Probabilistic approach to lattice path enumeration.

Vitali Wachtel

University of Munich

joint work with Denis Denisov (University of Cardiff)

Classical ballot problem

Candidates A and B received m_1 and m_2 votes respectively, where $m_1 < m_2$. The ballot problem asks for the number of ways to count these votes in such a way that at all times the partial scores satisfy $x_1 < x_2$. In other words, one counts all lattice walks from $(0, 0)$ to (m_1, m_2) with positive unit steps such that $x_1 < x_2$ holds for all nodes.

Probabilistic interpretation: Let $\{S(n)\}$ be the simple symmetric random walk and let τ denote the first descending ladder epoch, i.e. $\tau = \min\{k \geq 1 : S(k) \leq 0\}$. The ballot problem is then equivalent to the problem of computing $\mathbf{P}(S(m_1 + m_2) = m_2 - m_1, \tau > m_1 + m_2)$.

Highdimensional ballot problem

If we have k candidates with $m_1 < m_2 < \ldots \leq m_k$ votes, then the problem is calculate the number of walks in $\mathbb{Z}^k$ satisfying the condition $0 \leq x_1 < x_2 < \ldots < x_k$.

Probabilistic interpretation: Let $\{S(n)\}$ be a $(k-1)$-dimensional random walk such that

$$
\begin{aligned}
&\mathbf{P}(X = (-1, 0, \ldots, 0)) = k^{-1}, \\
&\mathbf{P}(X = (1, -1, 0, \ldots, 0)) = k^{-1}, \\
&\quad \ldots \\
&\mathbf{P}(X = (0, \ldots, 0, 1, -1)) = k^{-1}, \\
&\mathbf{P}(X = (0, \ldots, 0, 1)) = k^{-1}.
\end{aligned}
$$

Then, the problem is to compute
$\mathbf{P}(S(m_1 + m_2 + \ldots + m_k) = m, \tau > m_1 + m_2 + \ldots + m_k)$, where
$\tau = \min\{k \geq 1 : S(k) \notin \mathbb{Z}_+^{k-1}\}$.

The combinatorial approach consists in using a reflection principle. Andre (1887) suggested such a principle for the classical ballot problem. Gessel and Zeilberger (1992) gave a generalisation for a Weyl chamber generated by a root system.

In order to apply a reflection principle, it is necessary to assume that our random walk can not jump over the boundary.

Grabiner and Magyar (1993) found all random walks, to which the reflection principle of Gessel and Zeilberger can be applied.

Feierl (2011) suggested a generalisation of the reflection principle.

Walks in the quarter-plane:

Kreweras' walk: $\{W, S, NE\}$

Gessel's walk: $\{E, W, NE, SW\}$

Gouyou-Beauchamps's walk: $\{E, W, NW, SE\}$

Bousquet-Melou and Mishna (2010) have considered walks with small steps, i.e., $\mathcal{S} \subset \{-1, 0, 1\}^2 \setminus \{(0, 0)\}$. (79 different walks)

Kernel method. For the description see the book of Flajolet and Odlyzko. A probabilistic version of that method can be found in the book of Fayolle, Iasnogorodski and Malyshev. Fayolle and Raschel (2011) found exact expressions for generating functions of all 79 walks with small steps.

Compensation approach. This method allows one to find an explicit expression for the number of walks, and it has been developed by Adan, Wessels and Zijm (1990 and 1993). This mathod has recentely been applied to some walks confined to the positive quadrant, see Adan, Leeuwaarden and Raschel (2011).

Humphreys (2010): A history and a survey of lattice path enumeration

Combinatorics: Exact formulas for the number of walks or for the corresponding generating functions, quite hard restrictions on the step set or even very specific walks.

Probability: Asymptotic formulas for the number of walks, universality-idea, larger classes of walks.

Theorem 1.

There exists a positive harmonic function V such that :

$$\mathbf{P}(\tau_x > n) \sim \kappa V(x) n^{-p/2}, \quad x \in K,$$

and

$$\mathbf{P}\left(\frac{x+S(n)}{\sqrt{n}} \in \cdot \middle| \tau_x > n\right) \to \mu \quad \text{weakly,}$$

where μ is the probability measure on K with density proportional to $u(y)e^{-|y|^2/2}$.

Theorem 2. Assume that $S(n)$ takes values on some lattice and that $S(n)$ is aperiodic. Then

$$\sup_{y\in K}\left|n^{p/2+d/2}\mathbf{P}\left(x+S(n)=y,\tau_x>n\right)-\frac{V(x)}{(2\pi)^{d/2}}u\left(\frac{y}{\sqrt{n}}\right)e^{-|y|^2/2n}\right|\to 0.$$

Theorem 3. Under the assumptions of the preceding theorem, for every fixed $y \in K$,

$$\mathbf{P}\left(x + S(n) = y, \tau_x > n\right) \sim \varrho \frac{V(x)V'(y)}{n^{p+d/2}},$$

where V' is the harmonic function for the random walk $\{-S(n)\}$ and

$$\varrho = (2\pi)^{-d} \int_K u^2(w) e^{-|w|^2/2} dw.$$

Consider lattice paths with the step set $\mathcal{S} = \{s_1, s_2, \ldots, s_N\}$. If the vector sum of all s_i is not equal to zero, one has to perform the Cramer transformation with an appropriate parameter.

Set $R(h) = N^{-1} \sum_{i=1}^{N} e^{(h,s_i)}$ and define

$$\mathbf{P}(Y = s_i) = \frac{1}{R(h)} e^{(h,s_i)},$$

where h is such that $\mathbf{E}Y = 0$. (We assume that such a transformation is possible.)

Then we have the following formula for the number of walks with endpoints x and y

$$N_n(x,y) = N^n (R(h))^n e^{(h,x-y)} \mathbf{P}\left(x + \sum_{k=1}^{n} Y_k = y, \tau_x > n\right).$$

There exists a linear transformation with matrix M such that $X = MY$ has uncorrelated coordinates and

$$\mathbf{P}(x + \sum_{k=1}^{n} Y_k = y, \tau_x > n) = \mathbf{P}(Mx + S(n) = My, \tau_x > n).$$

Applying Theorem 3, we obtain

$$N_n(x, y) = V(x)V'(y)(NR(h))^n n^{-p-d/2}(1 + o(1)) \quad \text{as } n \to \infty,$$

where p, V and V' are characteristics of the cone MK.

Remarks.

- p, V and V' may depend on the parameter h. In other words, not only the cone K but also the step-set $\mathcal{S}$ affect these characteristics.
- The main disadvantage of this approach is the fact that we can not give an explicit expression for the functions V and V'. Therefore, we have determined the rate of growth of $N_n(x, y)$. Nevertheless, for large values of x and y inside the cone one can obtain an approximation for $V(x)V'(y)$ from the relation $V(x) \sim u(x)$.
- Upper bounds for $N_n(x, y)$ can be obtained from the estimates due to Varopoulos (1999). It follows from (0.7.4) in that paper that $V(x)V'(y)$ can be bounded from above by $Cu(x + y_0)u(y + y_0)$ with some appropriate y_0. An essential advantage of this bound consists in the fact, that u is more accessible than V.

Remark. An asymptotic formula for the number of walks with a free end can be obtained directly from Theorem 1:

$$N_n(x) = V(x)\,(NR(h))^n\, n^{-p/2}(1 + o(1)).$$

Example. Consider a walk in the positive quarter with possible steps

$$(-1,0),(1,0),(1,-1),(-1,1),(0,-1),(0,1).$$

It is clear that $\mathbf{E}X^2 = \mathbf{E}Y^2 = 2/3$ and $\mathbf{E}XY = -1/3$. Therefore, $\varrho = -1/2$ and $p = \pi/\arccos(-\varrho) = 3$.

We next compute the harmonic function V. We first note that, applying the transformation $y_1 = \sqrt{2}x_1 + x_2/\sqrt{2}$, $y_2 = \sqrt{3/2}x_2$, we get a random walk with uncorrelated coordinates in the cone $\{re^{i\varphi},\ \varphi(0,\pi/3)\}$.

The corresponging harmonic function for BM is easy to find:

$$u(y) = r^3\sin(3\varphi) = \mathrm{Im}\left[(y_1+iy_2)^3\right] = 3y_1^2y_2 - y_2^3.$$

Therefore, for the BM with the correlation $-1/2$ we get

$$\begin{aligned}u(x) &= 3\left(\sqrt{2}x_1 + x_2/\sqrt{2}\right)^2\sqrt{3/2}x_2 - \left(\sqrt{3/2}x_2\right)^3\\ &= 3^{3/2}2^{-1/2}x_1x_2(x_1+x_2)\end{aligned}$$

Since our random walk can not jump over the boundary, we have

$$u(x+S(\tau)) \equiv 0.$$

Consequently,

$$V(x) = u(x) + \mathbf{E}\sum_{k=1}^{\tau-1} f(x+S(k)),$$

where

$$f(x) = \mathbf{E}u(x+X) - u(x).$$

But in this particular case we have $\mathbf{E}u(x+X) = u(x)$. As a result

$$V(x) = \left(\frac{27}{2}\right)^{1/2} x_1 y_1 (x_1 + y_1).$$

Applying our theorems we then have

$$\mathbf{P}(\tau_x > n) \sim C_1 V(x) n^{-3/2}$$

and

$$N_n(x,y) \sim C_2 V(x) V(y) 6^n n^{-4}.$$

(C_1 and C_2 are easy to compute.)

(The asymptotic for $\mathbf{P}(\tau_x > n)$ has been obtained in Kurkova and Raschel (2011).)

Some other examples with computable $V(x)$ can be found in Raschel (2011):

$$p_{1,0} = p_{-1,0} = \frac{\sin^2(\pi/p)}{2}, \quad p_{-1,1} = p_{1,-1} = \frac{\cos^2(\pi/p)}{2}.$$

If $p = 3$ or $p = 4$ then, as in the previous example, $V(x) = u(x)$. But if $p \geq 5$, then, as it has been shown in Raschel (2011), $V(x) \neq u(x)$. Here we have $f(x) \neq 0$.

Let $\hat{V}(s_1, s_2)$ be the generating function of the harmonic function $V(x_1, x_2)$ of a random walk with small steps:

$$\hat{V}(s_1, s_2) = \sum_{x_1, x_2 \geq 1} V(x_1, x_2) s_1^{x_1} s_2^{x_2}.$$

Using the definition of the harmonicity one can easy obtain

$$\tilde{K}(s_1, s_2)\hat{V}(s_1, s_2) = g_1(s_1)\hat{V}(s_1, 0) + g_2(s_2)\hat{V}(0, s_2) - p_{11}V(1, 1),$$

where

$$\tilde{K}(s_1, s_2) = \sum_{-1 \leq i, j \leq 1} p_{ij} s_1^{1-i} s_2^{1-j} - s_1 s_2,$$

$$g_1(s_1) = \sum_{-1 \leq i \leq 1} p_{i1} s_1^{1-i}, \quad g_2(s_2) = \sum_{-1 \leq j \leq 1} p_{1j} s_2^{1-j}.$$

Question: Can one use the kernel method to solve the following problems:

- Exact (closed) expressions for harmonic functions;
- Number of harmonic functions.