LATTICE PATH COMBINATORICS — APPLICATIONS TO PROBABILITY AND STATISTICS

C. KRATTENTHALER AND S. G. MOHANTY

A *lattice path* P in $\mathbb{Z}^d$ is a path in the d-dimensional integer lattice $\mathbb{Z}^d$ which uses only points of the lattice, that is, it is a sequence $(P_0, P_1, \dots, P_l)$, where $P_i \in \mathbb{Z}^d$ for all i. The vectors $\overrightarrow{P_0P_1}, \overrightarrow{P_1P_2}, \dots, \overrightarrow{P_{l-1}P_l}$ are called the *steps* of P. The number of steps, l, is called the *length* of P, and denoted by $\ell(P)$. Figure 1 shows a lattice path in $\mathbb{Z}^2$ of length 11. (The horizontal line labelled $y = R - 2$ should be ignored at the moment.)

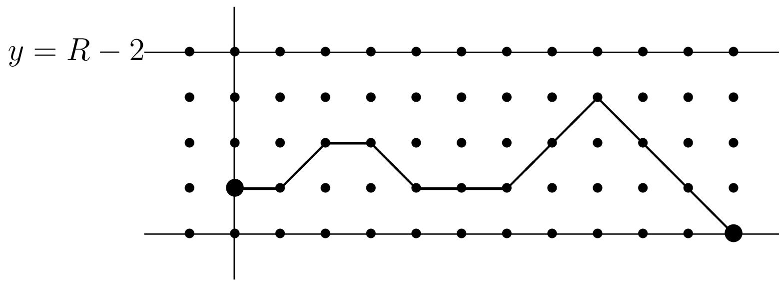

FIGURE 1. A lattice path

The interest in the combinatorial study of lattice paths in a statistical context arises primarily from three different sources: *random walks*, *rank order statistics* for *non-parametric testing*, and *queueing processes*. In most cases one is concerned with 2-dimensional lattice paths, i.e., with the case $d = 2$.

The prototypical example of a random walk problem is the *gambler's ruin problem* (see also [7, Ch. XIV]): two players A and B have initially a and $R - a$ dollar, respectively. They play several rounds, in each of which the probability that player A wins is p_A, the probability that player B wins is p_B, and that there is a tie is $p_T = 1 - p_A - p_B$. If one player wins, (s)he takes a dollar from the other. If there is a tie, nothing happens. The play stops when one of the players is bankrupt. What is the probability that player A, say, goes bankrupt after N rounds ?

By disregarding the last round (which is necessarily a round in which B wins), this problem can be represented by a lattice path starting at $(0, a-1)$, ending at $(N-1,0)$, with steps $(1,1)$ (corresponding to player A to win a round), $(1,-1)$ (corresponding to player B to win a round), and $(1,0)$ (corresponding to a tie), which does not pass below the x-axis, and which does not pass above the horizontal line $y=R-2$. For example, the lattice path in Figure 1 corresponds to the play, where player A starts with 2 dollar, player B starts with 4 dollar, the outcome of the rounds is in turn TATBTTAABBBB (the letter A symbolizing a round where A won, with an analogous meaning of the letter B, and the letter T symbolizing a tie), so that A goes bankrupt after $N=12$ rounds (while B did not).

If we assign the *weight* p_A to an up-step $(1,1)$, p_B to a down-step $(1,-1)$, and p_T to a level-step $(1,0)$, then the probability of this play is the product of the weights of all the steps of the path P times p_B (corresponding to the last round where B wins and A goes bankrupt; in our example, it is $p_T p_A p_T p_B p_T p_T p_A p_A p_B p_B p_B p_B$). If we write $p(P)$ for the product of the weights of the steps of P, then, in order to solve the problem, we need to compute the sum $\sum_P p_B p(P)$, where the sum is over all the above described paths from $(0,a-1)$ to $(N-1,0)$.

This problem is solved, in terms of generating functions, by the general theorem below, which shows that the solution is provided by *orthogonal polynomials*. For the statement of the theorem, we slightly modify and generalize the lattice path problem. We consider three-step lattice paths as above[1] (i.e., consisting of up-steps $(1,1)$, down-steps $(1,-1)$, and level-steps $(1,0)$), starting at $(0,r)$ and ending at (ℓ,s), which do not pass below the x-axis and do not pass above the horizontal line $y=K$. Furthermore, we assign the weight 1 to an up-step, the weight b_h to a level-step at height h, and the weight λ_h to a down-step from height h to $h-1$. The weight $w(P)$ of a path P is defined as the product of the weights of all its steps.[2] The gambler's ruin problem that we considered above corresponds to the choices $K=R-2$, $r=a-1$, $s=0$, $b_h=p_T$, and $\lambda_h=p_A p_B$ in the more general problem. (This is because, for any three-step path from $(0,a-1)$ to $(\ell,0)$, the difference of the numbers of down- and up-steps is $a-1$, and, thus, the described weighting differs from the weight which results by weighting an up-step by p_A, a down-step by p_B, and a level-step by p_t always by a factor of p_A^{a-1}, regardless of ℓ.)

---

[1]In the combinatorial literature, the term *Motzkin path* is often used for the special three-step lattice paths that start at the origin, return to the x-axis, and do not pass below the x-axis. Furthermore, the term *Dyck path* is used for Motzkin paths without any level-step.

[2]Clearly, these three-step paths can also be interpreted as a discrete-time birth-death process with stay, where the number of individuals in the process does not exceed K (see [4]). There is as well a Markov chain interpretation, see [7, Ch. XVI].

Theorem 1.[3] *Define the sequence $(p_n(x))_{n\ge 0}$ of polynomials by*

$$xp_n(x) = p_{n+1}(x) + b_np_n(x) + \lambda_np_{n-1}(x), \qquad \text{for } n \ge 1, \tag{1}$$

with initial conditions $p_0(x) = 1$ and $p_1(x) = x - b_0$. Furthermore, define $(Sp_n(x))_{n\ge 0}$ to be the sequence of polynomials which arises from the sequence $(p_n(x))$ by replacing λ_i by λ_{i+1} and b_i by b_{i+1}, $i = 0, 1, 2, \ldots$, everywhere in the three-term recurrence (1) and in the initial conditions. Finally, given a polynomial $p(x)$ of degree n, we denote the corresponding reciprocal polynomial $x^np(1/x)$ by $p^*(x)$.

With the weight w defined as before, the generating function $\sum_P w(P)x^{\ell(P)}$, where the sum is over all three-step paths which start at $(0, r)$, terminate at height s, do not pass below the x-axis, and do not pass above the line $y = K$, is given by

$$\begin{cases} \dfrac{x^{s-r}p_r^*(x)S^{s+1}p_{K-s}^*(x)}{p_{K+1}^*(x)} & r \le s, \\ \lambda_r \cdots \lambda_{s+1}\dfrac{x^{r-s}p_s^*(x)S^{r+1}p_{K-r}^*(x)}{p_{K+1}^*(x)} & r \ge s. \end{cases} \tag{2}$$

We remark that in the case that $r = s = 0$ there is also an elegant expression for the generating function due to Flajolet [9] in terms of a continued fraction.

In order to solve our problem, we just have to extract the coefficient of x^ℓ in (2). By a partial fraction expansion, a formula of the type

$$\sum_m c_m\xi_m^\ell, \tag{3}$$

results, where the ξ_m's are the zeroes of $p_{K+1}(x)$, and the c_m's are some coefficients, only a finite number of them being non-zero. In particular, the asymptotic behaviour of (3) is typically governed by the ξ_m's of largest absolute value, i.e., it is of the form

$$\sum_{m:|\xi_m|=M} c_m\xi_m^\ell, \tag{4}$$

where M denotes $\max_m |\xi_m|$. A difficulty would occur when this expression vanishes. However, under "normal" circumstances, this event rarely occurs. For, by Favard's theorem (see [43, Théorème 9 on p. I-4] or [44, Theorem 50.1]), the sequence of polynomials $(p_n(x))_{n\ge 0}$ is in fact a sequence of *orthogonal polynomials*. It is then known by general facts about orthogonal

[3] For self-contained derivations see e.g. [23, Appendix] or [28, Eqs. (4)/(5)].

polynomials (see e.g. [40, Theorem 3.3.1]) that if all b_n's are real and $\lambda_n > 0$ for all n, all the zeroes of $p_n(x)$ are in fact real and simple.

It should be noted that, because of the many available parameters (the b_n's and λ_n's), by appropriate specializations one can also obtain numerous results about enumerating three-step paths according to various statistics, such as the number of touchings on the bounding lines, etc. (See also [20]; see [4, Theorem 3.1] for alternative combinatorial formulas.)

There are two important special cases, in which a completely explicit solution in terms of elementary functions can be given.

The first case occurs for $b_i = 0$ and $\lambda_i = 1$ for all i, which (up to some multiplicative constant) corresponds to the game where each of the players A and B win a round with probability 1/2. In this case, the polynomials $p_n(x)$ defined by the three-term recurrence (1) are Chebyshev polynomials of the second kind,[4] $p_n(x) = U_n(x/2)$. The result which is then obtained from Theorem 1 (clearly, the zeroes of $U_n(x)$ are $x = \cos(2k\pi/(n+1))$, $k = 1, 2, \ldots, n$, and therefore the partial fraction expansion of (2) is easily determined) is that the number of lattice paths from $(0, r)$ to (ℓ, s) with only up- and down-steps, which always stay between the x-axis and the line $y = K$, is given by (see also [7, Ch. XIV, Eq. (5.7)])

$$\frac{2}{K+2} \sum_{k=1}^{K+1} \left(2 \cos \frac{\pi k}{K+2}\right)^{\ell} \cdot \sin \frac{\pi k(r+1)}{K+2} \cdot \sin \frac{\pi k(s+1)}{K+2}, \tag{5}$$

a formula which goes back to Lagrange. (An alternative expression is given in (9) with the replacements $s \to -r$, $t \to K - r$, $n \to \frac{1}{2}(\ell - s + r)$, $m \to \frac{1}{2}(\ell + s - r)$. It shows more clearly the integrality of the number. On the other hand, the advantage of the expression (5) is that it immediately allows to extract the asymptotic behaviour as the length ℓ of the paths tends to ∞.)

The second case occurs for $b_i = 1$ and $\lambda_i = 1$ for all i, which (up to some multiplicative constant) corresponds to the game where each of the three possibilities (player A wins, player B wins, a tie occurs) has equal probability 1/3. In this case, the polynomials $p_n(x)$ defined by the three-term recurrence (1) are again Chebyshev polynomials of the second kind, $p_n(x) = U_n((x-1)/2)$. The result which is then obtained from Theorem 1 is that the number of three-step lattice paths from $(0, r)$ to (ℓ, s), which always

---

[4] The Chebyshev polynomial of the second kind $U_n(x)$ is defined by $U_n(\cos t) = \sin((n+1)t)/\sin t$ (see [21] for almost exhaustive information on these polynomials and, more generally, on hypergeometric orthogonal polynomials).

stay between the x-axis and the line $y = K$, is given by

$$\frac{2}{K+2} \sum_{k=1}^{K+1} \left(2 \cos \frac{\pi k}{K+2} + 1 \right)^{\ell} \cdot \sin \frac{\pi k (r+1)}{K+2} \cdot \sin \frac{\pi k (s+1)}{K+2}. \tag{6}$$

We now turn to the second statistical motivation to study lattice paths: *rank order statistics* for *non-parametric testing.* There, we are given two sets of independent and identically distributed random variables $\mathcal{X} = \{X_1, X_2, \dots, X_m\}$ of size m and $\mathcal{Y} = \{Y_1, Y_2, \dots, Y_n\}$ of size n. These are then put together and ordered into $\mathcal{Z} = (Z_1, Z_2, \dots, Z_{m+n})$ according to size. To such an ordered sample one associates a lattice path in $\mathbb{Z}^2$ starting at the origin, in which the i-th step is a unit vertical step $(0,1)$ if Z_i belongs to $\mathcal{X}$, while it is a unit horizontal step $(1,0)$ if Z_i belongs to $\mathcal{Y}$. Thus, if $m = n = 5$, an example is $\mathcal{Z} = (X_1, Y_1, Y_2, Y_3, X_2, X_3, Y_4, X_4, X_5, Y_5)$, which would be represented as shown in Figure 2. (The diagonal lines should be ignored at the moment.) From now on, unless otherwise stated, the term "path" will always mean a lattice path in $\mathbb{Z}^2$ with unit horizontal and vertical steps in the positive direction.

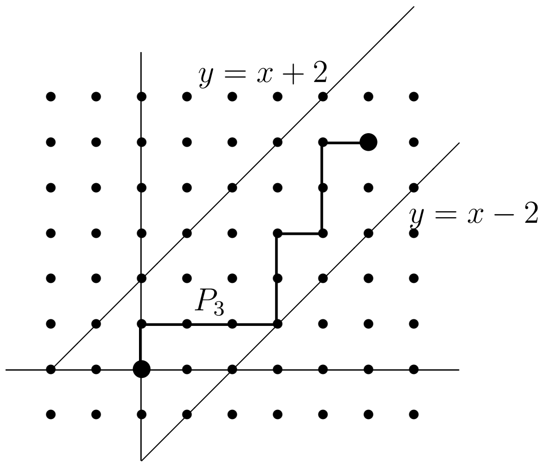

FIGURE 2

To test whether the underlying distribution functions for $\mathcal{X}$ and $\mathcal{Y}$ are equal or not, one introduces several statistics, all of which translate into lattice path statistics. For the purpose of testing, the distributions of these statistics are needed under the *null hypothesis* of equal distribution functions. In light of the preceding remark, this involves counting of certain paths.

The *one-sided Kolmogorov-Smirnov statistic* $D^+_{m,n}$ is defined by

$$D^+_{m,n} = \max_i \left\{ \frac{a_i}{m} - \frac{b_i}{n} \right\},$$

where a_i is the number of occurrences of X_j's in the initial segment $Z_1, Z_2, \dots,$ Z_i of $\mathcal{Z}$, while b_i is the number of occurrences of Y_j's in this initial segment. The *two-sided Kolmogorov-Smirnov statistic* $D_{m,n}$ is defined by

$$D_{m,n} = \max_i \left\{ \left| \frac{a_i}{m} - \frac{b_i}{n} \right| \right\}.$$

Thus we have $D^+_{5,5} = 1/5$ and $D_{5,5} = 2/5$ in our sample represented in Figure 2. The *run statistic* counts the number of maximal consecutive subsequences in $\mathcal{Z}$ the members of which belong to just one of the sets $\mathcal{X}$ or $\mathcal{Y}$. The number of runs in our example sample is 6. Other statistics that are considered are the *Galton statistic* (in the lattice path picture: the number of steps on the "positive" side of the main diagonal $y = x$), the *median statistic*, and the *rank sum statistic*, to mention a few, see [27, Ch. 4].

In the lattice path picture, the one-sided Kolmogorov-Smirnov statistic is basically the maximal deviation from the main diagonal in direction $(-1, 1)$. The two-sided Kolmogorov-Smirnov statistic is basically the maximal deviation from the main diagonal, in either direction. So in Figure 2, paths which stay in the region between the indicated lines $y = x + 2$ and $y = x - 2$ correspond to sequences $\mathcal{Z}$ with two-sided Kolmogorov-Smirnov statistic $D_{n,n} \le 2/5$. The run statistic obviously translates into the number of maximal straight pieces (horizontal or vertical) in the corresponding path.

Clearly, the number of all (unrestricted) paths from the origin to (n, m) is the binomial coefficient $\binom{n+m}{n}$. By the *reflection principle*, which is commonly attributed to D. André (see e.g. [5, p. 22]), it follows that the number of paths from the origin to (n, m) which do not pass above the line $y = x + t$, where $m \le n + t$, is given by[5]

$$\binom{n+m}{n} - \binom{n+m}{n+t+1}. \tag{7}$$

In the case that $n = m$, this translates easily into an expression for the probability of encountering $D^+_{n,n} \le t/n$ under the null hypothesis. If $n \neq m$, we

[5]Roughly, the reflection principle sets up a bijection between the paths from the origin to (n, m) which *do pass above* the line $y = x + t$ and *all* paths from $(-t - 1, t + 1)$ to (n, m), by reflecting the path portion between the origin and the last touching point on $y = x + t + 1$ in this latter line. Thus, the result of the enumeration problem is the number of *all* paths from $(0, 0)$ to (n, m), which is given by the binomial coefficient $\binom{n+m}{n}$, minus the number of *all* paths from $(-t - 1, t + 1)$ to (n, m), which is given by the binomial coefficient $\binom{n+m}{n+t+1}$, whence the formula (7).

need to count paths that do not pass above the line $my = nx + t$, for which there is no compact formula known (and is unlikely to exist). Sato [32, 33] has dealt with this problem (and its higher dimensional generalization) extensively. However, the most conceptual way to approach this problem seems to be through the so-called *kernel method* (see [1]), which, in combination with the saddle point method, allows one also to obtain strong asymptotic results. There is one special instance, however, which has a "nice" formula. The number of all lattice paths from the origin to (n, m) which never pass above $x = \mu y$, where μ is a positive integer, is given by

$$\frac{n - \mu m + 1}{n + m + 1}\binom{n + m + 1}{m}. \tag{8}$$

The most elegant way to prove this formula is by means of the *cycle lemma* of Dvoretzky and Motzkin [6] (see [27, p. 9], where the cycle lemma occurs under the name of "penetrating analysis").

Expression (7) with $t = 1$ and (8) provide solutions to the so-called *classical ballot problem*. Its generalization to the urn problem due to Takács (see [41, p. 2–4]), which has applications in queueing theory, has its solution through the form of a cycle lemma. Its most general form is due to Spitzer [35], and is known as "Spitzer's lemma," which has also many applications in random walk theory.

Iteration of the reflection principle shows that the number of paths from the origin to (n, m) which stay between the lines $y = x + t$ and $y = x + s$ (being allowed to touch them), where $t \geq 0 \geq s$ and $n + t \geq m \geq n + s$, is given by the finite (!) sum

$$\sum_{k\in\mathbb{Z}}\left(\binom{n + m}{n - k(t - s + 2)} - \binom{n + m}{n - k(t - s + 2) + t + 1}\right). \tag{9}$$

(See e.g. [27, p. 6]. Clearly, an alternative expression is provided by (5), under the substitutions $s \to m - n - s$, $r \to -s$, $K \to t - s$, $\ell \to n + m$. While the above expression shows clearly the integrality of the numbers, the expression (5) allows to extract the asymptotic behaviour as the length of the paths tends to ∞.) If $n = m$ and $t = s$, this translates in an obvious assertion about the probability of encountering $D_{n,n} \leq t/n$ under the assumption that the distribution functions of $\mathcal{X}$ and $\mathcal{Y}$ are the same. Again, if $n \neq m$, there is no compact formula. Sato [32, 33] has given formulas in terms of generating functions, but it seems difficult to work with them, except in a few special cases.

The enumeration of lattice paths restricted to regions bounded by hyperplanes has also been considered for other regions, such as quadrants, octants, and rectangles, as well as in higher dimensions. The papers [3, 11] (see also [12]) contain a general result on the number of lattice paths in a chamber (alcove) of an (affine) reflection group that shows how far one can go when one uses the reflection principle. In particular, this result covers (7) and (9), the enumeration of lattice paths in quadrants, octants, rectangles, and many other results that have appeared (before and after) in the literature. We present a particularly elegant (and frequently occurring) special case. (In the language of [3, 11], it corresponds to the reflection group of "type A_{n-1}". See [15] for terminology and information on reflection groups.)

Theorem 2. *Let $A = (a_1, a_2, \dots, a_d)$ and $E = (e_1, e_2, \dots, e_d)$ be points in $\mathbb{Z}^d$ with $a_1 \ge a_2 \ge \cdots \ge a_d$ and $e_1 \ge e_2 \ge \cdots \ge e_d$. The number of all paths from A to E in the integer lattice $\mathbb{Z}^d$, which consist of positive unit steps and which stay in the region $x_1 \ge x_2 \ge \cdots \ge x_d$, equals*

$$\left(\sum_{i=1}^{d}(e_i - a_i)\right)! \det_{1\le i,j\le d}\left(\frac{1}{(e_i - a_j - i + j)!}\right). \tag{10}$$

The counting problem of the theorem is equivalent to numerous other counting problems. It has been originally formulated as an n-candidate ballot problem (see e.g. [2]), but it is as well equivalent to counting the number of standard Young tableaux of a given shape (see e.g. [2, 45]). In the case that all a_j's are equal, the determinant does in fact evaluate into a closed form product. In Young tableaux theory a particular way to write the result is known as the *hook-length formula* (see e.g. [31, Sec. 3.10] or [36, Cor. 7.21.6]).

We return to lattice paths in the plane, mentioning some more closely related results. The first is a result of Mohanty [26], which expresses the number of all lattice paths from the origin to (n, m) which touch the line $y = x + t$ *exactly* r times, never crossing it, as the difference

$$\binom{n+m-r}{n+t-1} - \binom{n+m-r}{n+t}, \qquad r \ge 1. \tag{11}$$

Not forbidding that the paths cross the bounding line, we arrive at the problem of counting the lattice paths from the origin to (n, m), which cross the main diagonal $y = x$ exactly r times, the answer being [17]

$$\begin{cases} \frac{m-n+2r+1}{m+n+1}\binom{m+n+1}{n-r} & \text{if } m > n, \\ \frac{2r+2}{n}\binom{2n}{n-r-1} & \text{if } m = n. \end{cases} \tag{12}$$

Next, we give the number of lattice paths from the origin to (n,n) which have $2r$ steps on one side of the line $y=x$, as

$$\binom{2r}{r}\binom{2n-2r}{n-r}, \tag{13}$$

a result due to Sparre Andersen [34]. We refer the reader to [27, Ch. 3] for further results in this direction.

Enumerating lattice paths with a fixed number of maximal straight pieces (which correspond to runs), is intimately connected to another basic enumeration problem concerning lattice paths: the enumeration of lattice paths having a fixed number of *turns*. An effective way to attack the latter problem is by means of *two-rowed arrays*, see the survey article [22], where in particular analogues of the reflection principle for two-rowed arrays are developed. These imply formulas for the number of lattice paths with fixed starting and end points and a fixed number of *north-east* (respectively *east-north*) *turns*[6], for unrestricted paths, as well as for paths bounded by lines. In particular, analogues of (7)–(9) are known when the number of north-east (respectively east-north) turns is fixed.[7]

These formulas imply for example (see again [22, Sec. 3.5]) that the number of lattice paths from the origin to (n,n) which never pass above the line $y=x+t$ and have exactly $2r$ maximal straight pieces is given by

$$2\binom{n-1}{r-1}^2-\binom{n+t-1}{r-2}\binom{n-t-1}{r}-\binom{n+t-1}{r-1}\binom{n-t-1}{r-1}, \tag{14}$$

with a similar result for the case of $2r+1$ maximal straight pieces.[8] Furthermore, they imply that the number of lattice paths from the origin to (n,n) which never pass above the line $y=x+t$ and never below the line $y=x-t$

[6]A north-east turn in a lattice path is a point where the direction changes from "north" to "east." An east-north turn is defined analogously.

[7]These formulas imply directly formulas for the probability of the *correlated random walk* (see [27, Sec. 5.2]), which is a random walk with horizontal and vertical steps in which the probability of a step is not independent of the previous step made, subject to various restrictions.

[8]If $t=0$, the numbers in (14) become $\frac{1}{n}\binom{n}{r}\binom{n}{r-1}$, and they are known as the *Narayana numbers*.

and have exactly $2r$ maximal straight pieces is given by

$$\sum_{k=-\infty}^{\infty} \left\{ 2\binom{n-2kt-1}{r+k-1}\binom{n+2kt-1}{r-k-1} \right. \\ -\binom{n-2kt+t-1}{r+k-2}\binom{n+2kt-t-1}{r-k} \\ \left. -\binom{n-2kt+t-1}{r+k-1}\binom{n+2kt-t-1}{r-k-1} \right\}, \tag{15}$$

with a similar result for the case of $2r+1$ maximal straight pieces. Both, Eqs. (14) and (15), are results originally obtained by Vellore [42]. They translate into obvious assertions about the joint probability of encountering a fixed number of runs *and* $D^+_{n,n} \le t/n$, respectively $D_{n,n} \le t/n$, under the assumption that the distribution functions of $\mathcal{X}$ and $\mathcal{Y}$ are the same.

The most general boundary for lattice paths that one can imagine is the restriction that it stays between two given (fixed) paths. Let us assume that the horizontal steps of the upper (fixed) path are at heights $a_1 \le a_2 \le \dots \le a_n$, whereas the horizontal steps of the lower (fixed) path are at heights $b_1 \le b_2 \le \dots \le b_n$, $a_i \ge b_i$, $i = 1, 2, \dots, n$. Then the number of all paths from $(0, b_1)$ to (n, a_n) satisfying the property that for all $i = 1, 2, \dots, n$ the height of the i-th horizontal step is between b_i and a_i is given by the determinant

$$\det_{1\le i,j\le n}\left(\binom{a_i - b_j + 1}{j - i + 1}\right). \tag{16}$$

In the statistical literature, this formula is often known as "Steck's formula" [37], but it is actually a special case of a much more general theorem due to Kreweras [24]. A generalization of (16) to higher dimensional paths was given by Handa and Mohanty [14]. For a continuous analogue of (16), and its applications see [27, Sec. 4.5].

Several of the formulas presented have also applications in queueing theory (see [41] and [27, Sec. 5.3] for more extensive treatments of this aspect of lattice paths). Consider an $M/M/1$ queueing system in which the customers arrive individually at a counter in accordance with a Poisson process of density λ_1 and are served individually by a single server. The service times are i.i.d. random variables with distribution function

$$F(x) = \begin{cases} 1 - e^{-\lambda_0 x} & \text{if } x \ge 0, \\ 0 & \text{otherwise,} \end{cases}$$

and are independent of the arrival times. In the combined process, every event independent of others is either an arrival with probability $p = \lambda_1/(\lambda_0 + \lambda_1)$ or a departure with probability $q = 1-p$. If we represent a departure by a vertical unit and an arrival by a horizontal unit, then a busy period, initiated by j customers and consisting of $n+j$ services (departures) corresponds to the set of lattice paths from $(0,0)$ to $(n, n+j)$ that do not touch the line $y = x + j$ except at the end. Thus, the probability for a busy period initiated by j customers in which $n+j$ customers are served, is $p^n q^{n+j}$ times the number of these lattice paths. The latter number results from (7) by setting $m = n + j - 1$ and $t = j - 1$. Using some probabilistic reasoning, it is then possible to derive an explicit integral formula for the probability that such a busy period has length $\le t$. More refined queueing problems (allowing for arrivals/departures in batches, for bounds on the number of waiting customers, etc.) can also be treated by this lattice path method. We refer the reader again to [41] and [27, Sec. 5.3].

We remark that another area of application of lattice path enumeration has been in Discrete Distributions, see [27, Ch. 5].

Another vast topic is the enumeration of *several* paths which do not have any common points.[9] This problem has been introduced and was first studied by Karlin and McGregor [18, 19] (in the continuous case). If the starting and end points of the paths are fixed, the number of such families of non-touching paths is given (under mild restrictions on the starting and end points) by a determinant, where the (i, j)-entry of the determinant counts the number of paths from the j-th starting to the i-th end point. The most general form of this theorem has been obtained by Lindström [25] and has been rediscovered by Gessel and Viennot [10].[10] It covers numerous determinant formulas in the literature, including (16) and its higher dimensional generalization due to Handa and Mohanty (as shown by Sulanke [39]). The most widely used case of the theorem is the following.

Theorem 3. *Let G be an acyclic directed graph. Let $(A_1, A_2, \dots, A_n)$ and $(E_1, E_2, \dots, E_n)$ be sequences of vertices in G such that for $i < j$ and $k < l$ any path from A_i to E_l and any path from A_j to E_k have at least one point in common. Then the number of all families $(P_1, P_2, \dots, P_n)$ of non-touching*

[9] In the combinatorial literature, such families of paths are, slightly confusingly, called "*non-intersecting*" paths. In this text we use the term "*non-touching*" paths.

[10] Other occurrences of non-touching paths are in statistical physics under the name of *"vicious walkers"* as introduced by Fisher [8], and in combinatorial chemistry in the enumeration of perfect matchings of hexagonal graphs, see [13, 16].

paths, where P_i is a path running from A_i to E_i, $i = 1, 2, \ldots, n$, is given by

$$\det_{1 \le i,j \le n} \big(P(A_j \to E_i) \big), \tag{17}$$

where $P(A_j \to E_i)$ is the number of paths from A_j to E_i.

The most general theorem includes weighted counting as well, and it drops the condition involving i, j, k, l. The standard application of the above "weak" form of the theorem is to the enumeration of families of non-touching paths in the integer lattice.

Theorem 4. *Let $(A_1, A_2, \ldots, A_n)$ and $(E_1, E_2, \ldots, E_n)$ be sequences of lattice points in $\mathbb{Z}^2$ such that for $i < j$ and $k < l$ any path from A_i to E_l and any path from A_j to E_k have at least one point in common. Then the number of all families $(P_1, P_2, \ldots, P_n)$ of non-touching paths, where P_i consists of horizontal and vertical steps in the positive direction and runs from A_i to E_i, $i = 1, 2, \ldots, n$, is given by*

$$\det_{1 \le i,j \le n} \big(P(A_j \to E_i) \big), \tag{18}$$

where $P(A \to E)$ is the binomial coefficient $\binom{e_1+e_2-a_1-a_2}{e_1-a_1}$, given that $A = (a_1, a_2)$ and $E = (e_1, e_2)$.

If the starting points or/and the end points are not fixed, then the corresponding number is given by a Pfaffian, a result obtained by Okada [30] and Stembridge [38]. Refinements when the number of turns is fixed have been obtained by Krattenthaler, see [22].

For further reading, and many more references, we refer the reader to the textbooks [27] and [29].

References

[1] C. Banderier and P. Flajolet, *Basic analytic combinatorics of directed lattice paths*, Theoret. Computer Science **281** (2002), 37–80.

[2] D. E. Barton and C. L. Mallows, *Some aspects of the random sequences*, Amer. Math. Statist. **36** (1965), 236–260.

[3] P. Biane, *Minuscule weights and random walks on lattices*, in: Quantum probability & related topics, L. Accardi, ed., World Sci. Publ., River Edge, NJ, 1992, pp. 51–65.

[4] W. Böhm, A. Krinik and S. G. Mohanty, *The combinatorics of birth-death processes and applications to queues*, Queueing Systems **26** (1997), 255–267.

[5] L. Comtet, *Advanced Combinatorics*, D. Reidel, Dordrecht, Holland, 1974.

[6] A. Dvoretzky and T. Motzkin, *A problem of arrangements*, Duke Math. J. **14** (1947), 305–313.

[7] W. Feller, *An introduction to probability theory and its applications*, vol. 1, 2nd ed., Wiley, New York, 1957.

[8] M. E. Fisher, *Walks, walls, wetting, and melting*, J. Statist. Phys. **34** (1984), 667–729.

[9] P. Flajolet, *Combinatorial aspects of continued fractions*, Discrete Math. **32** (1980), 125–161.

[10] I. M. Gessel and X. Viennot, *Binomial determinants, paths, and hook length formulae*, Adv. in Math. **58** (1985), 300–321.

[11] I. M. Gessel and D. Zeilberger, *Random walk in a Weyl chamber*, Proc. Amer. Math. Soc. **115** (1992), 27–31.

[12] D. J. Grabiner, *Random walk in an alcove of an affine Weyl group, and non-colliding random walks on an interval*, J. Combin. Theory Ser. A **97** (2002), 285–306.

[13] H.-D. O. F. Gronau, W. Just, W. Schade, P. Scheffler and J. Wojciechowski, *Path systems in acyclic directed graphs*, Zastos. Mat. **19** (1988), 399–411.

[14] B. R. Handa and S. G. Mohanty, *Enumeration of higher-dimensional paths under restrictions*, Discrete Math. **26** (1979), 119–128.

[15] J. E. Humphreys, *Reflection groups and Coxeter groups*, Cambridge University Press, Cambridge, 1990.

[16] P. John and H. Sachs, *Wegesysteme und Linearfaktoren in hexagonalen und quadratischen Systemen*, Graphs in research and teaching (Kiel, 1985), Franzbecker, Bad Salzdetfurth, 1985, pp. 85–101.

[17] Kanwar Sen, *On some combinatorial relations concerning the symmetric random walk*, Publ. Math. Inst. Hungar. Acad. Sci. **9** (1965), 335–357.

[18] S. Karlin and J. L. McGregor, *Coincidence properties of birth-and-death processes*, Pacific J. Math. **9** (1959), 1109–1140.

[19] S. Karlin and J. L. McGregor, *Coincidence probabilities*, Pacific J. Math. **9** (1959), 1141–1164.

[20] W. Katzenbeisser and W. Panny, *Simple random walk statistics I: discrete time results*, J. Appl. Probab. **33** (1996), 311–330.

[21] R. Koekoek and R. F. Swarttouw, *The Askey–scheme of hypergeometric orthogonal polynomials and its q-analogue*, TU Delft, The Netherlands, 1998; on the WWW: `ftp://ftp.twi.tudelft.nl/TWI/publications/tech-reports/1998/DUT-TWI-98-17.ps.gz`.

[22] C. Krattenthaler, *The enumeration of lattice paths with respect to their number of turns*, in: Advances in Combinatorial Methods and Applications to Probability and Statistics, N. Balakrishnan, ed., Birkhäuser, Boston, 29–58.

[23] C. Krattenthaler, *Permutations with restricted patterns and Dyck paths*, Adv. Appl. Math. **27** (2001), 510–530.

[24] G. Kreweras, *Sur une classe de problèmes de dénombrement liés au treillis des partitions des entiers*, Cahiers du B.U.R.O. Univ. de Rech. Opér. **6** (1965), 5–105.

[25] B. Lindström, *On the vector representations of induced matroids*, Bull. London Math. Soc. **5** (1973), 85–90.

[26] S. G. Mohanty, *On some generalization of a restricted random walk*, Stud. Sci. Math. Hungarica **3** (1968), 225–241.

[27] S. G. Mohanty, *Lattice Path Counting and Applications*, Academic Press, New York, 1979.

[28] S. G. Mohanty, *On the transient behaviour of a finite discrete time birth-death process*, Assam Statist. Rev. **5** (1991), 1–7.

[29] T. V. Narayana, *Lattice path combinatorics with statistical applications*, Math. Expositions, no. 23, Univ. of Toronto Press, Toronto, 1979.

[30] S. Okada, *On the generating functions for certain classes of plane partitions*, J. Combin. Theory A **51** (1989), 1–23.

[31] B. E. Sagan, *The symmetric group*, 2nd edition, Springer–Verlag, New York, 2001.

[32] M. Sato, *Generating functions for the number of lattice paths between two parallel lines with a rational incline*, Math. Japonica **34** (1989), 123–137.

[33] M. Sato, *Generating functions for the number of lattice paths restricted by two parallel hyperplanes*, J. Statist. Plann. Inference **34** (1993), 251–258.

[34] E. Sparre Andersen, *On the number of positive sums of random variables*, Scand. Actuarietidskr. **32** (1949), 27–36.

[35] F. Spitzer, *A combinatorial lemma and its application to probability theory*, Trans. Amer. Math. Soc. **82** (1956), 323–339.

[36] R. P. Stanley, *Enumerative Combinatorics*, Vol. 2, Cambridge University Press, Cambridge, 1999.

[37] G. P. Steck, *The Smirnov two-sample tests as rank tests*, Ann. Math. Statist. **40** (1969), 1449–1466.

[38] J. R. Stembridge, *Nonintersecting paths, pfaffians and plane partitions*, Adv. in Math. **83** (1990), 96–131.

[39] R. A. Sulanke, *A determinant for q-counting lattice paths*, Discrete Math. **81** (1990), 91–96.

[40] G. Szegő, *Orthogonal polynomials*, American Mathematical Society Colloquium Publications Vol. XXIII, New York, 1959.

[41] L. Takács, *Combinatorial methods in the theory of stochastic processes*, Wiley, New York, 1967.

[42] S. Vellore, *Joint distributions of Kolmogorov–Smirnov statistics and runs*, Studia Scientiarum Math. Hungarica **7** (1972), 155–165.

[43] X. Viennot, *Une théorie combinatoire des polynômes orthogonaux généraux*, UQAM, Montréal, Québec, 1983.

[44] H. S. Wall, *Analytic Theory of Continued Fractions*, Van Nostrand, New York, 1948.

[45] D. Zeilberger, *André's reflection proof generalized to the many candidate ballot problem*, Discrete Math. **44** (1983), 325–326.