{1n},k_{21},\dots,k_{2n} : k_{1j}+k_{2j} \geq 0,$$
$$\sum_{j=1}^{n} k_{1j} = \lambda_1, \sum_{j=1}^{n} k_{2j} = \lambda_2, \sum_{j=1}^{n}(k_{1j}+k_{2j})j = n\}\,.$$

Proof: See [10]. $\square$

Nota bene: This huge sum has in fact a lot of 0 terms. It is possible to write more complicated expressions (by adding a nested sum) with less 0 terms; however this new expression would even not be more efficient (while trying to use it on a computer algebra system like Maple[(i)]).

Proposition 1 (Application of Faà di Bruno to $\partial_q^n F(z,uq^2,q)$)

$$\begin{aligned}\partial_q^n F(z,uq^2,q) = \quad & \sum_{\lambda=0}^{n}(\partial_u^{\lambda}\partial_q^{n-\lambda}F)(z,uq^2,q)\frac{(2uq)^{\lambda}n!}{(n-\lambda)!\lambda!} + (\partial_u^{n/2}F)(z,uq^2,q)\frac{n!u^{n/2}}{(n/2)!}\\ + \quad & \sum_{\lambda_1=1}^{n-2}\sum_{\lambda_2=\max(1,n-2\lambda_1)}^{n-\lambda_1-1}(\partial_u^{\lambda_1}\partial_q^{\lambda_2}F)(z,uq^2,q)\frac{n!u^{\lambda_1}(2q)^{2\lambda_1+\lambda_2-n}}{(2\lambda_1+\lambda_2-n)!(n-\lambda_1-\lambda_2)!\lambda_2!}\\ + \quad & \sum_{\lambda_1=n/2+1}^{n-1}(\partial_u^{\lambda_1}F)(z,uq^2,q)\frac{n!u^{\lambda_1}(2q)^{2\lambda_1-n}}{(2\lambda_1-n)!(n-\lambda_1)!}\,.\end{aligned}$$

Proof: Bivariate Faà di Bruno formula with $f = F(z,u,q)$, $g_1(q) = uq^2$, $g_2(q) = q$, and $f(g_1,g_2) = F(z,uq^2,q)$ (z is considered here as a parameter). $\square$

The following result was stated in [1], but the proof was just sketched.

(i) Note that Maple needs to be "human-helped" a lot, when dealing with this kind of formulae. This "pleasant" fact is due to bugs in Maple: Indeed Maple is sometimes converting its diff function into its D function without warning the user (which could miss this fact, cancelled after simplifications), and the D function is buggy while dealing with Faà di Bruno like differentiation. Another Maple trouble that we encounter is mul$(x, i = 3..1) = 1$ (which is fine) whereas product$(x, i = 3..1) = 1/x$ (sic!). The interested reader can check all our computations with the file `http://www-lipn.univ-paris13.fr/~banderier/Papers/area.mw`. At this occasion, we would like to advertise Salvy and Zimmermann's Gfun package, which allowed us to get automatic proofs for some of our generating functions.

Theorem 3 (Algebraicity of moments) *The moments of the area are algebraic, i.e.* $(\partial_q^n F)(z,u,1)$ *is an algebraic generating function. Furthermore, it can be expressed as a rational function in terms of the roots of the kernel, i.e.* $(\partial_q^n F)(z,u,1) \in \mathbb{Q}(z,u,u_1,\dots,u_c)$.

Comments: this field is in fact $\mathbb{Q}(z,u,u_1,\dots,u_c,p_{-c},\dots,p_d)$ if the Laurent polynomial $P(u)$ of the walk has non-rational coefficients. Note that the platypus trick (sic) presented in [4] implies that coefficients $\partial_q^r f_n(u,1)$ can be computed in linear time.

Proof: We make a proof by induction. Our claim is true for $n=0$: in this case $(\partial_q^n F)(z,u,1) = F(z,u,1)$, which is algebraic and in $\mathbb{Q}(z,u,u_1,\dots,u_c)$ (see [4] for a detailed proof). Now, let us assume that the induction hypothesis is true for all $\lambda_2 < n$ (which implies[(ii)] that $(\partial_u^k \partial_q^{\lambda_2} F)(z,u,1)$ is algebraic and in $\mathbb{Q}(z,u,u_1,\dots,u_c)$ for all k), then we will show that $\partial_q^n F(z,u,1)$ is algebraic and in $\mathbb{Q}(z,u,u_1,\dots,u_c)$.

To this aim, consider the nth derivative of the fundamental functional equation 1, using Leibniz rule and setting $G_k(z,q) := F_k(z,q)q^k$ gives:

$$\begin{aligned}(\partial_q^n F)(z,u,q) = \quad & zP(uq)\partial_q^n F(z,uq^2,q) + z\sum_{j=0}^{n-1}\binom{n}{j}\partial_q^{n-j}P(uq)\partial_q^j F(z,uq^2,q)\\ - \quad & z\sum_{k=0}^{c-1}\sum_{j=0}^{n}\binom{n}{j}\partial_q^{n-j}(r_k(uq))\partial_q^j G_k(z,q)\,.\end{aligned}$$

Define $R_n(z,u,q)$ as the "known" part (by induction, once one sets $q=1$) of the right hand side, that is

$$\begin{aligned}R_n(z,u,q) := \quad & zP(uq)\left(\partial_q^n F(z,uq^2,q) - \partial_q^n(F)(z,uq^2,q)\right)\\ + \quad & z\sum_{j=0}^{n-1}\binom{n}{j}\partial_q^{n-j}P(uq)\partial_q^j F(z,uq^2,q) - z\sum_{k=0}^{c-1}\sum_{j=0}^{n-1}\binom{n}{j}\partial_q^{n-j}(r_k(uq))\partial_q^j G_k(z,q)\,.\end{aligned}$$

Setting $q=1$, the equation is then simply:

$$(1-zP(u))(\partial_q^n F)(z,u,1) = R_n(z,u,1) - z\sum_{k=0}^{c-1} r_k(u)\partial_q^n G_k(z,1)\,.$$

R_n contains only derivatives of order $< n$ and therefore (by induction, see the previous footnote) is algebraic and belongs to $\mathbb{Q}(z,u,u_1,\dots,u_c)$. Plugging the c roots of the kernel in this equation (this substitution is analytically legitimate) and, taking all small branches into account, provides a system of c equations in the unknown functions $\partial_q^n G_0,\dots,\partial_q^n G_{c-1}$:

$$\begin{cases} R_n(z,u_1,1) - z\sum_{k=0}^{c-1}(r_k(u_1))\partial_q^n G_k(z,1) &= 0\,,\\ \vdots & \\ R_n(z,u_c,1) - z\sum_{k=0}^{c-1}(r_k(u_c))\partial_q^n G_k(z,1) &= 0.\end{cases}$$

One has $M.\overrightarrow{(\partial_q^n G_i)_i} = \overrightarrow{(R_n(u_i)/z)_i}$ where the matrix M of this system has the following shape:

$$M := \begin{pmatrix} (p_{-c}u_1^{-c} + u_1^{-c+1} + \ldots + u_1^{-1}) & \ldots & p_{-c}u_1^{-2} + p_{-c+1}u_1^{-1}) & (p_{-c}u_1^{-1})\\ \vdots & \vdots & \vdots & \vdots\\ (p_{-c}u_c^{-c} + u_c^{-c+1} + \ldots + u_c^{-1}) & \ldots & (p_{-c}u_c^{-2} + p_{-c+1}u_c^{-1}) & (p_{-c}u_c^{-1})\end{pmatrix}$$

The determinant of M is unchanged if we add/subtract rows between them. Then, subtracting iteratively a multiple of the i-th row to the $(i-1)$-th row gives a classical Vandermonde matrix V, multiplied by p_{-c}

$$V := \begin{pmatrix} p_{-c}u_1^{-c} & \ldots & p_{-c}u_1^{-2} & p_{-c}u_1^{-1}\\ \vdots & \vdots & \vdots & \vdots\\ p_{-c}u_c^{-c} & \ldots & p_{-c}u_c^{-2} & p_{-c}u_c^{-1}\end{pmatrix}$$

(ii) If $F(z,u)$ is algebraic, then $F(z,0)$ is algebraic. If $F(z,u)$ is algebraic, then $\partial_u F(z,u)$ is algebraic (this can be proven by a resultant with the derivative of the algebraic equation of F). From this, it is easy to get that if $F(z,u) = \sum u^k F_k(z)$ is algebraic, then each $F_k(z)$ is algebraic.

we can use the classical formula to get its determinant, and it gives

$$\det M = \det V = p_{-c}^{c} \prod_{i=1}^{c} u_i^{-c} \prod_{1 \le i < j \le n} (u_j(z) - u_i(z))$$

which is nonzero as $p_{-c} \neq 0$ (the multiplicity of the largest negative jump) and as all the roots are distinct (by Fact 1 in Section 2). Therefore our system has a unique solution, and all the unknowns are algebraic as they can be expressed thanks to the Cràmer formula as a fraction in the u_i's:

$$\partial_q^n G_k(z,1) = \frac{\det M_k}{\det M}.$$

Nota bene: it is in fact possible to express $\det M_k$ as a polynomial involving homogeneous symmetric functions. But this involves positive and negative integer coefficients, and therefore it seems hopeless to get a general asymptotic theory of the area from such a formula. □

In order to get a nicer expression for $\partial_q F(z,u,1)$, we need some intermediate computations, e.g. for computing $P\;\,(u_k), P\;\,(u_k)$ and other such expressions, for sake of simplicity, we simply include here the following Proposition, the other proofs having the same flavour.

Proposition 2 (Computations of $\partial_u F(z,u,1)$) *Writing u_k for $u_k(z)$, one has:*

$$\partial_u F(z,u_k,1) = \frac{\Pi\;\,(u_k)}{A\;\,(u_k)} \sum_{i=1..c,=k} \frac{1}{u_k - u_i} - \frac{\Pi\;\,(u_k) A\;\,(u_k)}{2A\;\,(u_k)^2} = -\frac{1}{2}(\frac{\Pi}{A})$$

where $\Pi(u) = \prod_{i=1..c}(u - u_i)$ *and* $A(u) = u^c(1 - zP(u))$.

Proof: Wh $\Pi(u)$ and $A(u)$ defined as above, derivating $F(z,u,1) = \frac{\Pi(u)}{A(u)}$ leads to

$$\partial_u F(z,u,1) = \frac{\Pi\;\,(u)}{A(u)} - \frac{A\;\,(u)}{A^2(u)}\Pi(u)$$

$$= \frac{\Pi(u)}{u - u_k(z)} \frac{u - u_k(z)}{A(u)} \left(\sum_{i=1..c,=k} \frac{1}{u - u_i(z)} + \frac{A(u) - A\;\,(u)(u - u_k(z))}{A(u)(u - u_k(z))} \right).$$

Dividing numerator and denominator of the last fraction by $(u - u_k(z))^2$ leads to

$$\partial_u F(z,u,1) = \frac{\Pi(u) - \Pi(u_k)}{u - u_k(z)} \frac{u - u_k(z)}{A(u) - A(u_k)} (\sum_{i=1..c,=k} \frac{1}{u - u_i(z)} + \frac{\frac{A(u)/(u-u_k(z)) - A'(u)}{u - u_k(z)}}{A(u)/(u - u_k(z))}).$$

Taking limit when u goes to u_k on both sides gives

$$\partial_u F(z,u_k(z),1) = \Pi\;\,(u_k(z)) \frac{1}{A\;\,(u_k)} (\sum_{i=1..c,=k} \frac{1}{u_k - u_i(z)} + \lim_{u \to u_k} \frac{\frac{A(u)/(u-u_k(z)) - A'(u)}{u - u_k(z)}}{A\;\,(u_k(z))}).$$

Using the Taylor expansion of $A(u)$ up to $O((u - u_k)^3)$ and of $A\;\,(u)$ up to $O((u - u_k)^2)$ gives :

$$\partial_u F(z,u_k(z),1) = \Pi\;\,(u_k(z)) \frac{1}{A\;\,(u_k)} (\sum_{i=1..c,=k} \frac{1}{u_k - u_i(z)} + \frac{-A\;\,(u_k)}{2A\;\,(u_k)}).$$ □

4 Average area for simple family of lattice paths

We call "simple family of lattice paths" or "Łukasiewicz walks" the important family of walks corresponding to the case $c = 1$ (there is only one jump of length 1 to the left). Note that we could allow $d = +\infty$, which means that $P(u)$ could be a Laurent series (this would correspond to càdlàg processes in probability theory); our formula would then be derived in the same way (and for asymptotics several subcases have to be considered, see [1, 2, 3]).

Łukasiewicz excursions with jumps in $\mathcal{S}$ correspond to trees whose node degrees are constrained to lie in $1+\mathcal{S}$. This holds by virtue of the Łukasiewicz encoding, a well-known correspondence. (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 the equation $1-zP(u)=0$ gives the GF of "simple families of trees"counted according to their number of nodes, an otherwise classical result [21]. By Lagrange inversion, the number of trees comprised of n nodes is $T_n = \frac{1}{n}[u^{n-1}]\varphi(u)^n$, where $\varphi = uP(u)$ can be directly interpreted as the characteristic polynomial of the allowed node (out)degrees. The rich combinatorics of these structures suggests that we could also get rather explicit formulae for the area of such walks. This is indeed the case, as shown in the next three theorems.

Theorem 4 *The generating function of the moments of the area below Łukasiewicz walks (i.e., $c=1$) simplify to:*

$$(\partial_q^n F)(z,u,1) = \frac{1}{1-zP(u)}\left(R_n(z,u,1) - u_1\frac{R_n(z,u_1,1)}{u}\right). \tag{2}$$

The generating function of the first moment is:

$$(\partial_q F)(z,u,1) = \frac{uz(1+zP(u))P\ (u)(u-u_1)}{(1-zP(u))^3} + \frac{2zP(u)u_1}{(1-zP(u))^2} - \frac{u_1+\frac{zu_1u_1''}{u_1'} - zu_1}{1-zP(u)}.$$

Proof: Propositions 2 and 3 allow to get the following expressions:

$$R_1(z,u,1) = 2zP(u)\partial_u F(z,u,1) + zuP\ (u)F(z,u,1) + \frac{zp_{-1}}{u}E(z), \tag{3}$$

$$R_1(z,u_1,1) = 2 + \frac{zu_1}{u_1} - \frac{zu_1}{u_1}. \tag{4}$$

Using the expressions for $F_0(z,1) = E(z)$ and $F(z,u,1)$ given in Section 2, and plugging this in Eq. 2 (for $n=1$) gives the first moment. □

Theorem 5 *The generating function of the average area below Łukasiewicz excursions is given by*

$$\partial_q F_0(z,1) = \frac{2}{zp_{-1}}u_1 + \frac{u_1u_1}{p_{-1}u_1} - \frac{2}{p_{-1}}u_1 = 2E\ + E\Theta\ln(zE) \tag{5}$$

where E is the GF of excursions and Θ stands for $z\partial_z$, the pointing operator(iii)*. The average area below an excursion is asymptotically:*

$$\frac{\tau\sqrt{\pi P(\tau)}}{\rho p_{-1}\sqrt{2P\ (\tau)}}n^{3/2} - \frac{3}{\rho p_{-1}}n - \frac{3\tau\sqrt{\pi P(\tau)}}{8p_{-1}\rho\sqrt{2P\ (\tau)}}\sqrt{n} + \frac{7}{2\rho p_{-1}} + O\left(\frac{1}{\sqrt{n}}\right).$$

Proof: Plugging $u=0$ in Theorem 4 and using several Taylor expansions gives

$$\partial_q F_0(z,1) = \frac{2}{zp_{-1}}u_1 + \frac{u_1u_1}{p_{-1}u_1} - \frac{2}{p_{-1}}u_1.$$

Then, we know by Fact 5 from Section 2 that $u_1(z) = \tau + K\sqrt{1-z/\rho} + \dots$ (with $K = 2P(\tau)/P\ (\tau)$). Using singularity analysis, this leads to

$$\partial_q F_0(z,1) = \frac{-K}{\sqrt{1-z/\rho}\rho p_{-1}} + 2\frac{\tau - K\sqrt{1-z/\rho}}{zp_{-1}} + \frac{\tau - K\sqrt{1-z/\rho}}{2((1-z/\rho)\rho p_{-1}}.$$

The average is then the quotient

$$\frac{[z^n]\partial_q F_0(z,1)}{[z^n]F_0(z,1)} = \frac{\tau\sqrt{\pi P(\tau)}}{\rho p_{-1}\sqrt{2P\ (\tau)}}n^{3/2} - \frac{3}{\rho p_{-1}}n - \frac{3\tau\sqrt{\pi P(\tau)}}{8p_{-1}\rho\sqrt{2P\ (\tau)}}\sqrt{n} + \frac{7}{2\rho p_{-1}} + O\left(\frac{1}{\sqrt{n}}\right)$$

(iii) Using the notations of symbolic combinatorics used in [15], note that we have both for EGF and OGF that $\Theta\ln\mathcal{E} = \mathcal{E}\Theta\mathcal{A}$ whenever $\mathcal{E} = \mathrm{Seq}\mathcal{A}$. For generating functions, this translates in $z\partial_z \ln E(z) = E(z)z\partial_z A(z)$. There is therefore a natural meaning for logarithms, also for non-labelled combinatorial structures: this is nothing else than another formulation of the cycle lemma (also known as Dvoretsky/Motzkin/Raney/Spitzer/Sparre Andersen principles). To give a bijective proof of our formula 5 is an interesting problem.

from which we can get as many terms as one wishes in the asymptotic expansion (note that the denominators are well defined: $P\ (\tau) > 0$ as P has ≥ 0 coefficients). □

Morality: whatever the drift of the excursion is, we have universality of the $n^{3/2}$ result for average area below an excursion.

Theorem 6 *The average area below Łukasiewicz meanders is given by*

$$(\partial_q F)(z, 1, 1)$$

counting "signed" area, counting other cumulative parameters... All these extensions are indeed suitable with the kernel method, via the approach presented in this article. For sake of simplicity, we preferred to stay here within the framework of a Laurent polynomial $P(u)$ whereas more general results can be written while using a Laurent series $P(z, x, y)$ (series in z and x, Laurent series in y).

For simple directed walks, we can also possible to compute the variance. In the Dyck case, the first asymptotic term of each moment follows from Louchard's result: the Airy distribution for area of Brownian excursions. Beyond the Dyck case, it is not known if discrete excursions are weakly converging towards the Brownian excursion. Tying down a random walk is not a continuous operation and therefore Donsker's theorem for unconstrained walks is not sufficient (note that the Brownian positivity constraint could – with respect to finite dimensional densities – also be realized by a $o(\sqrt{n})$ region constraint in the discrete case). So, in conclusion, for the general case ($c \geq 2$), the situation remains open: our computations become messy, and it is right now not clear if the symmetric functions we got are not hiding some tricky (asymptotical) cancellations. But we believe the pumping method should again lead to the Airy recurrence for the first asymptotic term of moments, and then to the Airy distribution in all cases.

Acknowledgements

The authors are grateful to Guy Louchard, Svante Janson, Philippe Chassaing, and Jean-François Marckert for interesting discussions. The first author is indebted to Donatella Merlini for discussions during autumn 2001 about the area and the kernel method, and to Mireille Bousquet-Mélou who asked the interesting question "when do you have a rational generating function for some kind of area ?", which then lead to the discussion of Section 5.

We thank Philippe Flajolet and Alois Panholzer for bibliographical references [28, 25].

References

[1] Cyril Banderier. *Combinatoire analytique des chemins et des cartes*. Ph.D. thesis, Université Paris VI, 2001.

[2] Cyril Banderier. Limit laws for basic parameters of lattice paths with unbounded jumps. In *Chauvin, Brigitte (ed.) et al., Mathematics and computer science II.*, pages 33–47. Birkhäuser. Trends in Mathematics., 2002.

[3] Cyril Banderier, Mireille Bousquet-Mélou, Alain Denise, Philippe Flajolet, and Danièle Gardy. Generating functions for generating trees. *Discrete Math.*, 246(1-3):29–55, 2002.

[4] Cyril Banderier and Philippe Flajolet. Basic analytic combinatorics of directed lattice paths. *Theor. Comput. Sci.*, 281(1-2):37–80, 2002.

[5] Mireille Bousquet-Mélou and Marko Petkovšek. Linear recurrences with constant coefficients: The multivariate case. *Discrete Math.*, 225(1-3):51–75, 2000.

[6] P. Chassaing and G. Louchard. Phase transition for parking blocks, Brownian excursion and coalescence. *Random Struct. Algorithms*, 21(1):76–119, 2002.

[7] Philippe Chassaing and Philippe Flajolet. Hashing, trees, walks and graphs. (Hachage, arbres, chemins & graphes). *Gaz. Math., Soc. Math. Fr.*, 95:29–49, 2003.

[8] Philippe Chassaing and Guy Louchard. Reflected Brownian bridge area conditioned on its local time at the origin. *J. Algorithms*, 44(1):29–51, 2002.

[9] Fan Chung, Ron Graham, John Morrison, and Andrew Odlyzko. Peppling a chessboard. *Am. Math. Mon.*, 102(2):113–123, 1995.

[10] G.M. Constantine and T.H. Savits. A multivariate Faa di Bruno formula with applications. *Trans. Am. Math. Soc.*, 348(2):503–520, 1996.

[11] Philippe Duchon. *Q*-grammars and wall polyominoes. *Ann. Comb.*, 3(2-4):311–321, 1999.

[12] Philippe Duchon. On the enumeration and generation of generalized Dyck words. *Discrete Math.*, 225(1-3):121–135, 2000.

[13] Guy Fayolle, Roudolf Iasnogorodski, and Vadim Malyshev. *Random walks in the quarter-plane. Algebraic methods, boundary value problems and applications.* Applications of Mathematics. 40. Berlin: Springer. xvi, 160 p. , 1999.

[14] P. Flajolet, P. Poblete, and A. Viola. On the analysis of linear probing hashing. *Algorithmica*, 22(4):490–515, 1998.

[15] Philippe Flajolet and Robert Sedgewick. *Analytic Combinatorics*. `http://algo.inria.fr/flajolet/Publications/books.html`, 2006.

[16] Donald Ervin Knuth. *The Art of Computer Programming*, volume 1. Addison-Wesley, third edition, 1997.

[17] Germain Kreweras. Aires des chemins surdiagonal et application à un problème économique. *Cahiers du Bulletin Universitaire de Recherche Opérationnelle*, 24:1–8, 1976.

[18] I. A. Kurkova and Y. M. Suhov. Malyshev's theory and JS-queues. Asymptotics of stationary probabilities. *Ann. Appl. Probab.*, 13(4):1313–1354, 2003.

[19] Jacques Labelle and Yeong-Nan Yeh. Generalized Dyck paths. *Discrete Math.*, 82(1):1–6, 1990.

[20] G. Louchard. Kac's formula, Levy's local time and Brownian excursion. *J. Appl. Probab.*, 21:479–499, 1984.

[21] Amram Meir and John W. Moon. On the altitude of nodes in random trees. *Canadian Journal of Mathematics*, 30:997–1015, 1978.

[22] D. Merlini, R. Sprugnoli, and M. C. M. C. Verri. The area determined by underdiagonal lattice paths. *Proceedings of CAAP'96, Lecture Notes in Computer Science 1059*, pages 59–71, 1996.

[23] Donatella Merlini. Generating functions for the area below some lattice paths. In *Banderier, Cyril & Krattenthaler, Christian (eds.), Proceedings of the conference Discrete random walks'2003*, pages 217–228. Discrete Math. Theor. Comput. Sci., Proc. AC, 2003.

[24] Michel Nguyên Thê. Area of Brownian motion with generating functionology. In *Banderier, Cyril & Krattenthaler, Christian (eds.), Proceedings of the conference Discrete random walks'2003*, pages 229–242. Discrete Math. Theor. Comput. Sci., Proc. AC, 2003.

[25] A. Panholzer and H. Prodinger. Towards a more precise analysis of an algorithm to generate binary trees: A tutorial. *The Computer Journal*, 41:201–204, 1998.

[26] E. Pergola, R. Pinzani, S. Rinaldi, and R.A. Sulanke. A bijective approach to the area of generalized Motzkin paths. *Adv. Appl. Math.*, 28(3-4):580–591, 2002.

[27] T. Prellberg. Uniform q-series asymptotics for staircase polygons. *J. Phys. A, Math. Gen.*, 28(5):1289–1304, 1995.

[28] Christoph Richard. Area distribution of the planar random loop boundary. *J. Phys. A, Math. Gen.*, 37(16):4493–4500, 2004.

[29] Martin Rubey. Transcendence of generating functions of walks on the slit plane. In *Drmota (ed.) et al., Mathematics and computer science III.*, pages 49–58. Birkhäuser. Trends in Mathematics., 2004.