

LATTICE PATHS, SAMPLING WITHOUT REPLACEMENT, AND THE KERNEL METHOD

MARKUS KUBA, ALOIS PANHOLZER, AND HELMUT PRODINGER

ABSTRACT. In this work we consider weighted lattice paths in the quarter plane $\mathbb{N}_0 \times \mathbb{N}_0$. The steps are given by $(m, n) \rightarrow (m-1, n)$, $(m, n) \rightarrow (m, n-1)$ and are weighted as follows: $(m, n) \rightarrow (m-1, n)$ by $m/(m+n)$ and step $(m, n) \rightarrow (m, n-1)$ by $n/(m+n)$. The considered lattice paths are absorbed at lines $y = x/t - s/t$ with $t \in \mathbb{N}$ and $s \in \mathbb{N}_0$. We provide explicit formulæ for the sum of the weights of paths, starting at (m, n) , which are absorbed at a certain height k at lines $y = x/t - s/t$ with $t \in \mathbb{N}$ and $s \in \mathbb{N}_0$, using a generating functions approach and the kernel method. Furthermore these weighted lattice paths can be interpreted as probability distributions arising in the context of Pólya-Eggenberger urn models, more precisely, the lattice paths are sample paths of the well known sampling without replacement urn. We provide limiting distribution results for the underlying random variable, surprisingly obtaining a total of five phase changes.

1. INTRODUCTION

1.1. Lattice paths. Let $\mathcal{S} \subseteq \mathbb{N}_0 \times \mathbb{N}_0$ denote a set of lattice points in the quarter plane¹. We consider lattice paths with steps $(m, n) \rightarrow (m-1, n)$ and $(m, n) \rightarrow (m, n-1)$, starting at $(m, n) \in \mathcal{S}$. The steps are weighted as follows: step $(m, n) \rightarrow (m-1, n)$ is weighted by $m/(m+n)$ and step $(m, n) \rightarrow (m, n-1)$ is weighted by $n/(m+n)$. We are interested in weighted lattice paths starting at $(m, n) \in \mathcal{S}$, which touch or cross a certain line $y = f(x)$ at height k , with $0 \leq k \leq n$ and $\mathcal{S} = \mathcal{S}_y$ suitably defined. We consider the following types of lines: $y = x/t - s/t$, with $t \in \mathbb{N}$ and $s \in \mathbb{N}_0$. In this case we have $\mathcal{S} = \{(m, n) \mid m \geq tn+s, n \in \mathbb{N}_0\}$. We are interested in the sum of the weights of all paths starting at $(m, n) \in \mathcal{S}$ and touching $y = x/t - s/t$ at height k , with $0 \leq k \leq n$, not touching the line y before, which we call *absorption at line y* .

From a probabilistic point of view we can interpret the desired numbers as probability distributions of a random variable $X_{m,n}$, which can be described as follows. A particle is

Date: April 11, 2008.

2000 Mathematics Subject Classification. 05A15, 60C05.

Key words and phrases. Lattice paths, Sampling without replacement, urn models, kernel method, Lévy distribution.

This work was supported by the Austrian Science Foundation FWF, grant S9608-N13 and by the South African Science Foundation NRF, grant 2053748. The second author wants to thank the Department of Mathematical Sciences, University of Stellenbosch, for its support and hospitality during a research visit, where a part of this work has been carried out.

¹Throughout this work we use the notations $\mathbb{N} := \{1, 2, 3, \dots\}$ and $\mathbb{N}_0 := \{0, 1, 2, \dots\}$.

located at a certain point $(m, n) \in \mathcal{S}$, and moves randomly to the left or downwards with probabilities depending on the actual position:

$$\mathbb{P}\{(m, n) \rightarrow (m-1, n)\} = \frac{m}{m+n}, \quad \mathbb{P}\{(m, n) \rightarrow (m, n-1)\} = \frac{n}{m+n}.$$

The random variable $X_{m,n}$ describes the height k at which a particle starting at (m, n) is absorbed, i.e., where it is touching or crossing a line $y = f(x)$ for the first time. The searched probability $\mathbb{P}\{X_{m,n} = k\}$ is then equal to the sum of the weights of all lattice paths, starting at (m, n) , which touch or cross the line $y = f(x)$ at height k . We can also formulate this problem in the context of certain urn models.

1.2. Pólya-Eggenberger urn models and the sampling without replacement urn. Pólya-Eggenberger urn models are defined as follows. We start with an urn containing n white balls and m black balls. The evolution of the urn occurs in discrete time steps. At every step a ball is chosen at random from the urn. The color of the ball is inspected and then the ball is reinserted into the urn. According to the observed color of the ball, balls are added/removed due to the following rules. If we have chosen a white ball, we put into the urn a white balls and b black balls, but if we have chosen a black ball, we put into the urn c white balls and d black balls. The values $a, b, c, d \in \mathbb{Z}$ are fixed integer values and the urn model is specified by the 2×2 ball replacement matrix $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

One of the most fundamental urn models is the so-called *sampling without replacement urn*, associated with the ball replacement matrix $M = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$. In this urn model a parameter of interest is the number $X_{m,n}$ of remaining white balls, after all black balls have been removed, starting with n white and m black balls. The formal setting is as follows. We have a state space \mathcal{S} , which is given by $\mathcal{S} := \{(m, n) \mid m, n \in \mathbb{N}_0\}$. Further we have a set of absorbing states $\mathcal{A} := \{(0, n) \mid n \in \mathbb{N}_0\}$, where the evolution of the urn stops.

The problem of absorption at line y can also be formulated in the context of Pólya-Eggenberger urn models, where the state space \mathcal{S} and the set of absorbing states \mathcal{A} are suitably modified. E. g. for $y = x/t - s/t$ we consider a sampling urn with ball replacement matrix $M = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$, where the state space is given by $\mathcal{S} = \{(m, n) \mid m \geq tn + s, n \in \mathbb{N}_0\}$ and the set of absorbing states by $\mathcal{A} = \{(tm + s, m) \mid m \in \mathbb{N}_0\}$.

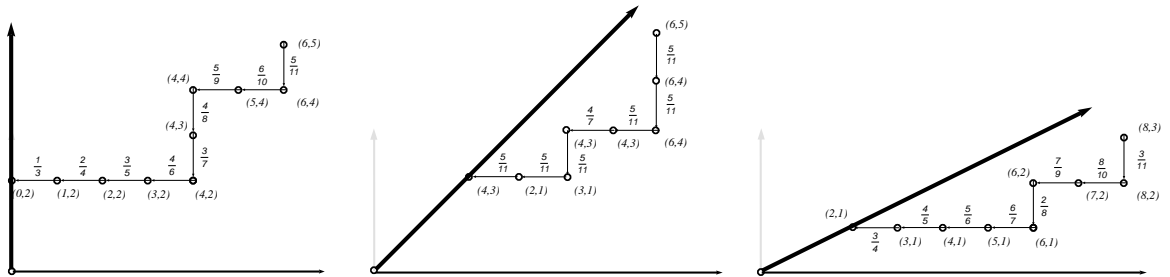


FIGURE 1. Sample paths for absorption at $x = 0$, $y = x$ and $y = x/2$.

1.3. Motivation and related work. The main motivation for this study is to combine the different areas of lattice path enumeration, see, i.e., Mohanty [10], Banderier and Flajolet [2] and Pólya-Eggenberger urn models, see Flajolet et al. [3], [4], and Hwang et al. [7]. Interestingly, for the weighted lattice paths and absorbing lines studied, we obtain closed formulæ for the probability of absorption at height k . This explicit results also allows a detailed study of the limiting behaviour of the random variable $X_{m,n}$. We could completely characterize the limit laws of $X_{m,n}$ and phase changes appearing depending on the growth behaviour of m and n of the starting point (m, n) .

The problem studied was also motivated by a combinatorial game involving card guessing, which has been analyzed by Levasseur [8], Zagier [13], Sulanke [12] and Knopfmacher and Prodinger [6]. One starts with a deck consisting of m red and n black cards. A guess is made as to the color of the top card, after which it is revealed and discarded. To maximize the number of correct guesses one chooses the color corresponding to the majority of cards remaining in the deck. We will revisit this problem and provide limiting distribution results.

Our analysis is based on a generating functions approach: we will derive functional equations for suitably defined generating functions and use then the kernel method (see Prodinger [11] for a survey about this method, and, i.e., the works of Banderier et al. [2], [1] for applications) in order to obtain explicit formulæ for the probability functions of the considered random variables.

1.4. Results. Here we collect the exact and asymptotic results obtained for the weighted lattice paths studied.

Theorem 1. *The probability $\mathbb{P}\{X_{m,n} = k\}$ that a particle starting at (m, n) is absorbed on the line $y = x/t - s/t$ at height k is given by the following explicit formula:*

$$\mathbb{P}\{X_{m,n} = k\} = \frac{\binom{(t+1)k+s}{k} \left(\binom{m+n-1-s-(t+1)k}{n-k} - t \binom{m+n-1-s-(t+1)k}{n-k-1} \right)}{\binom{m+n}{n}},$$

for $m \geq tn + s$ and $0 \leq k \leq n$, with $t \in \mathbb{N}$, $s \in \mathbb{N}_0$.

In the next theorem we state the obtained limiting distribution results of $X_{m,n}$, depending on the growth of m and n , for absorbing lines $y = x/t - s/t$, with $t \in \mathbb{N}$, $s \in \mathbb{N}_0$ fixed. We use here the notation $Y_n \xrightarrow{\mathcal{L}} Y$ for the weak convergence, i.e., the convergence in distribution, of a sequence of random variables Y_n to a random variable Y . Furthermore we use $Y \stackrel{\mathcal{L}}{=} Z$ for the equality in distribution of two random variables Y and Z .

Theorem 2. *The limiting distribution behaviour of $X_{m,n}$ is, for $m \rightarrow \infty$ and depending on the growth of $n = n(m)$, described as follows.*

(1) $n = o(m)$: *The random variable $X_{m,n}$ is asymptotically zero, as m tends to infinity:*

$$\mathbb{P}\{X_{m,n} = 0\} = \frac{(m-s-tn) \binom{m}{s}}{(m+n-s) \binom{m+n}{s}} \rightarrow 1.$$

- (2) $n \sim \rho m$, such that $0 < \rho < \frac{1}{t}$: The random variable $X_{m,n}$ weakly converges to a discrete random variable X_ρ :

$$X_{m,n} \xrightarrow{\mathcal{L}} X_\rho, \quad \text{with} \quad \mathbb{P}\{X_\rho = k\} = \frac{1 - t\rho}{1 + \rho} \cdot \binom{(t+1)k + s}{k} \frac{\rho^k}{(1 + \rho)^{(t+1)k + s}}, \quad k \geq 0.$$

- (3) $n \sim \frac{m}{t}$, such that $n = \frac{m}{t} - \frac{s}{t} - \ell$, with $m = o(\ell^2)$: The scaled random variable $\frac{\ell^2 t^2}{m^2} X_{m,n}$ is asymptotically Gamma distributed:

$$\frac{\ell^2 t^2}{m^2} X_{m,n} \xrightarrow{\mathcal{L}} X, \quad \text{with} \quad X \stackrel{\mathcal{L}}{=} \gamma\left(\frac{1}{2}, \frac{2(t+1)}{t}\right),$$

where $\gamma(k, \theta)$ denotes a Gamma distribution with shape parameter k and scale parameter θ . The density $f(x)$ of X is given by

$$f(x) = \frac{\sqrt{t}}{\sqrt{2(t+1)}\sqrt{\pi}\sqrt{x}} e^{-\frac{xt}{2(t+1)}}, \quad x > 0.$$

- (4) $n \sim \frac{m}{t}$, such that $n = \frac{m}{t} - \frac{s}{t} - \ell$, with $\ell \sim \rho\sqrt{m}$ and $\rho > 0$: The scaled random variable $\frac{t}{m} X_{m,n}$ weakly converges to X_ρ :

$$\frac{t}{m} X_{m,n} \xrightarrow{\mathcal{L}} X_\rho, \quad \text{with density} \quad f_\rho(x) = \frac{\rho t}{\sqrt{2\pi x(1+t)}(1-x)^{\frac{3}{2}}} e^{-\frac{\rho^2 t^2 x}{2(1+t)(1-x)}}, \quad 0 < x < 1.$$

- (5) $n \sim \frac{m}{t}$, such that $n = \frac{m}{t} - \frac{s}{t} - \ell$, with $\ell \rightarrow \infty$ and $\ell = o(\sqrt{m})$: The shifted and scaled random variable $\frac{1}{\ell^2}(\frac{m}{t} - X_{m,n})$ is asymptotically Lévy distributed:

$$\frac{1}{\ell^2} \left(\frac{m}{t} - X_{m,n} \right) \xrightarrow{\mathcal{L}} X, \quad \text{with} \quad X \stackrel{\mathcal{L}}{=} \text{Lévy}\left(\frac{t}{t+1}\right),$$

where Lévy(c) denotes a Lévy distribution with parameter c . The density $f(x)$ of X is given by

$$f(x) = \frac{\sqrt{t}}{\sqrt{2\pi(t+1)}x^{\frac{3}{2}}} e^{-\frac{t}{2x(t+1)}}, \quad x > 0.$$

- (6) $n \sim \frac{m}{t}$, such that $n \sim \frac{m}{t} - \frac{s}{t} - \ell$, with $\ell \in \mathbb{N}$ fixed: The shifted random variable $\frac{m}{t} - \frac{s}{t} - X_{m,n}$ weakly converges to X_ℓ :

$$\frac{m}{t} - \frac{s}{t} - X_{m,n} \xrightarrow{\mathcal{L}} X_\ell, \quad \text{with} \quad \mathbb{P}\{X_\ell = k\} = \frac{t^{tk}}{(1+t)^{tk-\ell}} \binom{(t+1)k - \ell - 1}{k - \ell} \frac{\ell}{k}, \quad k \geq \ell.$$

Remark 1. The Lévy distribution is a stable distribution. It is a special case of the Lévy skew alpha-stable distribution, which does not generally have an analytically expressible probability density. Furthermore the moments of the Lévy distribution do not exist. Hence, for $n \sim m/t$, such that $n = m/t - s/t - \ell$, with $\ell \rightarrow \infty$ and $\ell = o(\sqrt{m})$, the random variable $(m/t - X_{m,n})/\ell^2$ converges in distribution, but without convergence of any integer moment.

2. LINES $y = x/t - s/t$, WITH $t \in \mathbb{N}$, $s \in \mathbb{N}_0$

2.1. Recurrences. Let $\varphi_{m,n}(v) = \sum_{k \geq 0} \mathbb{P}\{X_{m,n} = k\}v^k$ denote the probability generating function of $X_{m,n}$, where $X_{m,n} = X_{m,n}(s, t)$. We usually drop the dependence of $X_{m,n}$ on s and t for the sake of simplicity. By using the basic decomposition of the paths according to the first step and taking into account the absorbing lines, the problem can be translated into the following recurrence:

$$\begin{aligned} \varphi_{m,n}(v) &= \frac{m}{m+n}\varphi_{m-1,n}(v) + \frac{n}{m+n}\varphi_{m,n-1}(v), \quad \text{for } m > tn + s, \quad n \geq 1, \\ \varphi_{tm+s,m}(v) &= v^m, \quad \text{for } m \geq 0, \quad \varphi_{m,0}(v) = 1, \quad \text{for } m \geq s. \end{aligned}$$

This recurrence will be treated by the introducing the normalized functions

$$\Phi_{m,n}(v) = \binom{m+n}{m} \varphi_{m,n}(v).$$

We obtain

$$\Phi_{m,n}(v) = \Phi_{m-1,n}(v) + \Phi_{m,n-1}(v), \quad \text{for } m > tn + s \geq 1, \quad (1a)$$

$$\Phi_{tm+s,m}(v) = \binom{(t+1)m+s}{m} v^m, \quad \text{for } m \geq 0, \quad \Phi_{m,0}(v) = 1, \quad \text{for } m \geq s. \quad (1b)$$

2.2. Generating functions. We introduce the trivariate generating function

$$F(z, u, v) = \sum_{n \geq 0} \sum_{m \geq tn+s} \Phi_{m,n}(v) z^m u^{m-tn-s},$$

and also $t+1$ auxiliary functions $F_k(z, v)$ defined by

$$F_k(z, v) = \sum_{n \geq 0} \Phi_{tn+s+k,n}(v) z^{tn+s+k}, \quad \text{for } 0 \leq k \leq t.$$

Due to (1b) the generating function $F_0(z, v)$ is already known:

$$F_0(z, v) = z^s \sum_{n \geq 0} \Phi_{tn+s,n}(v) z^{tn} = z^s \sum_{n \geq 0} \binom{(t+1)n+s}{n} (vz^t)^n. \quad (2)$$

Using (1a) we obtain the following functional equation for $F(z, u, v)$:

$$\left(1 - zu - \frac{1}{u^t}\right) F(z, u, v) = \left(1 - \frac{1}{u^t}\right) F_0(z, v) - \frac{1}{u^t} \sum_{k=1}^t u^k F_k(z, v). \quad (3)$$

It is advantageous to write equation (3) in the following form:

$$(zu^{t+1} - u^t + 1) F(z, u, v) = (1 - u^t) F_0(z, v) + \sum_{k=1}^t u^k F_k(z, v). \quad (4)$$

2.3. The kernel method. Equation (4) gives a simple relation between the unknown functions $F(z, u, v)$ and $F_k(z, v)$, $1 \leq l \leq t$. In order to solve (4) we will apply the kernel method. To do this we will consider the so-called characteristic equation

$$P(z, u) = zu^{t+1} - u^t + 1 = 0.$$

By general considerations on the roots of the characteristic polynomial $P(z, u)$ as figured out in [2], it follows that $P(z, u)$ can be written in the following form:

$$P(z, u) = (zu - \lambda_{t+1}(z))(u - \lambda_1(z))(u - \lambda_2(z)) \cdots (u - \lambda_t(z)),$$

with functions $\lambda_1(z), \dots, \lambda_{t+1}(z)$ analytic around $z = 0$. In the following we use the abbreviation $\lambda_i := \lambda_i(z)$, $1 \leq i \leq t + 1$, where we do not express explicitly the dependence of λ_i on z .

Now we use the fact that $F(z, u, v)$ is an analytic function in a neighbourhood of $z = 0$. Thus we can evaluate $F(z, u, v)$ and therefore equation (4) at $u = \lambda_1(z), \dots, u = \lambda_t(z)$ for z in a neighbourhood of 0. Since $P(z, \lambda_i(z)) = 0$, for $1 \leq i \leq t$, we obtain from (4) after plugging in $u = \lambda_i$, $1 \leq i \leq t$, the following system of linear equations for the unknown functions $F_k(z, v)$, $1 \leq k \leq t$:

$$\begin{aligned} (1 - \lambda_1^t)F_0(z, v) + \sum_{k=1}^t \lambda_1^k F_k(z, v) &= 0, \\ \vdots & \\ (1 - \lambda_t^t)F_0(z, v) + \sum_{k=1}^t \lambda_t^k F_k(z, v) &= 0. \end{aligned}$$

We rewrite this system of equations in matrix form:

$$\begin{pmatrix} \lambda_1 & \lambda_1^2 & \cdots & \lambda_1^t \\ \lambda_2 & \lambda_2^2 & \cdots & \lambda_2^t \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_t & \lambda_t^2 & \cdots & \lambda_t^t \end{pmatrix} \cdot \begin{pmatrix} F_1(z, v) \\ F_2(z, v) \\ \vdots \\ F_t(z, v) \end{pmatrix} = - \begin{pmatrix} 1 - \lambda_1^t \\ 1 - \lambda_2^t \\ \vdots \\ 1 - \lambda_t^t \end{pmatrix} F_0(z, v). \quad (5)$$

Applying Cramer's rule we can write the solution of this linear system of equations as a quotient of determinants, with $1 \leq k \leq t$:

$$F_k(z, v) = - \begin{vmatrix} \lambda_1 & \cdots & \lambda_1^{k-1} & 1 - \lambda_1^t & \lambda_1^{k+1} & \cdots & \lambda_1^t \\ \lambda_2 & \cdots & \lambda_2^{k-1} & 1 - \lambda_2^t & \lambda_2^{k+1} & \cdots & \lambda_2^t \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \lambda_t & \cdots & \lambda_t^{k-1} & 1 - \lambda_t^t & \lambda_t^{k+1} & \cdots & \lambda_t^t \end{vmatrix} \cdot \begin{vmatrix} \lambda_1 & \lambda_1^2 & \cdots & \lambda_1^t \\ \lambda_2 & \lambda_2^2 & \cdots & \lambda_2^t \\ \vdots & \vdots & \vdots & \vdots \\ \lambda_t & \lambda_t^2 & \cdots & \lambda_t^t \end{vmatrix}^{-1} F_0(z, v). \quad (6)$$

To evaluate the quotient of determinants appearing in (6) we distinguish between the cases $1 \leq k \leq t - 1$ and $k = t$. If $1 \leq k \leq t - 1$ we can add the t -th row to the k -th row in the

determinant appearing in the numerator of (6) and obtain after simple manipulations:

$$F_k(z, v) = \frac{(-1)^k}{\lambda_1 \lambda_2 \cdots \lambda_t} \begin{vmatrix} 1 & \lambda_1 & \cdots & \lambda_1^{k-1} & \lambda_1^{k+1} & \cdots & \lambda_1^t \\ 1 & \lambda_2 & \cdots & \lambda_2^{k-1} & \lambda_2^{k+1} & \cdots & \lambda_2^t \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_t & \cdots & \lambda_t^{k-1} & \lambda_t^{k+1} & \cdots & \lambda_t^t \end{vmatrix} \cdot \begin{vmatrix} 1 & \lambda_1 & \cdots & \lambda_1^{t-1} \\ 1 & \lambda_2 & \cdots & \lambda_2^{t-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_t & \cdots & \lambda_t^{t-1} \end{vmatrix}^{-1} F_0(z, v). \quad (7)$$

2.4. Elementary symmetric functions and the Jacobi-Trudi identity. The quotient of determinants appearing in (7) can be evaluated as elementary symmetric functions, which can be seen, e.g., by writing the determinants as Schur functions and applying the Jacobi-Trudi identity.

A partition is defined as a sequence $\alpha = (\alpha_1, \dots, \alpha_n)$ of non-negative integers in decreasing order, that is $\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_n$. The non-zero α_i are called the parts of α . Their number is the length of α and will be denoted by $l(\alpha)$. Further, given a partition α , the conjugate of the partition α is the partition α' defined by the sequence $\alpha'_i = |\{j : \alpha_j \geq i\}|$. Given a partition $\alpha = (\alpha_1, \dots, \alpha_n)$, the Schur symmetric function s_α is the symmetric function defined by the quotient

$$s_\alpha(x_1, \dots, x_n) = \frac{\det(x_i^{\alpha_j + n - j})_{1 \leq i, j \leq n}}{\det(x_i^{n - j})_{1 \leq i, j \leq n}}.$$

We will use the following form of the Jacobi-Trudi identity (see [9]).

$$s_\alpha(x_1, \dots, x_n) = \det(e_{\alpha'_i - i + j}(x_1, \dots, x_n))_{1 \leq i, j \leq n}, \quad \text{with } e_k = 0, \quad \text{for } k < 0 \text{ or } k > n. \quad (8)$$

Here $e_r(x_1, \dots, x_n)$ denotes the r -th elementary symmetric function with variables x_1, \dots, x_n :

$$e_r(x_1, \dots, x_n) = \sum_{1 \leq i_1 < i_2 < \cdots < i_r \leq n} x_{i_1} x_{i_2} \cdots x_{i_r}, \quad \text{with } 0 \leq r \leq n.$$

Hence we may express the functions $F_k(z, v)$ in terms of Schur functions using the partition $\alpha = (1, \dots, 1, 0, \dots, 0)$, with $\alpha_i = 1$ for $1 \leq i \leq t - k$ and $\alpha_i = 0$ for $t + 1 - k \leq i \leq t$:

$$F_k(z, v) = \frac{(-1)^k}{\lambda_1 \lambda_2 \cdots \lambda_t} s_\alpha(\lambda_1, \dots, \lambda_t) F_0(z, v). \quad (9)$$

Since $\alpha' = (t - k, 0, \dots, 0)$ we immediately obtain by applying the Jacobi-Trudi identity the following form of the solution:

$$F_k(z, v) = \frac{(-1)^k}{\lambda_1 \lambda_2 \cdots \lambda_t} e_{t-k}(\lambda_1, \lambda_2, \dots, \lambda_t) F_0(z, v), \quad 1 \leq k \leq t - 1. \quad (10)$$

For the case $k = t$ we can split the t -th row in the determinant appearing in the numerator of (6) and obtain easily:

$$F_t(z, v) = \left(\frac{(-1)^t}{\lambda_1 \lambda_2 \cdots \lambda_t} + 1 \right) F_0(z, v). \quad (11)$$

Thus we can write the solution of $F(z, u, v)$ as

$$\begin{aligned} F(z, u, v) &= \frac{\left((1 - u^t) + \sum_{k=1}^{t-1} u^k \frac{(-1)^k}{\lambda_1 \lambda_2 \cdots \lambda_t} e_{t-k}(\lambda_1, \dots, \lambda_t) + u^t \left(\frac{(-1)^k}{\lambda_1 \lambda_2 \cdots \lambda_t} + 1 \right) \right) F_0(z, v)}{zu^{t+1} - u^t + 1} \\ &= \frac{\left(\lambda_1 \cdots \lambda_t (-1)^t + \sum_{k=1}^{t-1} u^k (-1)^{t-k} e_{t-k}(\lambda_1, \dots, \lambda_t) + u^t \right) F_0(z, v)}{(-1)^t \lambda_1 \cdots \lambda_t (zu^{t+1} - u^t + 1)} \\ &= \frac{\left(\sum_{k=0}^t u^k (-1)^{t-k} e_{t-k}(\lambda_1, \dots, \lambda_t) \right) F_0(z, v)}{(-1)^t \lambda_1 \cdots \lambda_t (zu^{t+1} - u^t + 1)}. \end{aligned}$$

Now we use again the factorization of $P(z, u)$, which gives:

$$\begin{aligned} P(z, u) &= zu^{t+1} - u^t + 1 = (zu - \lambda_{t+1})(u - \lambda_1)(u - \lambda_2) \cdots (u - \lambda_t) \\ &= (zu - \lambda_{t+1}) \sum_{k=0}^t u^k (-1)^{t-k} e_{t-k}(\lambda_1, \dots, \lambda_t). \end{aligned}$$

Thus we get further:

$$F(z, u, v) = \frac{(-1)^t}{\lambda_1 \cdots \lambda_t (zu - \lambda_{t+1})} F_0(z, v) = \frac{\lambda_{t+1}}{\lambda_{t+1} - zu} F_0(z, v), \quad (12)$$

where we have used the identity $\lambda_1 \cdots \lambda_{t+1} = (-1)^{t+1}$, which is again a direct consequence of the factorization of $P(z, u)$.

2.5. An alternative approach. Starting with equation (6) one may also proceed differently, in a less standard manner, by using the following argument, which is based on arguments of Mireille Bousquet-Mélou [1]. In order to obtain $F(z, u, v)$ we derive at first $F_t(z, v)$. As remarked above we can split the t -th row in the determinant appearing in the numerator of (6) and obtain easily:

$$F_t(z, v) = \left(\frac{(-1)^t}{\lambda_1 \lambda_2 \cdots \lambda_t} + 1 \right) F_0(z, v). \quad (13)$$

Now let $N(z, u)$ denote the right-hand-side of the functional equation (4) for $F(z, u, v)$:

$$N(z, u) := (1 - u^t) F_0(z, v) + \sum_{k=1}^t u^k F_k(z, v)$$

The quantity $N(z, u)$ is a polynomial in u with leading coefficient $F_t(z, v) - F_0(z, v)$, whose zeros are exactly $\lambda_1, \dots, \lambda_t$. Hence, after normalization, we have a leading monomial u^t , and the normalized polynomial factors nicely into the following expression:

$$\frac{N(z, u)}{F_t(z, v) - F_0(z, v)} = \prod_{k=1}^t (u - \lambda_k).$$

Since

$$(-1)^{t+1} \lambda_1 \cdots \lambda_t \lambda_{t+1} = 1,$$

we get

$$\begin{aligned} F_t(z, v) - F_0(z, v) &= \left(\frac{(-1)^t}{\lambda_1 \cdots \lambda_t} + 1 \right) F_0(z, v) - F_0(z, v) = \frac{(-1)^t}{\lambda_1 \cdots \lambda_t} F_0(z, v) \\ &= -\lambda_{t+1} F_0(z, v). \end{aligned}$$

Hence we finally obtain:

$$F(z, u, v) = \frac{N(z, u)}{P(z, u)} = \frac{-\lambda_{t+1} F_0(z, v) \prod_{k=1}^t (u - \lambda_k)}{(zu - \lambda_{t+1}) \prod_{k=1}^t (u - \lambda_k)} = \frac{\lambda_{t+1}}{\lambda_{t+1} - zu} F_0(z, v). \quad (14)$$

2.6. Extracting coefficients. To obtain the required probabilities we only have to extract coefficients of (12). By using the definition of $F(z, u, v)$ we get

$$\begin{aligned} \mathbb{P}\{X_{m,n} = k\} &= \frac{1}{\binom{m+n}{m}} [z^m u^{m-tn-s} v^k] F(z, u, v) = \frac{1}{\binom{m+n}{m}} [z^{tn+s} (uz)^{m-tn-s} v^k] \frac{F_0(z, v)}{1 - \frac{zu}{\lambda_{t+1}}} \\ &= \frac{1}{\binom{m+n}{m}} [z^{tn+s} v^k] \frac{F_0(z, v)}{\lambda_{t+1}^{m-tn-s}} = \frac{1}{\binom{m+n}{m}} [z^{tn} v^k] \frac{\sum_{l \geq 0} \binom{(t+1)l+s}{l} (vz^t)^l}{\lambda_{t+1}^{m-tn-s}} \\ &= \frac{\binom{(t+1)k+s}{k}}{\binom{m+n}{m}} [z^{t(n-k)}] \frac{1}{\lambda_{t+1}(z)^{m-tn-s}}. \end{aligned}$$

To extract coefficients from this expression we consider the characteristic equation $zu^{t+1} - u^t + 1 = 0$. Multiplying with z^t and using the substitution $\lambda := zu$ leads then to the equation

$$z^t = \lambda^t (1 - \lambda).$$

Of course, $\lambda = \lambda_{t+1}(z)$ is exactly the function implicitly defined by this equation, which satisfies $\lambda_{t+1}(0) = 1$. To apply the Lagrange inversion formula we introduce the substitutions $\tilde{z} := z^t$ and $w := \lambda - 1$ leading to the following equation, which is suitable for that:

$$\tilde{z} = -(1 + w)^t w.$$

Thus we obtain further:

$$\begin{aligned} \mathbb{P}\{X_{m,n} = k\} &= \frac{\binom{(t+1)k+s}{k}}{\binom{m+n}{m}} [\tilde{z}^{n-k}] \frac{1}{(1+w)^{m-tn-s}} \\ &= \frac{\binom{(t+1)k+s}{k}}{\binom{m+n}{m}} \frac{1}{n-k} [w^{n-k-1}] \left(\frac{1}{(1+w)^{m-tn-s}} \right)' \left(-\frac{1}{(1+w)^t} \right)^{n-k} \\ &= \frac{\binom{(t+1)k+s}{k}}{\binom{m+n}{m}} \frac{(-1)^{n-k-1} (m-s-tn)}{n-k} [w^{n-k-1}] \frac{1}{(1+w)^{m+1-tk-s}} \\ &= \frac{\binom{(t+1)k+s}{k}}{\binom{m+n}{m}} \frac{m-s-tn}{n-k} \binom{m-s+n-1-(t+1)k}{n-k-1} \\ &= \frac{\binom{(t+1)k+s}{k}}{\binom{m+n}{m}} \frac{m-s-tn}{m-s+n-(t+1)k} \binom{m-s+n-(t+1)k}{n-k}, \end{aligned}$$

for $m \geq tn + s$ and $0 \leq k \leq n$. Now we rewrite the probabilities obtained in the following form, which is given in Theorem 1:

$$\mathbb{P}\{X_{m,n} = k\} = \frac{\binom{(t+1)k+s}{k} \left(\binom{m+n-1-s-(t+1)k}{n-k} - t \binom{m+n-1-s-(t+1)k}{n-k-1} \right)}{\binom{m+n}{n}},$$

for $m \geq tn + s$, with $0 \leq k \leq n$.

Remark 2. Due to the simplicity of the result concerning the lines $y = x/t - s/t$ it is natural to ask for a more direct derivation of the probabilities. In the following we will sketch such an alternative combinatorial derivation. It is well known that the number of (unweighted) lattice paths from (m, n) to the origin $(0, 0)$ with unit steps to the left or downwards is given by $\binom{m+n}{m}$. We need the following result, which can be found in Mohanty [10].

Lemma 1. *The number of (unweighted) lattice paths from the origin to (m, n) , which never pass above the line $y = x/t$, is given by*

$$\frac{m - tn + 1}{m + n + 1} \binom{m + n + 1}{n} = \binom{m + n}{n} - t \binom{m + n}{n - 1}.$$

Now we obtain the probabilities $\mathbb{P}\{X_{m,n} = k\}$ by fixing the last step and using Lemma 1:

$$\mathbb{P}\{X_{m,n} = k\} = \mathbb{P}\{(m, n) \rightarrow (tk + s, k) \mid y = (x - s)/t \text{ is not touched}\}.$$

By fixing the last step we get further

$$\begin{aligned} \mathbb{P}\{X_{m,n} = k\} &= \mathbb{P}\{(m, n) \rightarrow (tk + s + 1, k) \mid y = (x - s)/t \text{ is not touched}\} \\ &\quad \times \mathbb{P}\{(tk + s + 1, k) \rightarrow (tk + s, k)\}. \end{aligned}$$

Furthermore,

$$\begin{aligned} \mathbb{P}\{X_{m,n} = k\} &= \frac{\#\text{Paths}((m - tk - s - 1, n - k) \rightarrow (0, 0) \mid y = x/t \text{ is not passed})}{\#\text{Paths}((m, n) \rightarrow (0, 0))} \\ &\quad \times \#\text{Paths}((tk + s + 1, k) \rightarrow (0, 0)) \cdot \frac{tk + s + 1}{(t + 1)k + s + 1} \\ &= \frac{\left(\binom{m+n-1-s-(t+1)k}{n-k} - t \binom{m+n-1-s-(t+1)k}{n-k-1} \right) \binom{(t+1)k+s+1}{k}}{\binom{m+n}{m}} \cdot \frac{tk + s + 1}{(t + 1)k + s + 1} \\ &= \frac{\left(\binom{m+n-1-s-(t+1)k}{n-k} - t \binom{m+n-1-s-(t+1)k}{n-k-1} \right) \binom{(t+1)k+s}{k}}{\binom{m+n}{m}}. \end{aligned}$$

2.7. Deriving the limiting distributions. The main results of this paper, which describe the limiting distribution of $X_{m,n}$ depending on the growth behaviour of m and n , are obtained from the probability mass function given in Theorem 1 after a careful application of Stirling's formula

$$n! = \sqrt{2\pi n} \left(\frac{n}{e}\right)^n \left(1 + \frac{1}{12n} + \frac{1}{288n^2} + \mathcal{O}\left(\frac{1}{n^3}\right)\right). \quad (15)$$

Note that one main difficulty is to “guess” the right normalizations required. Once the right guess was made, we still had to carry out quite lengthy and tedious calculations in order to obtain the stated results. Since these calculations are very lengthy, we will only present as an example the derivation of the local limit for the case $s = 0$ and $n \sim \frac{m}{t}$, such that $n = \frac{m}{t} - \ell$, with $\ell \sim \rho\sqrt{m}$ and $\rho > 0$. Other computations are carried out in a similar manner.

$$\begin{aligned}
\frac{m}{t} \mathbb{P}\left\{\frac{tX_{m,n}}{m} = x\right\} &\sim \frac{\rho t}{\sqrt{2\pi(t+1)x(1-x)}^{\frac{3}{2}}} \frac{\left(1 - \frac{\rho}{\sqrt{m(1+\frac{1}{t})(1-x)}}\right)^{(1+\frac{1}{t})(1-x)m - \rho\sqrt{m}}}{\left(1 - \frac{\rho t}{\sqrt{m(1-x)}}\right)^{(1-x)\frac{m}{t} - \rho\sqrt{m}}} \\
&\times \frac{\left(1 - \frac{\rho t}{\sqrt{m}}\right)^{\frac{m}{t} - \rho\sqrt{m}}}{\left(1 - \frac{\rho}{\sqrt{m(1+\frac{1}{t})}}\right)^{(1+\frac{1}{t})m - \rho\sqrt{m}}} \\
&\sim \frac{\rho t}{\sqrt{2\pi(t+1)x(1-x)}^{\frac{3}{2}}} \left(\frac{\left(1 - \frac{\rho t}{\sqrt{m(1-x)}}\right)\left(1 - \frac{\rho t}{\sqrt{m(1+t)}}\right)}{\left(1 - \frac{\rho t}{\sqrt{m}}\right)\left(1 - \frac{\rho t}{\sqrt{m(1-x)(1+t)}}\right)}\right)^{\rho\sqrt{m}} \times \\
&\times \left(\frac{\left(1 - \frac{\rho t}{\sqrt{m(1+t)(1-x)}}\right)^{t+1}}{\left(1 - \frac{\rho t}{\sqrt{m(1-x)}}\right)}\right)^{\frac{m}{t}(1-x)} \left(\frac{\left(1 - \frac{\rho t}{\sqrt{m}}\right)}{\left(1 - \frac{\rho t}{\sqrt{m(1+t)}}\right)}\right)^{\frac{m}{t}} \\
&\sim \frac{\rho t}{\sqrt{2\pi(t+1)x(1-x)}^{\frac{3}{2}}} \frac{e^{-\frac{\rho^2 t}{1-x}} e^{-\frac{\rho^2 t}{1+t}}}{e^{-\rho^2 t} e^{-\frac{\rho^2 t}{(1+t)(1-x)}}} \frac{\rho^2 t \binom{t+1}{2}}{e^{(1+t)^2(1-x)}} e^{-\frac{\rho^2 t \binom{t+1}{2}}{(1+t)^2}} \\
&\sim \frac{\rho t}{\sqrt{2\pi x(1+t)(1-x)}^{\frac{3}{2}}} e^{-\frac{\rho^2 t^2 x}{2(1+t)(1-x)}}.
\end{aligned}$$

3. LIMIT LAWS FOR THE CARD GUESSING GAME AGAIN REVISITED

One starts with a deck consisting of m red and n black cards. A guess is made as to the color of the top card, after which it is revealed and discarded. To maximize the number of correct guesses one chooses the color corresponding to the majority of cards remaining in the deck. Let $X_{m,n}$ denote the random variable counting the number of correct guesses starting with m red and n black cards. The following result was obtained by Sulanke, and also by Knopfmacher and Prodinger.

Theorem 3 (Sulanke; Knopfmacher and Prodinger). *The exact distribution of the random variable $X_{m,n}$ counting the number of correct guesses in the card guessing game starting with m red and n black cards is given as follows:*

$$\mathbb{P}\{X_{m,n} = k\} = \frac{\binom{m+n}{k} - \binom{m+n}{k+1}}{\binom{m+n}{m}}, \quad \text{for } m \leq k \leq m+n.$$

Since no limit laws were derived in [6, 12], we complete the analysis by stating the limit laws for $X_{m,n}$.

Corollary 1. *The random variable $X_{m,n}$ counting the number of correct guesses in the card guessing game starting with m red and n black cards satisfies, for $m \rightarrow \infty$ and depending on the growth of $n = n(m)$, the following limit laws.*

- *The region $n = o(m)$: the centered random variable $X_{m,n}^* = X_{m,n} - m$ is asymptotically zero:*

$$\mathbb{P}\{X_{m,n}^* = 0\} = 1 - \frac{n}{m+1} \rightarrow 1.$$

- *The region $n \sim \rho m$, with $0 < \rho < 1$: the centered random variable $X_{m,n}^* = X_{m,n} - m$ is asymptotically geometrically distributed with parameter ρ :*

$$\mathbb{P}\{X_{m,n}^* = k\} \rightarrow (1 - \rho)\rho^k, \quad k \in \mathbb{N}_0.$$

- *The region $n = m - \ell$, with $m = o(\ell^2)$: the centered and scaled random variable $X_{m,n}^* = \frac{\ell}{m}(X_{m,n} - m)$ is asymptotically exponential distributed:*

$$X_{m,n}^* \xrightarrow{\mathcal{L}} X, \quad \text{with density } f_X(x) = e^{-x}, \quad x \geq 0.$$

- *The region $n = m - \ell$, with $\ell \sim \rho\sqrt{m}$ and $\rho > 0$: the centered and scaled random variable $X_{m,n}^* = (X_{m,n} - m)/\sqrt{m}$ weakly converges to X_ρ :*

$$X_{m,n}^* \xrightarrow{\mathcal{L}} X_\rho, \quad \text{with density } f_{X_\rho}(x) = (\rho + 2x)e^{-x(\rho+x)}, \quad x \geq 0.$$

- *The region $n = m - \ell$, with $\ell = o(\sqrt{m})$: the centered and scaled random variable $X_{m,n}^* = (X_{m,n} - m)/\sqrt{m}$ is asymptotically Rayleigh distributed:*

$$X_{m,n}^* \xrightarrow{\mathcal{L}} X, \quad \text{with density } f_X(x) = 2te^{-t^2}, \quad x \geq 0.$$

Proof sketch. The limiting distributions can be derived by an application of Stirling's formula. The calculations are again quite lengthy, so we leave the details to the interested reader. \square

CONCLUSION

We have analysed the distribution of weighted lattice paths absorbed at certain lines using a generating function approach and the kernel method. Moreover, we have derived limiting distributions for the underlying random variable. Surprisingly, a wealth of different distributions arise in the limit. We address here the interesting open problem of extending the studies of the limiting distribution behaviour of the height of the absorption to other absorbing lines, and also to other classes of weighted lattice paths, or equivalently to other types of Pólya-Eggenberger urn models, such as the OK-Corral urn model, $M = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$.

REFERENCES

- [1] C. Banderier, M. Bousquet-Mélou, A. Denise, P. Flajolet, D. Gardy, D. Gouyou-Beauchamps. Generating Functions of Generating Trees. *Discrete Mathematics* 246, 29–55, 2002.
- [2] C. Banderier and P. Flajolet. Basic analytic combinatorics of directed lattice paths. *Theoretical Computer Science* 281, 37–80, 2002.
- [3] P. Flajolet, P. Dumas, and V. Puyhaubert. Some exactly solvable models of urn process theory. *Discrete Mathematics and Computer Science, Proceedings of Fourth Colloquium on Mathematics and Computer Science*, Ph. Chassaing Editor., vol. AG, 59–118, 2006.
- [4] P. Flajolet, J. Gabarró, H. Pekari. Analytic Urns, *Annals of Probability* 33, 1200–1233, 2005.
- [5] R. L. Graham, D. E. Knuth, and O. Patashnik. *Concrete Mathematics*. Addison-Wesley, Reading, 1994.

- [6] A. Knopfmacher and H. Prodinger, A simple card guessing game revisited, *Electronic journal of combinatorics* 8, paper R13, 9 pages, 2001.
- [7] H.-K. Hwang, M. Kuba and A. Panholzer, Analysis of some exactly solvable diminishing urn models. *Proceedings of the 19th International Conference on Formal Power Series and Algebraic Combinatorics*, Nankai University, Tianjin, 2007.
Online available at <http://www.fpsac.cn/PDF-Proceedings/Posters/43.pdf>
- [8] K. Levasseur, How to beat your kids at their own game, *Mathematical Magazine* 61, 301–305, 1988.
- [9] I. G. Macdonald. *Symmetric functions and Hall polynomials*, Second Ed., Oxford Univ. Press, Oxford, 1995.
- [10] S.G. Mohanty, *Lattice path counting and applications*, Academic Press, New York, 1979.
- [11] H. Prodinger, The kernel method: a collection of examples, *Séminaire Lotharingien de Combinatoire*, B50f, 19 pp., 2004.
- [12] R. A. Sulanke, Guessing, Ballot Numbers, and Refining Pascal’s Triangle. Manuscript.
Online available at <http://math.boisestate.edu/~sulanke/PAPERS/guessingweb.ps>
- [13] D. Zagier, How often should you beat your kids? *Mathematical Magazine* 63, 89–92, 1990.

MARKUS KUBA, INSTITUT FÜR DISKRETE MATHEMATIK UND GEOMETRIE, TECHNISCHE UNIVERSITÄT WIEN, WIEDNER HAUPTSTR. 8-10/104, 1040 WIEN, AUSTRIA
E-mail address: markus.kuba@tuwien.ac.at

ALOIS PANHOLZER, INSTITUT FÜR DISKRETE MATHEMATIK UND GEOMETRIE, TECHNISCHE UNIVERSITÄT WIEN, WIEDNER HAUPTSTR. 8-10/104, 1040 WIEN, AUSTRIA
E-mail address: Alois.Panholzer@tuwien.ac.at

HELMUT PRODINGER, MATHEMATICS DEPARTMENT, STELLENBOSCH UNIVERSITY, 7602 STELLENBOSCH, SOUTH AFRICA
E-mail address: hprodinger@sun.ac.za