ESAIM: PS 16 (2012) 324–351
DOI: 10.1051/ps/2010013

ESAIM: Probability and Statistics
www.esaim-ps.org

SOJOURN TIME IN $\mathbb{Z}^+$ FOR THE BERNOULLI RANDOM WALK ON $\mathbb{Z}$

Aimé LACHAL[1]

Abstract. Let $(S_k)_{k\geqslant 1}$ be the classical Bernoulli random walk on the integer line with jump parameters $p \in (0,1)$ and $q = 1-p$. The probability distribution of the sojourn time of the walk in the set of non-negative integers up to a fixed time is well-known, but its expression is not simple. By modifying slightly this sojourn time through a particular counting process of the zeros of the walk as done by Chung & Feller [*Proc. Nat. Acad. Sci. USA* **35** (1949) 605–608], simpler representations may be obtained for its probability distribution. In the aforementioned article, only the symmetric case ($p = q = 1/2$) is considered. This is the discrete counterpart to the famous Paul Lévy's arcsine law for Brownian motion.

In the present paper, we write out a representation for this probability distribution in the general case together with others related to the random walk subject to a possible conditioning. The main tool is the use of generating functions.

Mathematics Subject Classification. 60G50, 60J22, 60J10, 60E10.

Received October 5, 2009. Revised March 29, 2010.

1. Introduction

Let $(X_k)_{k\geqslant 1}$ be a sequence of Bernoulli random variables with parameters $p = \mathbb{P}\{X_k = 1\} \in (0,1)$ and $q = 1-p = \mathbb{P}\{X_k = -1\}$, and $(S_k)_{k\geqslant 0}$ be the random walk defined on the set of integers $\mathbb{Z} = \{\ldots,-1,0,1,\ldots\}$ as $S_k = S_0 + \sum_{j=1}^{k} X_j$, $k \geqslant 1$ with initial location S_0. For brevity, we write $\mathbb{P}_i = \mathbb{P}\{\ldots|S_0 = i\}$ and $\mathbb{P}_0 = \mathbb{P}$.

The probability distribution of the sojourn time of the walk $(S_k)_{k\geqslant 0}$ in $\mathbb{Z}^+ = \mathbb{Z}\cap[0,+\infty)$ up to a fixed step $n \geqslant 1$, $N_n = \sum_{j=0}^{n} 1\!\!1_{\mathbb{Z}^+}(S_j) = \#\{j \in \{0,\ldots,n\} : S_j \geqslant 0\}$, is well-known. A representation for this probability distribution can be derived with the aid of Sparre Andersen's theorem (see [9,10] and, *e.g.*, [11] Chap. IV, Sect. 20). This latter can be stated as the remarkable relationship, setting $N_0 = 0$,

$$\mathbb{P}\{N_n = k\} = \mathbb{P}\{N_k = k\}\mathbb{P}\{N_{n-k} = 0\} \text{ for } 0 \leqslant k \leqslant n,$$

Keywords and phrases. Random walk, sojourn time, generating function.

[1] Institut National des Sciences Appliquées de Lyon, Pôle de Mathématiques/Institut Camille Jordan, Bâtiment Léonard de Vinci, 20 avenue Albert Einstein, 69621 Villeurbanne Cedex, France. `aime.lachal@insa-lyon.fr`; `http://maths.insa-lyon.fr`.

Article published by EDP Sciences

where the probabilities $\mathbb{P}\{N_k = 0\}$ and $\mathbb{P}\{N_k = k\}$, $k \in \mathbb{N}$, are implicitly known through their generating functions:

$$\sum_{k=0}^{\infty} \mathbb{P}\{N_k = 0\} z^k = \exp\left[\sum_{k=1}^{\infty} \mathbb{P}\{S_k < 0\} \frac{z^k}{k}\right],$$
$$\sum_{k=0}^{\infty} \mathbb{P}\{N_k = k\} z^k = \exp\left[\sum_{k=1}^{\infty} \mathbb{P}\{S_k \geqslant 0\} \frac{z^k}{k}\right].$$

Nevertheless, the result is not so simple. Rescaling the random walk and passing to the limit, we get Paul Lévy's famous arcsine law for Brownian motion.

By modifying slightly the counting process of the positive terms of the random walk as done by Chung & Feller (see [4] and, *e.g.*, [5], Chap. III, Sect. 4 and [8], Chap. 8, Sect. 11), an alternative sojourn time of the walk $(S_k)_{k\in\mathbb{N}}$ in $\mathbb{Z}^+$ up to n can be defined as $T_n = \sum_{j=1}^{n} \delta_j$ with

$$\delta_j = \begin{cases} 1 & \text{if } (S_j > 0) \text{ or } (S_j = 0 \text{ and } S_{j-1} > 0), \\ 0 & \text{if } (S_j < 0) \text{ or } (S_j = 0 \text{ and } S_{j-1} < 0). \end{cases}$$

We put $T_0 = 0$. We obviously have $0 \leqslant T_n \leqslant n$. In T_n, $n \geqslant 1$, one counts each step j such that $S_j > 0$ and only those steps such that $S_j = 0$ which correspond to a downstep: $S_{j-1} = 1$. This convention is described in [4] (and, *e.g.*, in [5] and [8]) in the symmetric case $p = q = 1/2$ when n is an even integer and, as written in [4] – *"The elegance of the results to be announced depends on this convention"* (*sic*) –, it produces a remarkable result. Indeed, in this case, the sojourn time is even and its probability distribution takes the simple following form: for even integers k such that $0 \leqslant k \leqslant n$, as in Sparre Andersen's theorem,

$$\mathbb{P}\{T_n = k\} = \mathbb{P}\{T_k = k\}\mathbb{P}\{T_{n-k} = 0\} = \frac{1}{2^n}\binom{k}{k/2}\binom{(n-k)}{(n-k)/2}.$$

In this paper, we derive explicit expressions for the probability distribution of T_n in any case, that is for any $p \in (0, 1)$ and any integer (even or odd) $n \geqslant 1$. The main results are displayed in Theorems 4.2 and 5.3. We also compute the distribution of T_n under various constraints at the last step: $S_n = 0$, $S_n > 0$ or $S_n < 0$. The constraint $S_n = 0$ (with $S_0 = 0$) corresponds to the bridge of the random walk. The related results are respectively included in Theorems 6.2 and 7.3. The main tool for this study is the use of generating functions, excursions theory associated with random walks together with clever algebra. The intermediate results are contained in Theorems 4.1, 5.1, 6.1 and 7.1.

Finally, by rescaling suitably the random walk, we retrieve the distribution of the sojourn time in $(0, +\infty)$ for the Brownian motion with a possible drift. This includes of course Paul Lévy's arcsine law for Brownian motion without drift.

Although this problem is old and classical, we are surprised not to have found any related reference in the literature.

2. Settings and mathematical background

2.1. Some preliminary identities

Let $\mathbb{N} = \mathbb{Z}^+ = \{0, 1, 2, \ldots\}$ be the usual set of non-negative integers, $\mathbb{N}^* = \mathbb{Z}^{+*} = \mathbb{N} \setminus \{0\} = \{1, 2, 3, \ldots\}$ that of positive integers, $\mathbb{Z}^{-*} = \{\ldots, -3, -2, -1\}$ that of negative integers and $\mathbb{Z}^* = \mathbb{Z} \setminus \{0\}$. Let $\mathcal{E} = \{0, 2, 4, \ldots\}$ denote the set of even non-negative integers, $\mathcal{E}^* = \mathcal{E} \setminus \{0\} = \{2, 4, 6, \ldots\}$ the set of even positive integers and $\mathcal{O} = \{1, 3, 5, \ldots\}$ the set of odd positive integers. Set, for suitable real z,

$$A(z) = \sqrt{1 - 4z^2} \quad \text{and} \quad \mathbf{A}(z) = \sqrt{1 - 4pqz^2}.$$

Set also for $i \in \mathcal{E}$

$$a_i = \frac{1}{i+2}\binom{i}{i/2} = \frac{1}{4(i+1)}\binom{i+2}{(i+2)/2} \quad \text{and} \quad b_i = (i+2)a_i = \binom{i}{i/2}.$$

The a_i's are closely related to the famous Catalan numbers: $a_i = C_{i/2}/2$. We shall make use of the following elementary identities.

Proposition 2.1. *For any z such that $|z| < 1/2$,*

$$\begin{aligned} A(z) &= \sqrt{1-4z^2} = -\sum_{i\in\mathcal{E}} \frac{1}{i-1}\binom{i}{i/2} z^i = 1 - 4\sum_{i\in\mathcal{E}^*} a_{i-2} z^i, \\ \frac{1}{A(z)} &= \frac{1}{\sqrt{1-4z^2}} = \sum_{i\in\mathcal{E}} \binom{i}{i/2} z^i = \sum_{i\in\mathcal{E}} b_i z^i. \end{aligned} \tag{2.1}$$

We have the following convolution relationships.

Proposition 2.2. *For any even integer $i \geqslant 0$,*

$$\sum_{\substack{j\in\mathcal{E}:\\ j\leqslant i}} a_j a_{i-j} = \frac{1}{2}\, a_{i+2} \quad \textit{and} \quad \sum_{\substack{j\in\mathcal{E}:\\ j\leqslant i}} a_j b_{i-j} = \frac{1}{4}\, b_{i+2}. \tag{2.2}$$

Proof. The generating function of the left-hand side of the first equality in (2.2) can be evaluated as

$$\sum_{i\in\mathcal{E}} \left(\sum_{\substack{j\in\mathcal{E}:\\ j\leqslant i}} a_j a_{i-j} \right) z^i = \left(\sum_{i\in\mathcal{E}} a_i z^i \right)^2 = \left(\frac{1-A(z)}{4z^2} \right)^2 = \frac{1}{8z^4}(1-2z^2-A(z)) = \frac{1}{2}\sum_{i\in\mathcal{E}} a_{i+2} z^i.$$

By identification of the coefficients of the foregoing generating functions, we immediately obtained the first equality in (2.2). Analogously, for the second equality in (2.2),

$$\sum_{i\in\mathcal{E}} \left(\sum_{\substack{j\in\mathcal{E}:\\ j\leqslant i}} a_j b_{i-j} \right) z^i = \left(\sum_{i\in\mathcal{E}} a_i z^i \right)\left(\sum_{i\in\mathcal{E}} b_i z^i \right) = \frac{1-A(z)}{4z^2 A(z)} = \frac{1}{4z^2 A(z)} - \frac{1}{4z^2} = \frac{1}{4}\sum_{i\in\mathcal{E}} b_{i+2} z^i$$

and the second equality in (2.2) holds. □

We shall also use the identity below.

Proposition 2.3. *For any x and y such that $|x| < 1/2$ and $|y| < 1/2$,*

$$\frac{1}{A(x)+A(y)} = \sum_{i,j\in\mathcal{E}} a_{i+j}\, x^i y^j. \tag{2.3}$$

Proof. Let us write $\dfrac{1}{A(x)+A(y)} = \dfrac{A(x)-A(y)}{4(y^2-x^2)}$. On one hand,

$$A(x) - A(y) = 4\sum_{k\in\mathcal{E}} a_k (y^{k+2} - x^{k+2}).$$

On the other hand,

$$\frac{y^{k+2}-x^{k+2}}{y^2-x^2} = \sum_{\substack{i,j\in\mathcal{E}:\\ i+j=k}} x^i y^j.$$

Therefore,

$$\frac{1}{A(x)+A(y)} = \sum_{k\in\mathcal{E}} a_k \sum_{\substack{i,j\in\mathcal{E}:\\ i+j=k}} x^i y^j = \sum_{i,j\in\mathcal{E}} a_{i+j}\, x^i y^j. \qquad \square$$

2.2. Some well-known identities on random walks

Here, we recall several well-known formulas in the theory of random walks. We refer, *e.g.*, to [11]. We have, for $j\in\mathbb{Z}$ and $k\in\mathbb{N}$ such that $|j|\leqslant k$ and $k-j\in\mathcal{E}$,

$$\mathbb{P}\{S_k=j\} = \binom{k}{(j+k)/2} p^{(j+k)/2} q^{(k-j)/2}. \tag{2.4}$$

Using the representation $\binom{k}{(k+1)/2} = b_{k+1}/2$, we have in particular

$$\mathbb{P}\{S_k=0\} = \begin{cases} b_k (pq)^{k/2} & \text{if } k \text{ is even,}\\ 0 & \text{if } k \text{ is odd,}\end{cases}$$

$$\mathbb{P}\{S_k=1\} = \mathbb{P}_{-1}\{S_k=0\} = \begin{cases} \frac12\, b_{k+1}\, p^{(k+1)/2} q^{(k-1)/2} & \text{if } k \text{ is odd,}\\ 0 & \text{if } k \text{ is even,}\end{cases} \tag{2.5}$$

$$\mathbb{P}\{S_k=-1\} = \mathbb{P}_{1}\{S_k=0\} = \begin{cases} \frac12\, b_{k+1}\, p^{(k-1)/2} q^{(k+1)/2} & \text{if } k \text{ is odd,}\\ 0 & \text{if } k \text{ is even.}\end{cases}$$

We define several generating functions. For $j\in\mathbb{Z}$, let H_j be the generating function of the $\mathbb{P}\{S_k=j\}, k\in\mathbb{N}$, and, for $i\in\mathbb{Z}$, H_i^F be the generating function of the $\mathbb{P}_i\{S_k\in F\}$, $k\in\mathbb{N}$:

$$H_j(z) = \sum_{k=0}^{\infty} \mathbb{P}\{S_k=j\} z^k \quad \text{and} \quad H_i^F(z) = \sum_{k=0}^{\infty} \mathbb{P}_i\{S_k\in F\} z^k = \sum_{j\in F-i} H_j(z) \tag{2.6}$$

where $F-i$ is the set of the numbers of the form $j-i, j\in F$. We explicitly have

$$H_j(z) = \begin{cases} \dfrac{1}{\mathbf{A}(z)} \left(\dfrac{1-\mathbf{A}(z)}{2qz}\right)^j = \dfrac{1}{\mathbf{A}(z)} \left(\dfrac{2pz}{1+\mathbf{A}(z)}\right)^j & \text{if } j\geqslant 0,\\[2ex] \dfrac{1}{\mathbf{A}(z)} \left(\dfrac{1-\mathbf{A}(z)}{2pz}\right)^{|j|} = \dfrac{1}{\mathbf{A}(z)} \left(\dfrac{2qz}{1+\mathbf{A}(z)}\right)^{|j|} & \text{if } j\leqslant 0.\end{cases} \tag{2.7}$$

We need to introduce the first hitting time of a level $a\in\mathbb{Z}$ for the random walk: $\tau_a=\min\{n\geqslant 1 : S_n=a\}$. The probability distribution of τ_a for $a\in\mathbb{Z}^*$ can be expressed by means of the probabilities $\mathbb{P}\{S_k=a\}$, $k\in\mathbb{N}$, as

$$\mathbb{P}\{\tau_a=k\} = \frac{|a|}{k}\,\mathbb{P}\{S_k=a\} = \frac{|a|}{k}\binom{k}{(k+a)/2} p^{(k+a)/2} q^{(k-a)/2} \text{ for } k\geqslant |a|. \tag{2.8}$$

In some particular cases, we have for $k\in\mathcal{E}^*$:

$$\mathbb{P}\{\tau_0=k\} = \frac{1}{k-1}\binom{k}{k/2}(pq)^{k/2}$$

and for $k \in \mathcal{O}$:

$$\mathbb{P}\{\tau_1 = k\} = \frac{1}{k}\binom{k}{(k+1)/2} p^{(k+1)/2} q^{(k-1)/2},$$
$$\mathbb{P}\{\tau_{-1} = k\} = \frac{1}{k}\binom{k}{(k+1)/2} p^{(k-1)/2} q^{(k+1)/2}.$$

We sum up these formulas as follows:

- for $k \in \mathcal{E}^*$,

$$\mathbb{P}\{\tau_0 = k\} = 4a_{k-2}(pq)^{k/2}; \tag{2.9}$$

- for $k \in \mathcal{O}$,

$$\begin{aligned}\mathbb{P}\{\tau_1 = k\} = \mathbb{P}_{-1}\{\tau_0 = k\} = \frac{1}{2q}\,\mathbb{P}\{\tau_0 = k+1\} = 2p\,a_{k-1}(pq)^{(k-1)/2},\\ \mathbb{P}\{\tau_{-1} = k\} = \mathbb{P}_{1}\{\tau_0 = k\} = \frac{1}{2p}\,\mathbb{P}\{\tau_0 = k+1\} = 2q\,a_{k-1}(pq)^{(k-1)/2}.\end{aligned} \tag{2.10}$$

Remark 2.4. The convolution identities (2.2) can be interpreted as Darling–Siegert-type equations (see, *e.g.*, [3]) which are due to the Markov property of the random walk. More precisely, the second identity of (2.2) is the analytic form of the probabilistic equality

$$\mathbb{P}\{S_n = 0\} = \sum_{\substack{j\in\mathcal{E}:\\ j\leqslant n}} \mathbb{P}\{\tau_0 = j\}\mathbb{P}\{S_{n-j} = 0\}$$

which is obtained by remarking that a trajectory starting at zero and terminating at zero at time n necessarily passes through zero at a time equal to or less than n and then $\tau_0 \leqslant n$. The first identity of (2.2) is the analytic form of the probabilistic equality

$$\mathbb{P}\{\tau_2 = n\} = \sum_{\substack{j\in\mathcal{O}:\\ j\leqslant n-1}} \mathbb{P}\{\tau_1 = j\}\mathbb{P}\{\tau_1 = n-j\}$$

which is obtained by observing that a trajectory starting at zero and passing at level two at time n for the first time necessarily crosses level one at a time less than n: $\tau_1 < n$.

The corresponding generating functions K_a defined as

$$K_a(z) = \mathbb{E}(z^{\tau_a}) = \sum_{k=1}^{\infty} \mathbb{P}\{\tau_a = k\} z^k$$

are explicitly expressed by

$$K_a(z) = \begin{cases} \left(\dfrac{1-\mathbf{A}(z)}{2qz}\right)^a & \text{if } a \geqslant 1,\\[2ex] \left(\dfrac{1-\mathbf{A}(z)}{2pz}\right)^{|a|} & \text{if } a \leqslant -1,\\[2ex] 1-\mathbf{A}(z) & \text{if } a = 0. \end{cases} \tag{2.11}$$

We state a last elementary result. Noticing, by the Markov property, that $\mathbb{P}\{\tau_0 = j, S_1 > 0\} = p\,\mathbb{P}_1\{\tau_0 = j-1\}$ and $\mathbb{P}\{\tau_0 = j, S_1 < 0\} = q\,\mathbb{P}_1\{\tau_0 = j-1\}$, we get, by (2.10), $\mathbb{P}\{\tau_0 = j, S_1 > 0\} = \mathbb{P}\{\tau_0 = j, S_1 < 0\} = \frac{1}{2}\mathbb{P}\{\tau_0 = j\}$ and then

$$\mathbb{E}(z^{\tau_a}1\!\!1_{\{S_1>0\}}) = \mathbb{E}(z^{\tau_a}1\!\!1_{\{S_1<0\}}) = \frac{1}{2}(1-\mathbf{A}(z)). \tag{2.12}$$

2.3. Purpose of the article

The aim of this paper is to compute the probabilities

$$r_{k,n}^F = \mathbb{P}\{T_n = k, S_n \in F\}, 0 \leqslant k \leqslant n,$$

in the four cases $F = \mathbb{Z}$, $F = \{0\}$, $F = \mathbb{Z}^{+*}$ and $F = \mathbb{Z}^{-*}$. The corresponding results are displayed in Theorems 5.1, 6.1 and 7.1. The adopted way consists of first calculating the generating function G_F of the $r_{k,n}^F$'s:

$$G^F(x,y) = \sum_{\substack{k,n\in\mathbb{N}:\\ k\leqslant n}} r_{k,n}^F\, x^k y^{n-k} = \sum_{n\in\mathbb{N}} \mathbb{E}\left(x^{T_n} y^{n-T_n} 1\!\!1_{\{S_n\in F\}}\right)$$

with the help of the excursions theory associated with random walks, and next of inverting this function. Notice that $r_{0,0}^F = 1\!\!1_F(0)$ and for $n \geqslant 1$, the events $\{T_n = 0\}$ and $\{T_n = n\}$ respectively coincide with $\{S_1 \leqslant 0, \ldots, S_n \leqslant 0\} = \{\tau_1 > n\}$ and $\{S_1 \geqslant 0, \ldots, S_n \geqslant 0\} = \{\tau_{-1} > n\}$. Therefore, for $n \geqslant 1$,

$$r_{0,n}^F = \mathbb{P}\{\tau_1 > n, S_n \in F\} \quad \text{and} \quad r_{n,n}^F = \mathbb{P}\{\tau_{-1} > n, S_n \in F\}. \tag{2.13}$$

The particular probabilities $r_{0,n}^F$ and $r_{n,n}^F$, $n \in \mathbb{N}$, are generated by the partial functions of G^F, namely

$$G^F(x,0) = \sum_{n\in\mathbb{N}} r_{n,n}^F\, x^k \quad \text{and} \quad G^F(0,y) = \sum_{n\in\mathbb{N}} r_{0,n}^F\, y^n. \tag{2.14}$$

3. Generating function

Theorem 3.1. *The generating function G^F can be written as*

$$G^F(x,y) = \frac{[1+\mathbf{A}(x)]G^F(x,0) + [1+\mathbf{A}(y)]G^F(0,y) - 2\,1\!\!1_F(0)}{\mathbf{A}(x)+\mathbf{A}(y)} \tag{3.1}$$

where $\mathbf{A}(z) = \sqrt{1-4pqz^2}$. The quantities $G^F(x,0)$ and $G^F(0,y)$ can be expressed as

$$G^F(x,0) = H_0^F(x) - \frac{1-\mathbf{A}(x)}{2px} H_{-1}^F(x) \quad \text{and} \quad G^F(0,y) = H_0^F(y) - \frac{1-\mathbf{A}(y)}{2qy} H_1^F(x) \tag{3.2}$$

where H_0^F, H_1^F and H_{-1}^F are defined by (2.6).

Proof. Let us introduce the successive passage times by 0 recursively defined by $N_0 = 0$ and for $m \in \mathbb{N}$, $N_{m+1} = \min\{n \geqslant N_m + 1 : S_n = 0\}$. We decompose the random walk into excursions and this yields the following decomposition for the generating function G^F:

$$G^F(x,y) = \sum_{n\in\mathbb{N}} \mathbb{E}\left(x^{T_n} y^{n-T_n} 1\!\!1_{\{S_n\in F\}}\right) = 1\!\!1_F(0) + \mathbb{E}\left[\sum_{m\in\mathbb{N}} \sum_{n=N_m+1}^{N_{m+1}} x^{T_n} y^{n-T_n} 1\!\!1_{\{S_n\in F\}}\right].$$

For $N_m+1 \leqslant n \leqslant N_{m+1}$, we have $T_n = T_{N_m} + (n-N_m)1\!\!1_{\{e_m \geqslant 0\}}$ where e_m is the excursion of the random walk on the interval $[N_m, N_{m+1}]$. In particular, e_0 is the first excursion, it is defined on the interval $[0,\tau_0]$. The fact that $e_0 \geqslant 0$ (respectively $e_0 \leqslant 0$) is equivalent to that $S_1 > 0$ (respectively $S_1 < 0$). We have also $n - T_n = N_m - T_{N_m} + (n-N_m)1\!\!1_{\{e_m \leqslant 0\}}$ and then, by the independence of the excursions,

$$\begin{aligned} G^F(x,y) &= 1\!\!1_F(0) + \mathbb{E}\bigg[\sum_{m\in\mathbb{N}} x^{T_{N_m}} y^{N_m - T_{N_m}} \sum_{n=N_m+1}^{N_{m+1}} \big(x^{n-N_m} 1\!\!1_{\{e_m\geqslant 0\}} + y^{n-N_m} 1\!\!1_{\{e_m \leqslant 0\}}\big)\, 1\!\!1_{\{S_n\in F\}}\bigg] \\ &= 1\!\!1_F(0) + \mathbb{E}\bigg[\sum_{m\in\mathbb{N}} x^{T_{N_m}} y^{N_m - T_{N_m}} \sum_{n=1}^{N_{m+1}-N_m} \big(x^{n} 1\!\!1_{\{e_m\geqslant 0\}} + y^{n} 1\!\!1_{\{e_m \leqslant 0\}}\big)\, 1\!\!1_{\{S_{n+N_m}\in F\}}\bigg] \\ &= 1\!\!1_F(0) + \mathbb{E}\bigg[\sum_{m\in\mathbb{N}} x^{T_{N_m}} y^{N_m - T_{N_m}}\bigg] \mathbb{E}\bigg[\sum_{n=1}^{\tau_0} \big(x^{n} 1\!\!1_{\{e_0\geqslant 0\}} + y^{n} 1\!\!1_{\{e_0 \leqslant 0\}}\big)\, 1\!\!1_{\{S_n\in F\}}\bigg]. \end{aligned}$$

Hence, we have obtained a representation of the form

$$G^F(x,y) = 1\!\!1_F(0) + G^0(x,y)[B_+^F(x) + B_-^F(y)] \tag{3.3}$$

with

$$G^0(x,y) = \mathbb{E}\bigg[\sum_{m\in\mathbb{N}} x^{T_{N_m}} y^{N_m - T_{N_m}}\bigg]$$

and

$$B_+^F(x) = \mathbb{E}\bigg[\sum_{n=1}^{\tau_0} 1\!\!1_{\{S_1>0, S_n\in F\}} x^n\bigg], \quad B_-^F(y) = \mathbb{E}\bigg[\sum_{n=1}^{\tau_0} 1\!\!1_{\{S_1<0, S_n\in F\}} y^n\bigg].$$

Actually, G^0 is nothing but $G^{\{0\}}$ (*i.e.*, G^F associated with $F=\{0\}$) as it is immediately seen by definition. The quantity $G^0(x,y)$ can be easily evaluated as follows. Set $u_m = \mathbb{E}\left(x^{T_{N_m}} y^{N_m - T_{N_m}}\right)$. We have $u_0 = 0$ and for $m \in \mathbb{N}$, $u_{m+1} = u_m\, \mathbb{E}\big(x^{\tau_0} 1\!\!1_{\{S_1>0\}} + y^{\tau_0} 1\!\!1_{\{S_1<0\}}\big)$. Now, by (2.12),

$$\mathbb{E}\big(x^{\tau_0} 1\!\!1_{\{S_1>0\}} + y^{\tau_0} 1\!\!1_{\{S_1<0\}}\big) = \frac{1}{2}\,[2 - \mathbf{A}(x) - \mathbf{A}(y)].$$

As a byproduct, the series $(\sum_{m\in\mathbb{N}} u_m)$ is geometrical with ratio $[2-\mathbf{A}(x)-\mathbf{A}(y)]/2$ and its sum is given by $2/[\mathbf{A}(x)+\mathbf{A}(y)]$. This yields the following result which will be stated in Theorem 6.1:

$$G^0(x,y) = \frac{2}{\mathbf{A}(x)+\mathbf{A}(y)}\cdot$$

Thus, (3.3) can be written as

$$G^F(x,y) = 1\!\!1_F(0) + \frac{2B_+^F(x) + 2B_-^F(y)}{\mathbf{A}(x)+\mathbf{A}(y)}\cdot \tag{3.4}$$

We then deduce

$$G^F(x,0) = 1\!\!1_F(0) + \frac{2B_+^F(x)}{1+\mathbf{A}(x)}, \quad G^F(0,y) = 1\!\!1_F(0) + \frac{2B_-^F(y)}{1+\mathbf{A}(y)}$$

from which we extract

$$2B_+^F(x) = [1+\mathbf{A}(x)][G^F(x,0) - 1\!\!1_F(0)], \quad 2B_-^F(y) = [1+\mathbf{A}(y)][G^F(0,y) - 1\!\!1_F(0)]. \tag{3.5}$$

Finally, putting (3.5) into (3.4) gives (3.1).

Let us check (3.2). In view of (2.10), the probability $r^F_{n,n}$ admits the following representations:

$$\begin{aligned} r^F_{n,n} &= \mathbb{P}\{S_n \in F\} - \mathbb{P}\{\tau_{-1} \leqslant n, S_n \in F\} = \mathbb{P}\{S_n \in F\} - \sum_{j=1}^{n} \mathbb{P}\{\tau_{-1} = j\}\mathbb{P}_{-1}\{S_{n-j} \in F\} \\ &= \mathbb{P}\{S_n \in F\} - \frac{1}{2p} \sum_{\substack{j\in\mathcal{O}:\\ j\leqslant n}} \mathbb{P}\{\tau_0 = j+1\}\mathbb{P}_{-1}\{S_{n-j} \in F\}. \end{aligned} \tag{3.6}$$

Therefore, we have by (2.14) and (3.6)

$$G^F(x,0) = \sum_{n\in\mathbb{N}} r^F_{n,n} x^n = \sum_{n\in\mathbb{N}} \mathbb{P}\{S_n \in F\}x^n - \frac{1}{2p} \sum_{n\in\mathbb{N}^*} \sum_{\substack{j\in\mathcal{O}:\\ j\leqslant n}} \mathbb{P}\{\tau_0 = j+1\}\mathbb{P}_{-1}\{S_{n-j} \in F\}x^n.$$

The first sum in the above equality is $H^F_0(x)$ while the second can be computed as follows: by (2.10),

$$\begin{aligned} \sum_{n\in\mathbb{N}^*} \sum_{\substack{j\in\mathcal{O}:\\ j\leqslant n}} \mathbb{P}\{\tau_0 = j+1\}\mathbb{P}_{-1}\{S_{n-j} \in F\}x^n &= \sum_{j\in\mathcal{O}} \mathbb{P}\{\tau_0 = j+1\}x^j \sum_{\substack{n\in\mathbb{N}^*:\\ n\geqslant j}} \mathbb{P}_{-1}\{S_{n-j} \in F\}x^{n-j} \\ &= 2p \sum_{j\in\mathcal{O}} \mathbb{P}\{\tau_{-1} = j\}x^j \sum_{n\in\mathbb{N}} \mathbb{P}_{-1}\{S_n \in F\}x^n \\ &= 2pK_{-1}(x)H^F_{-1}(x) \end{aligned}$$

and the expression of $G^F(x,0)$ in (3.2) ensues. The derivation of that of $G^F(0,y)$ is quite similar. The Proof of Theorem 3.1 is finished. □

In [7], we propose a Proof of (3.1) based on a recursive relationship concerning the probabilities $r_{k,n}$ which is similar to the proof originally described in [4]. We also refer the reader to the book [6] for a combinatorial approach for tackling such problems.

4. The case $F = \{j\}$ for $j \in \mathbb{Z}^*$

We suppose that $F = \{j\}$ for a fixed $j \in \mathbb{Z}^*$. So, we are dealing with a random walk with a prescribed location after the nth step. The case where $j = 0$ will be considered in Section 6. We set for simplicity $r^{\{j\}}_{k,n} = r^j_{k,n}$ and $G^{\{j\}}(x,y) = G^j(x,y)$, $H^{\{j\}}_i(z) = H^j_i(z)$.

4.1. **Generating function**

In order to write the generating function $G^j(x,y)$, in view of (3.1), we need to know that $H^j_i(z) = H_{j-i}(z)$ and to evaluate the functions $G^j(x,0)$ and $G^j(0,y)$. We have

$$G^j(x,0) = H^j_0(x) - \frac{1-\mathbf{A}(x)}{2px}\, H^j_{-1}(x).$$

But

$$H^j_{-1}(x) = \begin{cases} \dfrac{2px}{1+\mathbf{A}(x)}\, H^j_0(x) & \text{if } j \in \mathbb{Z}^{+*}, \\[2ex] \dfrac{2px}{1-\mathbf{A}(x)}\, H^j_0(x) & \text{if } j \in \mathbb{Z}^{-*}. \end{cases}$$

Then, for $j \in \mathbb{Z}^{+*}$,

$$G^j(x,0) = H_0^j(x)\left[1 - \frac{1-\mathbf{A}(x)}{2px}\,\frac{2px}{1+\mathbf{A}(x)}\right] = \frac{2K_j(x)}{1+\mathbf{A}(x)}$$

and, for $j \in \mathbb{Z}^{+*}$,

$$G^j(x,0) = H_0^j(x)\left[1 - \frac{1-\mathbf{A}(x)}{2px}\,\frac{2px}{1-\mathbf{A}(x)}\right] = 0.$$

Similarly, we have

$$G^j(0,y) = \begin{cases} \dfrac{2K_j(y)}{1+\mathbf{A}(y)} & \text{if } j \in \mathbb{Z}^{-*}, \\ 0 & \text{if } j \in \mathbb{Z}^{+*}. \end{cases}$$

From this and (3.1), we immediately derive the function $G(x,y)$.

Theorem 4.1. *The generating function G is given by*

$$G^j(x,y) = \begin{cases} \dfrac{2K_j(x)}{\mathbf{A}(x)+\mathbf{A}(y)} & \textit{if } j \in \mathbb{Z}^{+*} \\[2ex] \dfrac{2K_j(y)}{\mathbf{A}(x)+\mathbf{A}(y)} & \textit{if } j \in \mathbb{Z}^{-*} \end{cases} \tag{4.1}$$

where $\mathbf{A}(z) = \sqrt{1-4pqz^2}$.

4.2. Distribution of the sojourn time

We now invert the generating function G^j given by (4.1) in order to derive the coefficients $r^j_{k,n}$.

Theorem 4.2. *The probability $r^j_{k,n} = \mathbb{P}\{T_n = k, S_n = j\}$ admits the following expression: for $0 \leqslant k \leqslant n$ such that $n-k$ is even,*
if $j \geqslant 1$:

$$r^j_{k,n} = \begin{cases} 2j \displaystyle\sum_{\substack{i\in\mathbb{N}:\\ j\leqslant i\leqslant k,\\ i-j\in\mathcal{E}}} \frac{a_{n-i}}{i}\binom{i}{(i+j)/2} p^{(n+j)/2} q^{(n-j)/2} & \textit{if } j \leqslant k \textit{ and } k-j \textit{ is even,} \\ 0 & \textit{if } j > k \textit{ or } k-j \textit{ is odd;} \end{cases} \tag{4.2}$$

if $j \leqslant -1$:

$$r^j_{k,n} = \begin{cases} 2|j| \displaystyle\sum_{\substack{i\in\mathbb{N}:\\ |j|\leqslant i\leqslant k,\\ i+j\in\mathcal{E}}} \frac{a_{n-i}}{i}\binom{i}{(i+j)/2} p^{(n+j)/2} q^{(n-j)/2} & \textit{if } j \geqslant k-n \textit{ and } k-j \textit{ is even,} \\ 0 & \textit{if } j < k-n \textit{ or } k-j \textit{ is odd,} \end{cases} \tag{4.3}$$

where $a_i = \dfrac{1}{i+2}\dbinom{i}{i/2}$ *for* $i \in \mathcal{E}$.

Proof. Assume first that $j \geqslant 1$. We expand $G^j(x,y)$ by using (2.3):

$$G^j(x,y) = 2\sum_{i=0}^{\infty} \mathbb{P}\{\tau_j = i\} x^i \sum_{l,m\in\mathcal{E}} a_{l+m}\, x^l y^m = 2 \sum_{\substack{i,l,m\in\mathbb{N}:\\ l,m\in\mathcal{E}}} a_{l+m}\, \mathbb{P}\{\tau_j = i\}\, x^{i+l} y^m.$$

By performing the transformations $i+l=k$ and $m=n-k$ in the last sum, we get

$$G^j(x,y) = 2 \sum_{\substack{i,k,n\in\mathbb{N}:\\ k-i\in\mathcal{E},n-k\in\mathcal{E}}} a_{n-i}\,\mathbb{P}\{\tau_j=i\}\,x^k y^{n-k} = \sum_{\substack{k,n\in\mathbb{N}:\\ n-k\in\mathcal{E}}} \Bigg(\sum_{\substack{i\in\mathbb{N}:\\ k-i\in\mathcal{E}}} 2a_{n-i}\,\mathbb{P}\{\tau_j=i\}\Bigg) x^k y^{n-k}.$$

We finally obtain (4.2) by identifying the coefficients of the two expansions of G^j and using (2.8). For the case where $j\leqslant -1$, we invoke an argument of duality which is explained in Remark 4.3 below. The expression of $r^j_{k,n}$ for $j\leqslant -1$ can be deduced from that related to the case where $j\geqslant 1$ by interchanging p and q, k and $n-k$, j and $-j$, and this proves (4.3). □

Remark 4.3. Let us introduce the dual random walk $(S^*_n)_{n\geqslant 0}$ with $S^*_k=\sum_{j=1}^k X^*_j=-S_k$. This sequence is the Bernoulli random walk with interchanged parameters $q=\mathbb{P}\{X^*_j=+1\}$ and $p=\mathbb{P}\{X^*_j=-1\}$. The corresponding sojourn time is defined as $T^*_n=\sum_{j=1}^n \delta^*_j$ with

$$\delta^*_j = \begin{cases} 1 & \text{if } (S^*_j>0) \text{ or } (S^*_j=0 \text{ and } S^*_{j-1}>0),\\ 0 & \text{if } (S^*_j<0) \text{ or } (S^*_j=0 \text{ and } S^*_{j-1}<0),\end{cases}$$
$$= \begin{cases} 1 & \text{if } (S_j<0) \text{ or } (S_j=0 \text{ and } S_{j-1}<0),\\ 0 & \text{if } (S_j>0) \text{ or } (S_j=0 \text{ and } S_{j-1}>0).\end{cases}$$

We see that $\delta^*_j=1-\delta_j$ and then $T^*_n=n-T_n$ which implies $r^j_{k,n}=\mathbb{P}\{T^*_n=n-k,S^*_n=-j\}$. As a result, the probability $r^j_{k,n}$ can be deduced from the probability $r^{-j}_{n-k,n}$ by interchanging p and q.

5. The case $F=\mathbb{Z}$ (random walk without conditioning)

In this part, we study the sojourn time without conditioning the extremity of the random walk. This corresponds to the case $F=\mathbb{Z}$. We set for simplifying the notations $r^{\mathbb{Z}}_{k,n}=r_{k,n}$ and $G^{\mathbb{Z}}(x,y)=G(x,y)$. A possible expression for the $r_{k,n}$'s can be obtained from Theorem 4.2 by summing the $r^j_{k,n}$, $j\in\mathbb{Z}$. Actually, $r^j_{k,n}=0$ for $|j|>n$; so,

$$r_{k,n}=\sum_{j=-n}^{n} r^j_{k,n}.$$

We propose another representation which can be deduced from the generating function G.

5.1. **Generating function**

In order to derive the generating function $G(x,y)$, in view of (3.1), we need to evaluate the functions $H^{\mathbb{Z}}_i(z)$, $G(x,0)$ and $G(0,y)$. On one hand, by definition,

$$H^{\mathbb{Z}}_i(z)=\sum_{k=0}^{\infty}\mathbb{P}_i\{S_k\in\mathbb{Z}\}z^k=\sum_{k=0}^{\infty}z^k=\frac{1}{1-z}\cdot$$

On the other hand, by (3.2),

$$G(x,0)=H^{\mathbb{Z}}_0(x)-\frac{1-\mathbf{A}(x)}{2px}\,H^{\mathbb{Z}}_{-1}(x)=\frac{1}{1-x}\left[1-\frac{1-\mathbf{A}(x)}{2px}\right]=\frac{2px-1+\mathbf{A}(x)}{2px(1-x)}\cdot$$

Similarly,

$$G(0,y)=\frac{2qy-1+\mathbf{A}(y)}{2qy(1-y)}\cdot$$

We can then write out $G(x,y)$.

Theorem 5.1. *The generating function G is given by*

$$G(x,y)=\frac{1}{\mathbf{A}(x)+\mathbf{A}(y)}\left[\frac{2p-1+\mathbf{A}(x)}{1-x}+\frac{2q-1+\mathbf{A}(y)}{1-y}\right] \tag{5.1}$$

where $\mathbf{A}(z)=\sqrt{1-4pqz^2}$.

Proof. We have by (3.1)

$$G(x,y)=\frac{[1+\mathbf{A}(x)]G(x,0)+[1+\mathbf{A}(y)]G(0,y)-2}{\mathbf{A}(x)+\mathbf{A}(y)}\cdot$$

Let us compute the terms $[1+\mathbf{A}(x)]G(x,0)$ and $[1+\mathbf{A}(y)]G(0,y)$. We have

$$[1+\mathbf{A}(x)][2px-1+\mathbf{A}(x)]=2px-1+2px\mathbf{A}(x)+\mathbf{A}(x)^2=2px[1-2qx+\mathbf{A}(x)]$$

and then

$$[1+\mathbf{A}(x)]G(x,0)=\frac{1-2qx+\mathbf{A}(x)}{1-x}\cdot$$

Similarly,

$$[1+\mathbf{A}(y)]G(0,y)=\frac{1-2py+\mathbf{A}(y)}{1-y}\cdot$$

Therefore, we get

$$\begin{aligned}
[1+\mathbf{A}(x)]G(x,0)+[1+\mathbf{A}(y)]G(0,y)-2&=\left[\frac{1-2qx}{1-x}+\frac{1-2py}{1-y}-2\right]+\frac{\mathbf{A}(x)}{1-x}+\frac{\mathbf{A}(y)}{1-y}\\
&=\left[\left(\frac{1-2qx}{1-x}-2q\right)+\left(\frac{1-2py}{1-y}-2p\right)\right]+\frac{\mathbf{A}(x)}{1-x}+\frac{\mathbf{A}(y)}{1-y}\\
&=\frac{2p-1+\mathbf{A}(x)}{1-x}+\frac{2q-1+\mathbf{A}(y)}{1-y}
\end{aligned}$$

from which we obtain (5.1). □

An interesting consequence of Theorem 5.1 concerns the "partial" generating function $\tilde{G}$ of the probabilities $r_{k,n}$ limited to the even indices k and n:

$$\tilde{G}(x,y)=\sum_{k,n\in\mathcal{E}}r_{k,n}x^ky^{n-k}.$$

For this function, we have the simple result below.

Corollary 5.2. *The generating function $\tilde{G}$ is given by*

$$\tilde{G}(x,y)=\frac{4pq}{[1-2p+\mathbf{A}(x)][1-2q+\mathbf{A}(y)]} \tag{5.2}$$

where $\mathbf{A}(z) = \sqrt{1-4pqz^2}$. *In particular,*

$$\tilde{G}(x,y) = \tilde{G}(x,0)\tilde{G}(0,y). \tag{5.3}$$

Proof. We first try to relate $\tilde{G}$ to G. For this, we observe that

$$\frac{1}{2}[G(x,y)+G(-x,-y)] = \sum_{\substack{k,n\in\mathbb{N}:\\ k\leqslant n}} \frac{1+(-1)^n}{2} r_{k,n}x^k y^{n-k} = \sum_{\substack{k\in\mathbb{N},n\in\mathcal{E}:\\ k\leqslant n}} r_{k,n}x^k y^{n-k}.$$

We know that, for all even integer n, if k is odd, then $r_{k,n}=0$. We therefore have checked that

$$\tilde{G}(x,y) = \frac{1}{2}[G(x,y)+G(-x,-y)].$$

Thus,

$$\tilde{G}(x,y) = \frac{1}{\mathbf{A}(x)+\mathbf{A}(y)}\left[\frac{2p-1+\mathbf{A}(x)}{1-x^2}+\frac{2q-1+\mathbf{A}(y)}{1-y^2}\right].$$

We have now

$$2p-1+\mathbf{A}(x) = \frac{(2p-1)^2-\mathbf{A}(x)^2}{2p-1-\mathbf{A}(x)} = 4pq\,\frac{1-x^2}{1-2p+\mathbf{A}(x)}$$

and similarly,

$$2q-1+\mathbf{A}(y) = 4pq\,\frac{1-y^2}{1-2q+\mathbf{A}(y)}\cdot$$

This gives

$$\tilde{G}(x,y) = \frac{4pq}{\mathbf{A}(x)+\mathbf{A}(y)}\left[\frac{1}{1-2p+\mathbf{A}(x)}+\frac{1}{1-2q+\mathbf{A}(y)}\right]$$

which in turn, with $[1-2p+\mathbf{A}(x)]+[1-2q+\mathbf{A}(y)] = \mathbf{A}(x)+\mathbf{A}(y)$, yields (5.2). This formula supplies in particular

$$\tilde{G}(x,0) = \frac{2q}{1-2p+\mathbf{A}(x)} \quad\text{and}\quad \tilde{G}(0,y) = \frac{2p}{1-2q+\mathbf{A}(y)}$$

and we immediately obtain (5.3). □

5.2. **Distribution of the sojourn time**

In this part, we invert the generating function G given by (5.1) in order to derive the coefficients $r_{k,n}$.

Theorem 5.3. *The probability* $r_{k,n} = \mathbb{P}\{T_n = k\}$ *admits the following expression: for* $0 \leqslant k \leqslant n$,

$$\begin{aligned} r_{k,n} = \mathbb{1}_{\mathcal{E}}(n-k)\Bigg[2p\sum_{\substack{i\in\mathcal{E}:\\ n-k\leqslant i\leqslant n}} a_i(pq)^{i/2} - 4\sum_{\substack{i\in\mathcal{E}:\\ n-k\leqslant i\leqslant n-2}}\Bigg(\sum_{\substack{j\in\mathcal{E}:\\ j\leqslant i+k-n}} a_j a_{i-j}\Bigg)(pq)^{i/2+1}\Bigg] \\ + \mathbb{1}_{\mathcal{E}}(k)\Bigg[2q\sum_{\substack{i\in\mathcal{E}:\\ k\leqslant i\leqslant n}} a_i(pq)^{i/2} - 4\sum_{\substack{i\in\mathcal{E}:\\ k\leqslant i\leqslant n-2}}\Bigg(\sum_{\substack{j\in\mathcal{E}:\\ j\leqslant i-k}} a_j a_{i-j}\Bigg)(pq)^{i/2+1}\Bigg] \end{aligned} \tag{5.4}$$

where $a_i = \dfrac{1}{i+2}\dbinom{i}{i/2}$ *for* $i \in \mathcal{E}$.

More specifically:

- *For odd n*

$$r_{k,n}=\begin{cases} 2p\displaystyle\sum_{\substack{i\in\mathcal{E}:\\ n-k\leqslant i\leqslant n-1}} a_i(pq)^{i/2}-4\sum_{\substack{i\in\mathcal{E}:\\ n-k\leqslant i\leqslant n-3}}\Bigg(\sum_{\substack{j\in\mathcal{E}:\\ j\leqslant i+k-n}} a_j a_{i-j}\Bigg)(pq)^{i/2+1} & \textit{if } k \textit{ is odd,}\\ 2q\displaystyle\sum_{\substack{i\in\mathcal{E}:\\ k\leqslant i\leqslant n-1}} a_i(pq)^{i/2}-4\sum_{\substack{i\in\mathcal{E}:\\ k\leqslant i\leqslant n-3}}\Bigg(\sum_{\substack{j\in\mathcal{E}:\\ j\leqslant i-k}} a_j a_{i-j}\Bigg)(pq)^{i/2+1} & \textit{if } k \textit{ is even;}\end{cases} \tag{5.5}$$

- *For even n*

$$r_{k,n}=\begin{cases} 2p\displaystyle\sum_{\substack{i\in\mathcal{E}:\\ n-k\leqslant i\leqslant n}} a_i(pq)^{i/2}-4\sum_{\substack{i\in\mathcal{E}:\\ n-k\leqslant i\leqslant n-2}}\Bigg(\sum_{\substack{j\in\mathcal{E}:\\ j\leqslant i+k-n}} a_j a_{i-j}\Bigg)(pq)^{i/2+1} & \\ +2q\displaystyle\sum_{\substack{i\in\mathcal{E}:\\ k\leqslant i\leqslant n}} a_i(pq)^{i/2}-4\sum_{\substack{i\in\mathcal{E}:\\ k\leqslant i\leqslant n-2}}\Bigg(\sum_{\substack{j\in\mathcal{E}:\\ j\leqslant i-k}} a_j a_{i-j}\Bigg)(pq)^{i/2+1} & \textit{if } k \textit{ is even,}\\ 0 \textit{ if } k \textit{ is odd.} & \end{cases} \tag{5.6}$$

Proof. Rewrite (5.1) as

$$\begin{aligned} G(x,y)=&\frac{2p}{1-x}\,\frac{1}{\mathbf{A}(x)+\mathbf{A}(y)}-\frac{1-\mathbf{A}(x)}{1-x}\,\frac{1}{\mathbf{A}(x)+\mathbf{A}(y)}\\ &+\frac{2q}{1-y}\,\frac{1}{\mathbf{A}(x)+\mathbf{A}(y)}-\frac{1-\mathbf{A}(y)}{1-y}\,\frac{1}{\mathbf{A}(x)+\mathbf{A}(y)}. \end{aligned} \tag{5.7}$$

We expand the first term of (5.7). By (2.1) and (2.3),

$$\begin{aligned} \frac{2p}{1-x}\,\frac{1}{\mathbf{A}(x)+\mathbf{A}(y)}&=2p\Bigg(\sum_{m\in\mathbb{N}} x^m\Bigg)\Bigg(\sum_{j,l\in\mathcal{E}} a_{j+l}(pq)^{(j+l)/2}x^j y^l\Bigg)\\ &=2p\sum_{j,l\in\mathcal{E},m\in\mathbb{N}} a_{j+l}(pq)^{(j+l)/2}x^{j+m}y^l\\ &=2p\sum_{k,n\in\mathbb{N}}\Bigg[\sum_{\substack{j,l\in\mathcal{E},m\in\mathbb{N}:\\ j+m=k,l=n-k}} a_{j+l}(pq)^{(j+l)/2}\Bigg]x^k y^{n-k} \qquad (5.8)\\ &=2p\sum_{\substack{k,n\in\mathbb{N}:\\ k\leqslant n,n-k\in\mathcal{E}}}\Bigg[\sum_{\substack{m\in\mathbb{N}:\\ m\leqslant k,n-m\in\mathcal{E}}} a_{n-m}(pq)^{(n-m)/2}\Bigg]x^k y^{n-k} \qquad (5.9)\\ &=2p\sum_{\substack{k,n\in\mathbb{N}:\\ k\leqslant n,}} 1\!\!1_{\mathcal{E}}(n-k)\Bigg[\sum_{\substack{i\in\mathcal{E}:\\ n-k\leqslant i\leqslant n}} a_i(pq)^{i/2}\Bigg]x^k y^{n-k}. \qquad (5.10) \end{aligned}$$

Let us explain certain transformations made in the above calculations:

(1) In (5.8), we have introduced the indices $k=j+m$ and $n=j+l+m$. Then $j=k-m$ and $l=n-k$, the conditions $j,l\in\mathcal{E}$ are equivalent to $k-m\geqslant 0, n-k\geqslant 0, k-m\in\mathcal{E}$, $n-k\in\mathcal{E}$, or $m\leqslant k\leqslant n, n-k\in\mathcal{E}, n-m\in\mathcal{E}$. This leads to (5.9);

(2) From (5.9) to (5.10), we have put $i=n-m$.

We expand the second term of (5.7). By (2.1) and (2.3),

$$\begin{aligned}
\frac{1-\mathbf{A}(x)}{1-x}\,\frac{1}{\mathbf{A}(x)+\mathbf{A}(y)} &= 4\Big(\sum_{m\in\mathbb{N}} x^m\Big)\Big(\sum_{i\in\mathcal{E}^*} a_{i-2}(pq)^{i/2}x^i\Big)\Big(\sum_{j,l\in\mathcal{E}} a_{j+l}(pq)^{(j+l)/2}x^jy^l\Big)\\
&= 4\sum_{i\in\mathcal{E}^*,j,l\in\mathcal{E},m\in\mathbb{N}} a_{i-2}a_{j+l}(pq)^{(i+j+l)/2}x^{i+j+m}y^l\\
&= 4\sum_{k,n\in\mathbb{N}}\Bigg[\sum_{\substack{i\in\mathcal{E}^*,j,l\in\mathcal{E},m\in\mathbb{N}:\\ i+j+m=k,l=n-k}} a_{i-2}a_{j+l}(pq)^{(i+j+l)/2}\Bigg]x^ky^{n-k} \qquad (5.11)\\
&= 4\sum_{\substack{k,n\in\mathbb{N}:\\ k\leqslant n,n-k\in\mathcal{E}}}\Bigg[\sum_{\substack{i\in\mathcal{E}^*,m\in\mathbb{N}:\\ i+m\leqslant k,i+n-m\in\mathcal{E}}} a_{i-2}a_{n-m-i}(pq)^{(n-m)/2}\Bigg]x^ky^{n-k} \qquad (5.12)\\
&= 4\sum_{\substack{k,n\in\mathbb{N}:\\ k\leqslant n,n-k\in\mathcal{E}}}\Bigg[\sum_{\substack{m\in\mathbb{N}:\\ n-m\in\mathcal{E}}}\Bigg(\sum_{\substack{i\in\mathcal{E}^*:\\ i\leqslant k-m}} a_{i-2}a_{n-m-i}\Bigg)(pq)^{(n-m)/2}\Bigg]x^ky^{n-k} \qquad (5.13)\\
&= 4\sum_{\substack{k,n\in\mathbb{N}:\\ k\leqslant n,n-k\in\mathcal{E}}}\Bigg[\sum_{\substack{m\in\mathbb{N}:\\ n-m\in\mathcal{E}}}\Bigg(\sum_{\substack{j\in\mathcal{E}:\\ j\leqslant k-m-2}} a_ja_{n-m-j-2}\Bigg)(pq)^{(n-m)/2}\Bigg]x^ky^{n-k} \qquad (5.14)\\
&= 4\sum_{\substack{k,n\in\mathbb{N}:\\ k\leqslant n,n-k\in\mathcal{E}}}\Bigg[\sum_{\substack{m\in\mathbb{N}\setminus\{0,1\}:\\ n-m\in\mathcal{E}}}\Bigg(\sum_{\substack{j\in\mathcal{E}:\\ j\leqslant k-m}} a_ja_{n-m-j}\Bigg)(pq)^{(n-m)/2+1}\Bigg]x^ky^{n-k} \qquad (5.15)\\
&= 4\sum_{\substack{k,n\in\mathbb{N}:\\ k\leqslant n}}1\!\!1_{\mathcal{E}}(n-k)\Bigg[\sum_{\substack{i\in\mathcal{E}:\\ i\leqslant n-2}}\Bigg(\sum_{\substack{j\in\mathcal{E}:\\ j\leqslant i+k-n}} a_ja_{i-j}\Bigg)(pq)^{i/2+1}\Bigg]x^ky^{n-k}. \qquad (5.16)
\end{aligned}$$

Let us explain the transformations used in the above computations:

(1) In (5.11), we have introduced the indices $k=i+j+m$ and $n=i+j+l+m$. Then $j=-i+k-m, l=n-k$ and $i\in\mathcal{E}$, the conditions $j,l\in\mathcal{E}$ are equivalent to $k-i-m\geqslant 0, n-k\geqslant 0, k-i-m\in\mathcal{E}, n-k\in\mathcal{E}$, or $i+m\leqslant k\leqslant n, n-k\in\mathcal{E}, n-m\in\mathcal{E}$. This leads to (5.12);

(2) From (5.13) to (5.14), we have put $i=j+2$;

(3) From (5.14) to (5.15), we have changed m into $m-2$;

(4) From (5.15) to (5.16), we have put $i=n-m$.

By invoking the argument of duality which is explained in Remark 4.3, the two last terms of (5.7) can be deduced from the two first ones by interchanging p and q, and x and y (that is k and $n-k$). This yields

$$\frac{2q}{1-y}\,\frac{1}{\mathbf{A}(x)+\mathbf{A}(y)} = 2q\sum_{\substack{k,n\in\mathbb{N}:\\ k\leqslant n}}1\!\!1_{\mathcal{E}}(k)\Bigg[\sum_{\substack{i\in\mathcal{E}:\\ k\leqslant i\leqslant n}} a_i(pq)^{i/2}\Bigg]x^ky^{n-k}, \qquad (5.17)$$

$$\frac{1-\mathbf{A}(y)}{1-y}\,\frac{1}{\mathbf{A}(x)+\mathbf{A}(y)} = 4\sum_{\substack{k,n\in\mathbb{N}:\\ k\leqslant n}}1\!\!1_{\mathcal{E}}(k)\Bigg[\sum_{\substack{i\in\mathcal{E}:\\ i\leqslant n-2}}\Bigg(\sum_{\substack{j\in\mathcal{E}:\\ j\leqslant i-k}} a_ja_{i-j}\Bigg)(pq)^{i/2+1}\Bigg]x^ky^{n-k}. \qquad (5.18)$$

By identifying the coefficients of the series lying in the definition of G and in (5.10), (5.16), (5.17) and (5.18), we extract (5.4). □

Remark 5.4. The probabilities

$$\begin{aligned}
r_{0,n} &= \mathbb{P}\{T_n=0\} = \mathbb{P}\{S_1\leqslant 0,\dots,S_n\leqslant 0\} = \mathbb{P}\{\tau_1>n\},\\
r_{n,n} &= \mathbb{P}\{T_n=n\} = \mathbb{P}\{S_1\geqslant 0,\dots,S_n\geqslant 0\} = \mathbb{P}\{\tau_{-1}>n\}
\end{aligned}$$

are given by

$$r_{0,n}=1-2p\sum_{\substack{i\in\mathcal{E}:\\ i\leqslant n-1}}a_i(pq)^{i/2}\quad\text{and}\quad r_{n,n}=1-2q\sum_{\substack{i\in\mathcal{E}:\\ i\leqslant n-1}}a_i(pq)^{i/2}.\tag{5.19}$$

The probabilities

$$r_{1,n}=\mathbb{P}\{T_n=1\}=\mathbb{P}\{S_1\leqslant 0,\ldots,S_{n-1}\leqslant 0,S_n>0\}=\mathbb{P}\{\tau_1=n\},$$
$$r_{n-1,n}=\mathbb{P}\{T_n=n-1\}=\mathbb{P}\{S_1\geqslant 0,\ldots,S_{n-1}\geqslant 0,S_n<0\}=\mathbb{P}\{\tau_{-1}=n\}$$

are given by $r_{1,n}=r_{n-1,n}=0$ if n is even and, if n is odd,

$$r_{1,n}=2a_{n-1}\,p^{(n+1)/2}q^{(n-1)/2}\quad\text{and}\quad r_{n-1,n}=2a_{n-1}\,p^{(n-1)/2}q^{(n+1)/2}.\tag{5.20}$$

Formulas (5.20) can be easily deduced from (5.5) and (5.6). Formulas (5.19) could be deduced from Theorem 5.3. The computations are postponed to the report [7].

Proposition 5.5. *For even n,k such that $0\leqslant k\leqslant n$, the following relationship holds:*

$$r_{k,n}=r_{k,k}r_{0,n-k}.\tag{5.21}$$

Proof. The probabilities $r_{k,n}$, $k,n\in\mathcal{E}$, are generated by the function $\tilde{G}$ which was introduced in Section 5.2. By (5.3), the quantity $\tilde{G}(x,y)$ can be factorized into the product of $\tilde{G}(x,0)$ and $\tilde{G}(0,y)$. Writing that

$$\tilde{G}(x,0)\tilde{G}(0,y)=\sum_{k,l\in\mathcal{E}}(r_{k,k}r_{0,l})x^ky^l=\sum_{\substack{k,n\in\mathcal{E}:\\ k\leqslant n}}(r_{k,k}r_{0,n-k})x^ky^{n-k},$$

we conclude, by identification, that (5.21) holds true. Formula (5.21) could be obtained by using the explicit representation (5.19) of the probabilities $r_{k,k}$ and $r_{0,n-k}$ and the representation (5.6) of $r_{k,n}$. Nevertheless, the computations are very cumbersome. They are included in the report [7]. In [7], the reader can find also a proof by induction. □

Corollary 5.6. *In the symmetric case ($p=q=1/2$), the well-known following expression holds for even integers n,k such that $0\leqslant k\leqslant n$:*

$$r_{k,n}=\frac{1}{2^n}\binom{k}{k/2}\binom{(n-k)}{(n-k)/2}.$$

Proof. In the case where $p=1/2$, we have the particular identities

$$\mathbb{P}\{\tau_0=j\}=\mathbb{P}\{S_{j-2}=0\}-\mathbb{P}\{S_j=0\}\quad\text{and}\quad\mathbb{P}\{\tau_{-1}=j\}=\mathbb{P}\{\tau_0=j+1\}$$

which are due to the fact that $b_{j-2}-b_j/4=a_{j-2}$ and to (2.10) respectively, and then

$$\begin{aligned}r_{k,k}&=\mathbb{P}\{\tau_{-1}\geqslant k+1\}=\sum_{\substack{j\in\mathcal{O}:\\ j\geqslant k+1}}\mathbb{P}\{\tau_{-1}=j\}=\sum_{\substack{j\in\mathcal{O}:\\ j\geqslant k+1}}\mathbb{P}\{\tau_0=j+1\}=\sum_{\substack{j\in\mathcal{E}:\\ j\geqslant k+2}}\mathbb{P}\{\tau_0=j\}\\&=\sum_{\substack{j\in\mathcal{E}:\\ j\geqslant k+2}}[\mathbb{P}\{S_{j-2}=0\}-\mathbb{P}\{S_j=0\}]=\mathbb{P}\{S_k=0\}=\frac{1}{2^k}\binom{k}{k/2}.\end{aligned}$$

Analogously,

$$r_{0,n-k}=\mathbb{P}\{\tau_1\geqslant n-k+1\}=\mathbb{P}\{S_{n-k}=0\}=\frac{1}{2^{n-k}}\binom{(n-k)}{(n-k)/2}.$$

We conclude with the help of (5.21). □

Proposition 5.7. *For odd integers n, k such that $0 \leqslant k \leqslant n-1$, the following relationship holds:*

$$r_{k,n} + r_{k+1,n} = r_{k+1,n+1}. \tag{5.22}$$

Proof. Pick two odd integers n, k such that $0 \leqslant k \leqslant n-1$. We have

$$r_{k,n} + r_{k+1,n} = \mathbb{P}\{T_n = k\} + \mathbb{P}\{T_n = k+1\} = \mathbb{P}\{T_n \in \{k, k+1\}\}. \tag{5.23}$$

Let us introduce the last passage time by 0, say σ_0, for the random walk. Is is plain that σ_0 is even and that there is an even number of δ_j up to time σ_0 which are equal to one. Afterwards, it remains an odd number (this is $n - \sigma_0$) of steps up to time n. One has either $S_{\sigma_0+1} > 0, S_{\sigma_0+2} > 0, \ldots, S_n > 0$ or $S_{\sigma_0+1} < 0, S_{\sigma_0+2} < 0, \ldots, S_n < 0$. The corresponding δ_j, $\sigma_0 + 1 \leqslant j \leqslant n$, are either all equal to one or all equal to zero. More precisely:

- If $T_n = k$, since k is odd, one has in this case $S_{\sigma_0+1} > 0, S_{\sigma_0+2} > 0, \ldots, S_n > 0$. Then, necessarily, $S_{n+1} \geqslant 0$ which entails that $\delta_{n+1} = 1$ and $T_{n+1} = T_n + 1 = k+1$;
- If $T_n = k+1$, since $k+1$ is even, one has in this case $S_{\sigma_0+1} < 0, S_{\sigma_0+2} < 0, \ldots, S_n < 0$. Then, necessarily, $S_{n+1} \leqslant 0$ which entails that $\delta_{n+1} = 0$ and $T_{n+1} = T_n = k+1$.

This discussion implies the inclusion $\{T_n \in \{k, k+1\}\} \subset \{T_{n+1} = k+1\}$. Conversely, since $T_{n+1} - T_n = \delta_n \in \{0,1\}$, the equality $T_{n+1} = k+1$ implies $T_n \in \{k, k+1\}$ which proves the inclusion $\{T_{n+1} = k+1\} \subset \{T_n \in \{k, k+1\}\}$. As a byproduct, the equality $\{T_n \in \{k, k+1\}\} = \{T_{n+1} = k+1\}$ holds and this proves, referring to (5.23), the relationship (5.22). In [7], we propose another proof of (5.22) which uses the explicit results obtained in Theorem 5.3. $\square$

6. The case $F = \{0\}$ (bridge of the random walk)

In this part, we shall set for simplifying $r_{k,n}^{\{0\}} = r_{k,n}^0$ and $G^{\{0\}}(x,y) = G^0(x,y)$. We consider the distribution of the sojourn time T_n subject to the condition that $S_n = 0$, that is we are dealing with the so-called bridge of the random walk pinned at zero at times 0 and n. The condition $S_n = 0$ can be fulfilled only when n is even. So, we make here the assumption that n is an even integer throughout this section.

6.1. Generating function

In the Proof of Theorem 3.1, we have derived the remarkably simple result below.

Theorem 6.1. *The generating function G^0 is given by*

$$G^0(x,y) = \frac{2}{\mathbf{A}(x) + \mathbf{A}(y)} \tag{6.1}$$

where $\mathbf{A}(z) = \sqrt{1 - 4pqz^2}$.

6.2. Distribution of the sojourn time

Theorem 6.2. *Assume that n is even. The probability $r_{k,n}^0 = \mathbb{P}\{T_n = k, S_n = 0\}$ admits the following expression:*

$$r_{k,n}^0 = \begin{cases} \dfrac{2}{n+2}\dbinom{n}{n/2}(pq)^{n/2} & \text{if } k \text{ is even such that } 0 \leqslant k \leqslant n, \\ 0 & \text{if } k \text{ is odd or such that } k \geqslant n+1. \end{cases} \tag{6.2}$$

The distribution of T_n for the bridge of the random walk is given by

$$\mathbb{P}\{T_n = k \,|\, S_n = 0\} = \begin{cases} \dfrac{2}{n+2} & \text{if } k \text{ is even such that } 0 \leqslant k \leqslant n, \\ 0 & \text{if } k \text{ is odd or such that } k \geqslant n+1, \end{cases}$$

that is, the conditioned random variable $(T_n = k \,|\, S_n = 0)$ is uniformly distributed on the set $\{0, 2, 4, \ldots, n-2, n\}$.

Proof. By (2.3), we rewrite $G^0(x,y)$ as

$$G^0(x,y) = 2\sum_{i,j\in\mathcal{E}} a_{i+j}(pq)^{(i+j)/2}x^iy^j = 2\sum_{\substack{k,n\in\mathcal{E}:\\ k\leqslant n}} a_n(pq)^{n/2}x^ky^{n-k}.$$

From this, we immediately extract the announced expression for $r^0_{k,n}$. Moreover, we plainly have

$$\mathbb{P}\{T_n = k \,|\, S_n = 0\} = \frac{\mathbb{P}\{T_n = k, S_n = 0\}}{\mathbb{P}\{S_n = 0\}}$$

with

$$\mathbb{P}\{S_n = 0\} = \binom{n}{n/2}(pq)^{n/2}$$

and this prove the assertion related to the uniform law. $\square$

7. The cases $F = \mathbb{Z}^{+*}$ and $F = \mathbb{Z}^{-*}$

In this part, we shall set for simplifying $r^{\mathbb{Z}^{+*}}_{k,n} = r^+_{k,n}$, $G^{\mathbb{Z}^{+*}}(x,y) = G^+(x,y)$ and $H^{\mathbb{Z}^{+*}}(z) = H^+(z)$. We shall also use similar notations with minus signs for the study of the case $F = \mathbb{Z}^{-*}$. Some expressions for the $r^+_{k,n}$'s can be obtained from Theorem 4.2 by summing the $r^j_{k,n}$, $j \in \mathbb{Z}^{+*}$ or $j \in \mathbb{Z}^{-*}$. Since $r^j_{k,n} = 0$ for $|j| > n$, the corresponding sums reduce to

$$r^+_{k,n} = \sum_{j=1}^{n} r^j_{k,n} \quad\text{and}\quad r^-_{k,n} = \sum_{j=-n}^{-1} r^j_{k,n}.$$

We propose other expressions which can be deduced from the generating functions G^+ and G^-.

7.1. Generating function

We first consider the case where $F = \mathbb{Z}^{+*}$. We need to evaluate the functions $H^+_i(z)$ for $i \in \{-1,0,1\}$, $G^+(x,0)$ and $G^+(0,y)$. On one hand, we have for $i \in \{-1,0,1\}$ (and then $i \geqslant 0$)

$$\begin{aligned} H^+_i(z) &= \sum_{j\in\mathbb{Z}^{+*}-i} H_j(z) = \frac{1}{\mathbf{A}(z)}\sum_{j=1-i}^{\infty}\left(\frac{1-\mathbf{A}(z)}{2qz}\right)^j \\ &= \frac{1}{\mathbf{A}(z)(1-\frac{1-\mathbf{A}(z)}{2qz})}\left(\frac{1-\mathbf{A}(z)}{2qz}\right)^{1-i} = \frac{1-\mathbf{A}(z)}{\mathbf{A}(z)[2qz-1+\mathbf{A}(z)]}\left(\frac{1-\mathbf{A}(z)}{2qz}\right)^{-i}. \end{aligned}$$

Invoking (2.7) and observing that

$$[2pz-1+\mathbf{A}(z)][2qz-1+\mathbf{A}(z)] = 4pqz^2-2z+1+2(z-1)\mathbf{A}(z)+\mathbf{A}(z)^2 = 2(1-z)[1-\mathbf{A}(z)],$$

we find that

$$H^+_i(z) = \frac{2pz-1+\mathbf{A}(z)}{2(1-z)\mathbf{A}(z)}\left(\frac{1-\mathbf{A}(z)}{2qz}\right)^{-i} = \frac{2pz-1+\mathbf{A}(z)}{2(1-z)\mathbf{A}(z)}\left(\frac{1+\mathbf{A}(z)}{2pz}\right)^{i}.$$

On the other hand, by (3.2),

$$G^+(x,0) = H_0^+(x) - \frac{1-\mathbf{A}(x)}{2px}\, H_{-1}^+(x) = H_0^+(x)\left[1 - \frac{1-\mathbf{A}(x)}{2px}\,\frac{2px}{1+\mathbf{A}(x)}\right]$$
$$= \frac{2\mathbf{A}(x)}{1+\mathbf{A}(x)}\, H_0^+(x) = \frac{2px-1+\mathbf{A}(x)}{(1-x)[1+\mathbf{A}(x)]}\cdot$$

Similarly,

$$G^+(0,y) = H_0^+(y) - \frac{1-\mathbf{A}(y)}{2qy}\, H_1^+(y) = H_0^+(y)\left[1 - \frac{1-\mathbf{A}(y)}{2qy}\,\frac{2qy}{1-\mathbf{A}(y)}\right] = 0.$$

With this at hand, we can derive $G^+(x,y)$. Indeed, using the general formula (3.1), we obtain

$$G^+(x,y) = \frac{1+\mathbf{A}(x)}{\mathbf{A}(x)+\mathbf{A}(y)}\, G^+(x,0) = \frac{2px-1+\mathbf{A}(x)}{(1-x)[\mathbf{A}(x)+\mathbf{A}(y)]}\cdot$$

Exactly in the same way, we could find the result related to the case $F = \mathbb{Z}^{-*}$. We state both results in the theorem below.

Theorem 7.1. *The generating functions G^+ and G^- are given by*

$$G^+(x,y) = \frac{2px-1+\mathbf{A}(x)}{(1-x)[\mathbf{A}(x)+\mathbf{A}(y)]} \quad \textit{and} \quad G^-(x,y) = \frac{2qy-1+\mathbf{A}(y)}{(1-y)[\mathbf{A}(x)+\mathbf{A}(y)]} \tag{7.1}$$

where $\mathbf{A}(z) = \sqrt{1-4pqz^2}$.

Remark 7.2. We have the following relationship between the generating functions G, G^+, G^-, G^0:

$$G^+(x,y) + G^-(x,y) + G^0(x,y) = G(x,y)$$

which can be directly checked by using the expressions (5.1), (6.1) and (7.1). In fact, it is due to the fact that $\{S_n \in \mathbb{Z}^{+*}\} \cup \{S_n \in \mathbb{Z}^{-*}\} \cup \{S_n = 0\} = \Omega$ which implies that

$$r_{k,n}^+ + r_{k,n}^- + r_{k,n}^0 = r_{k,n}.$$

7.2. Distribution of the sojourn time

From Theorem 7.1, we derive the coefficients $r_{k,n}^+$ and $r_{k,n}^-$.

Theorem 7.3. *The probabilities $r_{k,n}^+ = \mathbb{P}\{T_n = k, S_n > 0\}$ and $r_{k,n}^- = \mathbb{P}\{T_n = k, S_n < 0\}$ admit the following expressions: for $0 \leqslant k \leqslant n$,*

$$r_{k,n}^+ = \mathbb{1}_{\mathcal{E}}(n-k)\left[2p \sum_{\substack{i\in\mathcal{E}:\\ n-k\leqslant i\leqslant n-1}} a_i (pq)^{i/2} - 4 \sum_{\substack{i\in\mathcal{E}:\\ n-k\leqslant i\leqslant n-2}} \left(\sum_{\substack{j\in\mathcal{E}:\\ j\leqslant i+k-n}} a_j a_{i-j}\right)(pq)^{i/2+1}\right], \tag{7.2}$$

$$r_{k,n}^- = \mathbb{1}_{\mathcal{E}}(k)\left[2q \sum_{\substack{i\in\mathcal{E}:\\ k\leqslant i\leqslant n-1}} a_i (pq)^{i/2} - 4 \sum_{\substack{i\in\mathcal{E}:\\ k\leqslant i\leqslant n-2}} \left(\sum_{\substack{j\in\mathcal{E}:\\ j\leqslant i-k}} a_j a_{i-j}\right)(pq)^{i/2+1}\right], \tag{7.3}$$

where $a_i = \dfrac{1}{i+2}\dbinom{i}{i/2}$ *for* $i \in \mathcal{E}$.

Proof. We invert the generating function G^+. For its expansion, we refer to the Proof of Theorem 5.3. We have

$$\begin{aligned}G^+(x,y) &= \frac{2px}{1-x}\,\frac{1}{\mathbf{A}(x)+\mathbf{A}(y)} - \frac{1-\mathbf{A}(x)}{1-x}\,\frac{1}{\mathbf{A}(x)+\mathbf{A}(y)}\\ &= 2p\sum_{\substack{k,n\in\mathbb{N}:\\ k\leqslant n,}} \mathbb{1}_{\mathcal{E}}(n-k)\Bigg[\sum_{\substack{i\in\mathcal{E}:\\ n-k\leqslant i\leqslant n}} a_i(pq)^{i/2}\Bigg]x^{k+1}y^{n-k}\\ &\quad -4\sum_{\substack{k,n\in\mathbb{N}:\\ k\leqslant n}} \mathbb{1}_{\mathcal{E}}(n-k)\Bigg[\sum_{\substack{i\in\mathcal{E}:\\ i\leqslant n-2}}\Bigg(\sum_{\substack{j\in\mathcal{E}:\\ j\leqslant i+k-n}} a_ja_{i-j}\Bigg)(pq)^{i/2+1}\Bigg]x^ky^{n-k}.\end{aligned}$$

Performing the substitution $(k,n)\mapsto(k-1,n-1)$ in the first term of the last above equality, we get

$$\begin{aligned}G^+(x,y) &= 2p\sum_{\substack{k,n\in\mathbb{N}:\\ k\leqslant n,}} \mathbb{1}_{\mathcal{E}}(n-k)\Bigg[\sum_{\substack{i\in\mathcal{E}:\\ n-k\leqslant i\leqslant n-1}} a_i(pq)^{i/2}\Bigg]x^{k}y^{n-k}\\ &\quad -4\sum_{\substack{k,n\in\mathbb{N}:\\ k\leqslant n}} \mathbb{1}_{\mathcal{E}}(n-k)\Bigg[\sum_{\substack{i\in\mathcal{E}:\\ i\leqslant n-2}}\Bigg(\sum_{\substack{j\in\mathcal{E}:\\ j\leqslant i+k-n}} a_ja_{i-j}\Bigg)(pq)^{i/2+1}\Bigg]x^ky^{n-k}.\end{aligned}$$

Formula (7.2) ensues by identification. Formula (7.3) can be deduced from (7.2) by invoking the duality argument mentioned in Remark 4.3: it suffices to interchange p and q on one hand, and k and $n-k$ on the other hand. □

Remark 7.4. By comparing (5.4) and (7.2), we can see that, for odd integer n, $r^+_{k,n}=r_{k,n}$ if k is odd, $r^+_{k,n}=0$ if k is even, and, for even integer n, $r^+_{0,n}=r^+_{1,n}=0$ and $r^+_{2,n}=p^2r^0_{0,n-2}$. These relations can be directly checked. For instance, in the last case, we can easily observe that the conditions $T_n=2$ and $S_n>0$ are fulfilled only in the case where $S_0\leqslant 0, S_1\leqslant 0,\dots,S_{n-3}\leqslant 0$, $S_{n-2}=0, S_{n-1}=1$ and $S_n=2$. Thus,

$$\begin{aligned}\mathbb{P}\{T_n=2,S_n>0\} &= \mathbb{P}\{S_0\leqslant 0,S_1\leqslant 0,\dots,S_{n-3}\leqslant 0,S_{n-2}=0,X_{n-1}=X_n=1\}\\ &= p^2\mathbb{P}\{T_{n-2}=0,S_{n-2}=0\}\end{aligned}$$

which is nothing but $r^+_{2,n}=p^2r^0_{0,n-2}$.

Remark 7.5. In Remark 7.2, we mentioned the relationship $r^+_{k,n}+r^-_{k,n}+r^0_{k,n}=r_{k,n}$. This one can be checked by adding (6.2), (7.2) and (7.3) after noticing that

$$r^0_{k,n}=2\mathbb{1}_{\mathcal{E}}(k)\mathbb{1}_{\mathcal{E}}(n)a_n(pq)^{n/2}=\mathbb{1}_{\mathcal{E}}(n-k)\mathbb{1}_{\mathcal{E}}(n)2p\,a_n(pq)^{n/2}+\mathbb{1}_{\mathcal{E}}(k)\mathbb{1}_{\mathcal{E}}(n)2q\,a_n(pq)^{n/2},$$

the foregoing sum coincides with (5.4).

8. Asymptotics: Brownian motion with a linear drift

In this part, our aim is to retrieve certain probability distributions related to the sojourn time in $(0,+\infty)$ of Brownian motion with a linear drift.

8.1. Rescaled random walk

We consider a sequence of random walks $(S^N_k)_{k\in\mathbb{N}}$ indexed by $N\in\mathbb{N}$, defined by

$$S^N_k=S^N_0+\sum_{j=1}^k X^N_j,\ k\geqslant 1,$$

where for each $N \in \mathbb{N}^*$, $(X_j^N)_{j\in\mathbb{N}^*}$ is a sequence of independent Bernoulli variables with jump probabilities depending on N as follows:

$$p_N = \mathbb{P}\{X_k^N = +1\} = \frac{1}{2} + \frac{\rho}{2\sqrt{N}} \quad \text{and} \quad q_N = \mathbb{P}\{X_k^N = -1\} = \frac{1}{2} - \frac{\rho}{2\sqrt{N}},$$

ρ being a fixed parameter. We also define the centered random walk $(\tilde{S}_k^N)_{k\in\mathbb{N}}$ as

$$\tilde{S}_k^N = S_k^N - \mathbb{E}(S_k^N) = S_k^N - \rho\,\frac{k}{\sqrt{N}}\cdot$$

Let T_n^N be the corresponding sojourn time in $\mathbb{Z}^+$: $T_n^N = \sum_{j=1}^n \delta_j^N$ with

$$\delta_j^N = \begin{cases} 1 & \text{if } (S_j^N > 0) \text{ or } (S_j^N = 0 \text{ and } S_{j-1}^N > 0), \\ 0 & \text{if } (S_j^N < 0) \text{ or } (S_j^N = 0 \text{ and } S_{j-1}^N < 0). \end{cases}$$

Let us introduce the rescaled random walks

$$B_t^N = \frac{1}{\sqrt{N}}\,S_{[Nt]}^N \quad \text{and} \quad \tilde{B}_t^N = B_t^N - \mathbb{E}(B_t^N) = \frac{1}{\sqrt{N}}\,\tilde{S}_{[Nt]}^N$$

defined on continuous time $t \geqslant 0$. We have

$$B_t^N = \tilde{B}_t^N + \rho\,\frac{[Nt]}{N},\ t \geqslant 0.$$

Donsker's theorem (see, *e.g.*, [1], p. 68) asserts that the sequence of processes $(\tilde{B}_t^N)_{t\geqslant 0}$, $N \in \mathbb{N}$, weakly converges to the standard linear Brownian motion $(\tilde{B}_t)_{t\geqslant 0}$ and the sequence of processes $(\tilde{B}_t^N)_{t\geqslant 0}$, $N \in \mathbb{N}$, weakly converges to the drifted Brownian motion $(B_t)_{t\geqslant 0}$ defined as

$$B_t = \tilde{B}_t + \rho t,\ t \geqslant 0.$$

We now introduce the sojourn times in $\mathbb{R}^+$ of the processes $(B_t^N)_{t\geqslant 0}$ and $(B_t)_{t\geqslant 0}$:

$$\mathbf{T}_t^N = \int_0^t 1\!\!1_{\mathbb{R}^+}(B_s^N)\,\mathrm{d}s \quad \text{and} \quad \mathbf{T}_t = \int_0^t 1\!\!1_{\mathbb{R}^+}(B_s)\,\mathrm{d}s.$$

We have

$$\begin{aligned} \mathbf{T}_t^N &= \sum_{j=0}^{[Nt]} \int_{j/N}^{(j+1)/N} 1\!\!1_{\mathbb{R}^+}(S_j^N)\,\mathrm{d}s - \int_t^{([Nt]+1)/N} 1\!\!1_{\mathbb{R}^+}(S_{[Nt]}^N)\,\mathrm{d}s \\ &= \frac{1}{N}\sum_{j=0}^{[Nt]} 1\!\!1_{\mathbb{R}^+}(S_j^N) - \Big(\frac{[Nt]+1}{N} - t\Big) 1\!\!1_{\mathbb{R}^+}(S_{[Nt]}^N) \\ &= \frac{1}{N}\,T_{[Nt]}^N + \frac{1}{N}\sum_{j=1}^{[Nt]} \Big(1\!\!1_{\mathbb{R}^+}(S_j^N) - \delta_j^N\Big) + \frac{1}{N}\,1\!\!1_{\mathbb{R}^+}(S_0^N) - \Big(\frac{[Nt]+1}{N} - t\Big) 1\!\!1_{\mathbb{R}^+}(S_{[Nt]}^N). \end{aligned} \tag{8.1}$$

We compare the sojourn time of the random walk $(S_k^N)_{k\in\mathbb{N}}$ and that of the Brownian motion $(B_t)_{t\geqslant 0}$:

$$\frac{1}{N}\,T_{[Nt]}^N - \mathbf{T}_t = \left(\frac{1}{N}\,T_{[Nt]}^N - \mathbf{T}_t^N\right) + \left(\mathbf{T}_t^N - \mathbf{T}_t\right).$$

On one hand, for $j \geqslant 1$,

$$\mathbb{1}_{\mathbb{R}^+}(S_j^N) - \delta_j^N = \begin{cases} 0 & \text{if } (S_j^N \neq 0) \text{ or } (S_j^N = 0 \text{ and } S_{j-1}^N > 0), \\ 1 & \text{if } S_j^N = 0 \text{ and } S_{j-1}^N < 0, \end{cases}$$

and then,

$$|\mathbb{1}_{\mathbb{R}^+}(S_j^N) - \delta_j^N| \leqslant \mathbb{1}_{\{0\}}(S_j).$$

On the other hand, by (8.1), the following estimate holds:

$$\left|\frac{1}{N}\,T_{[Nt]}^N - \mathbf{T}_t^N\right| \leqslant \frac{1}{N}\sum_{j=1}^{[Nt]} \mathbb{1}_{\{0\}}(S_j^N) + \frac{2}{N}\cdot$$

We observe, by (2.1), that

$$\mathbb{E}\left[\sum_{j=1}^{[Nt]} \mathbb{1}_{\{0\}}(S_j^N)\right] \leqslant \sum_{j=1}^{\infty} \mathbb{P}\{S_j^N = 0\} \leqslant \sum_{j\in\mathcal{E}} b_j(p_N q_N)^{j/2} = \frac{1}{A(p_N q_N)} = \frac{\sqrt{N}}{\rho} = o(N)$$

which implies that $\frac{1}{N}\,T_{[Nt]}^N - \mathbf{T}_t$ tends to 0 in mean. Now, by Donsker's theorem, the sequence $(\mathbf{T}_t^N)_{N\in\mathbb{N}^*}$ weakly converges to $\mathbf{T}_t$ (see [1], p. 72). This discussion shows that $\frac{1}{N}\,T_{[Nt]}^N - \mathbf{T}_t$ converges to 0 in mean. As a result, the probability distribution of $\frac{1}{N}\,T_{[Nt]}^N$ converges to that of $\mathbf{T}_t$. In the following subsection, we compute this limit.

8.2. **Limiting distribution of the sojourn time**

Theorem 8.1. *The probability distribution function of the sojourn time $\mathbf{T}_t$ admits the following expression: for $0 < s < t$,*

$$\mathbb{P}\{\mathbf{T}_t \in \mathrm{d}s\}/\mathrm{d}s = \begin{cases} \dfrac{1}{\sqrt{2\pi}}\left[\rho + \dfrac{1}{2\sqrt{2\pi}}\displaystyle\int_s^\infty \frac{\mathrm{e}^{-\rho^2 z/2}}{z^{3/2}}\,\mathrm{d}z\right]\displaystyle\int_{t-s}^\infty \frac{\mathrm{e}^{-\rho^2 z/2}}{z^{3/2}}\,\mathrm{d}z & \text{if } \rho \geqslant 0, \\ \dfrac{1}{\sqrt{2\pi}}\left[|\rho| + \dfrac{1}{2\sqrt{2\pi}}\displaystyle\int_{t-s}^\infty \frac{\mathrm{e}^{-\rho^2 z/2}}{z^{3/2}}\,\mathrm{d}z\right]\displaystyle\int_s^\infty \frac{\mathrm{e}^{-\rho^2 z/2}}{z^{3/2}}\,\mathrm{d}z & \text{if } \rho \leqslant 0. \end{cases}$$

The integral $\int_\sigma^\infty \frac{\mathrm{e}^{-\rho^2 z/2}}{z^{3/2}}\,\mathrm{d}z$ can be expressed by means of the error function according as

$$\int_\sigma^\infty \frac{\mathrm{e}^{-\rho^2 z/2}}{z^{3/2}}\,\mathrm{d}z = \frac{2}{\sqrt{\sigma}}\,\mathrm{e}^{-\rho^2\sigma/2} - \sqrt{2\pi}\,\rho\,\mathrm{Erfc}(\rho\sqrt{\sigma/2}).$$

Let us point out that the density of $\mathbf{T}_t$ is known under another form, see formula 2.1.4.8 in [2].

Proof. We shall assume that $\rho \geqslant 0$, the case $\rho \leqslant 0$ being quite similar. We begin by computing the limit of the following probability as $N \to \infty$:

$$\mathbb{P}\{s_1 < \frac{1}{N}\,T_{[Nt]}^N \leqslant s_2\} = \mathbb{P}\{[Ns_1] < T_{[Nt]}^N \leqslant [Ns_2]\} = \sum_{k=[Ns_1]+1}^{[Ns_2]} \mathbb{P}\Big\{T_{[Nt]}^N = k\Big\}.$$

• *Case where $[Nt]$ is even.* Suppose that $[Nt]$ is even and set $\alpha_i^N = a_i(p_N q_N)^{i/2}$ with $p_N q_N = \frac{1}{4}(1-\rho^2/N)$. By using (5.19) and (5.21), for even integer k such that $[Ns_1] < k \leqslant [Ns_2]$, we get

$$\mathbb{P}\{T_{[Nt]}^N = k\} = \mathbb{P}\{T_k^N = k\}\mathbb{P}\{T_{[Nt]-k}^N = 0\} = \Bigg[1 - q_N \sum_{\substack{i\in\mathcal{E}:\\ 0\leqslant i\leqslant k-1}} \alpha_i^N\Bigg]\Bigg[1 - p_N \sum_{\substack{i\in\mathcal{E}:\\ 0\leqslant i\leqslant [Nt]-k-1}} \alpha_i^N\Bigg]. \tag{8.2}$$

Since ρ is assumed to be non negative, we have by (2.1)

$$\sum_{i\in\mathcal{E}} \alpha_i^N = \frac{1-A(\sqrt{p_N q_N})}{2p_N q_N} = \frac{1-|p_N - q_N|}{2p_N q_N} = \frac{1}{p_N},$$

and we can rewrite (8.2) as

$$\mathbb{P}\{T_{[Nt]}^N = k\} = \Bigg[1 - \frac{q_N}{p_N} + q_N \sum_{\substack{i\in\mathcal{E}:\\ i\geqslant k}} \alpha_i^N\Bigg]\Bigg[p_N \sum_{\substack{i\in\mathcal{E}:\\ i\geqslant [Nt]-k}} \alpha_i^N\Bigg]. \tag{8.3}$$

We aim to evaluate the limit of the foregoing quantity as $N\to\infty$. For this, we need an asymptotic for the sum $\sum_{\substack{i\in\mathcal{E}:\\ i\geqslant k}} \alpha_i^N$ as $k, N\to\infty$ with $c_1 N \leqslant k \leqslant c_2 N$ for any $c_1, c_2 > 0$.

Lemma 8.2. *The following asymptotics holds: for any $c_1, c_2 > 0$ such that $c_1 < c_2$,*

$$\sum_{\substack{i\in\mathcal{E}:\\ i\geqslant k}} \alpha_i^N \underset{\substack{k,N\to\infty\\ c_1 N\leqslant k\leqslant c_2 N}}{\sim} \sqrt{\frac{2}{\pi N}} \int_{k/N}^{\infty} \frac{\mathrm{e}^{-\rho^2 z/2}}{z^{3/2}}\,\mathrm{d}z. \tag{8.4}$$

The proof of this lemma is postponed to Appendix A. Set, for $z > 0$ and $s \in (0, t)$,

$$\varphi(z) = \frac{\mathrm{e}^{-\rho^2 z/2}}{z^{3/2}} \quad\text{and}\quad \psi(s,t) = \sqrt{\frac{2}{\pi}}\Bigg[\rho + \frac{1}{2\sqrt{2\pi}}\int_s^{\infty} \varphi(z)\,\mathrm{d}z\Bigg]\int_{t-s}^{\infty} \varphi(z)\,\mathrm{d}z.$$

In the light of (8.3) and (8.4), for even k such that $[Ns_1] < k \leqslant [Ns_2]$, we have

$$\mathbb{P}\{T_{[Nt]}^N = k\} \underset{\substack{N\to\infty\\ [Nt]\in\mathcal{E}}}{\sim} \Bigg[\frac{2\rho}{\sqrt{N}} + \frac{1}{\sqrt{2\pi N}}\int_{k/N}^{\infty} \varphi(z)\,\mathrm{d}z\Bigg]\Bigg[\frac{1}{\sqrt{2\pi N}}\int_{([Nt]-k)/N}^{\infty} \varphi(z)\,\mathrm{d}z\Bigg] \underset{\substack{N\to\infty\\ [Nt]\in\mathcal{E}}}{\sim} \frac{1}{N}\,\psi(k/n).$$

Finally,

$$\mathbb{P}\{s_1 < \frac{1}{N}\,T_{[Nt]}^N \leqslant s_2\} \underset{\substack{N\to\infty\\ [Nt]\in\mathcal{E}}}{\sim} \frac{1}{N} \sum_{\substack{k\in\mathcal{E}:\\ [Ns_1]\leqslant k\leqslant [Ns_2]}} \psi(k/N) \underset{\substack{N\to\infty\\ [Nt]\in\mathcal{E}}}{\to} \frac{1}{2}\int_{s_1}^{s_2} \psi(s,t)\,\mathrm{d}s.$$

• *Case where* $[Nt]$ *is odd.*
Assume that $[Nt]$ is odd. Then, by invoking (5.22), we obtain

$$\begin{aligned}\mathbb{P}\{s_1 < \frac{1}{N}\,T^{N}_{[Nt]} \leqslant s_2\} &= \sum_{\substack{k\in\mathcal{E}:\\ [Ns_1]<k\leqslant[Ns_2]}} \mathbb{P}\{T^{N}_{[Nt]} = k\} + \sum_{\substack{k\in\mathcal{O}:\\ [Ns_1]<k\leqslant[Ns_2]}} \mathbb{P}\{T^{N}_{[Nt]} = k\}\\ &= \sum_{\substack{k\in\mathcal{E}:\\ [Ns_1]<k\leqslant[Ns_2]}} \mathbb{P}\{T^{N}_{[Nt]} = k\} + \sum_{\substack{k\in\mathcal{E}:\\ [Ns_1]+1<k\leqslant[Ns_2]+1}} \mathbb{P}\{T^{N}_{[Nt]} = k-1\}\\ &= \sum_{k=[Ns_1]+1}^{[Ns_2]} \mathbb{P}\{T^{N}_{[Nt]+1} = k\} + \varepsilon_{N}\\ &= \mathbb{P}\{[Ns_1] < T^{N}_{[Nt]+1} \leqslant [Ns_2]\} + \varepsilon_{N}\end{aligned}$$

where $\varepsilon_{N} = 1\!\!1_{\mathcal{O}}([Ns_2])\mathbb{P}\{T^{N}_{[Nt]} = [Ns_2]\} - 1\!\!1_{\mathcal{O}}([Ns_1])\mathbb{P}\{T^{N}_{[Nt]} = [Ns_1]\}$. We have

$$\begin{aligned}|\varepsilon_{N}| &\leqslant \mathbb{P}\{T^{N}_{[Nt]} = [Ns_1]\} + \mathbb{P}\{T^{N}_{[Nt]} = [Ns_2]\}\\ &\leqslant \mathbb{P}\{T^{N}_{[Nt]+1} \in \{[Ns_1],[Ns_1]+1\}\} + \mathbb{P}\{T^{N}_{[Nt]+1} \in \{[Ns_2],[Ns_2]+1\}\}\\ &\underset{\substack{N\to\infty\\ [Nt]\in\mathcal{O}}}{\sim} \frac{1}{N}\left[\psi\Big(\frac{[Ns_1]}{N}\Big) + \psi\Big(\frac{[Ns_1]+1)}{N}\Big) + \psi\Big(\frac{[Ns_2]}{N}\Big) + \psi\Big(\frac{[Ns_2]+1)}{N}\Big)\right]\\ &\underset{\substack{N\to\infty\\ [Nt]\in\mathcal{O}}}{\sim} \frac{2}{N}\,[\psi(s_1)+\psi(s_1)].\end{aligned}$$

Since $[Nt]$ is odd, $[Nt]+1$ is even and we can use the foregoing analysis. Hence,

$$\mathbb{P}\{s_1 \leqslant \frac{1}{N}\,T^{N}_{[Nt]} \leqslant s_2\} \underset{\substack{N\to\infty\\ [Nt]\in\mathcal{O}}}{\sim} \frac{1}{2}\int_{s_1}^{s_2} \psi(s,t)\,\mathrm{d}s.$$

We know that $\frac{1}{N}\,T^{N}_{[Nt]} \underset{N\to\infty}{\longrightarrow} \mathbf{T}_t$, then

$$\mathbb{P}\{s_1 \leqslant \mathbf{T}_t \leqslant s_2\} = \frac{1}{2}\int_{s_1}^{s_2} \psi(s,t)\,\mathrm{d}s \quad\text{and}\quad \mathbb{P}\{\mathbf{T}_t \in \mathrm{d}s\}/\mathrm{d}s = \frac{1}{2}\,\psi(s,t)$$

which ends up the Proof of Theorem 8.1. □

For $\rho = 0$, we retrieve Paul Lévy's famous arcsine law for standard Brownian motion: for $s \in (0,t)$,

$$\mathbb{P}\{\mathbf{T}_t \in \mathrm{d}s\}/\mathrm{d}s = \frac{1}{\pi\sqrt{s(t-s)}}\cdot$$

Remark 8.3. Let t tend to ∞ in the formula of Theorem 8.1. The limiting random variable $\mathbf{T}_\infty$ denotes the total sojourn time in $\mathbb{R}^+$ and its probability distribution is given by

$$\mathbb{P}\{\mathbf{T}_\infty \in \mathrm{d}s\}/\mathrm{d}s = \begin{cases} 0 & \text{if } \rho \geqslant 0,\\ \dfrac{|\rho|}{\sqrt{2\pi}}\displaystyle\int_s^\infty \frac{\mathrm{e}^{-\rho^2 z/2}}{z^{3/2}}\,\mathrm{d}z & \text{if } \rho < 0.\end{cases}$$

From this, we deduce $\mathbb{P}\{\mathbf{T}_\infty = \infty\} = 1$ if $\rho \geqslant 0$. If $\rho < 0$,

$$\mathbb{P}\{\mathbf{T}_\infty < \infty\} = \frac{|\rho|}{\sqrt{2\pi}} \int_0^\infty \mathrm{d}s \int_s^\infty \frac{\mathrm{e}^{-\rho^2 z/2}}{z^{3/2}} \,\mathrm{d}z = \frac{|\rho|}{\sqrt{2\pi}} \int_0^\infty \frac{\mathrm{e}^{-\rho^2 z/2}}{\sqrt{z}} \,\mathrm{d}z = 1.$$

This is in good accordance with the effect of the drift near infinity.

Theorem 8.4. *The probability distribution functions of* $\mathbf{T}_t \mathbb{1}_{(-\infty,0)}(B_t)$ *and* $\mathbf{T}_t \mathbb{1}_{(0,+\infty)}(B_t)$ *admit the following expressions: for* $0 < s < t$,

$$\mathbb{P}\{\mathbf{T}_t \in \mathrm{d}s, B_t < 0\}/\mathrm{d}s = \frac{\rho^-}{\sqrt{2\pi}} \int_s^t \frac{\mathrm{e}^{-\rho^2 z/2}}{z^{3/2}} \,\mathrm{d}z + \frac{1}{4\pi} \int_s^t \frac{\mathrm{e}^{-\rho^2 u/2}}{u^{3/2}} \,\mathrm{d}u \int_{t-u}^\infty \frac{\mathrm{e}^{-\rho^2 v/2}}{v^{3/2}} \,\mathrm{d}v,$$
$$\mathbb{P}\{\mathbf{T}_t \in \mathrm{d}s, B_t > 0\}/\mathrm{d}s = \frac{\rho^+}{\sqrt{2\pi}} \int_{t-s}^t \frac{\mathrm{e}^{-\rho^2 z/2}}{z^{3/2}} \,\mathrm{d}z + \frac{1}{4\pi} \int_{t-s}^t \frac{\mathrm{e}^{-\rho^2 u/2}}{u^{3/2}} \,\mathrm{d}u \int_{t-u}^\infty \frac{\mathrm{e}^{-\rho^2 v/2}}{v^{3/2}} \,\mathrm{d}v.$$

Proof. Set $\alpha_i = a_i (pq)^{i/2}$. We first begin by rewriting $r^-_{k,n}$ as follows. By (7.3), we have

$$r^-_{k,n} = \mathbb{1}_{\mathcal{E}}(k) \Bigg[2q \sum_{\substack{i \in \mathcal{E}:\\ k \leqslant i \leqslant n-1}} \alpha_i - 4pq \sum_{\substack{i \in \mathcal{E}:\\ k \leqslant i \leqslant n-2}} \Bigg(\sum_{\substack{j \in \mathcal{E}:\\ j \leqslant i-k}} \alpha_j \alpha_{i-j} \Bigg) \Bigg].$$

The double sum of the foregoing expression can be written as

$$\begin{aligned}
\sum_{\substack{i \in \mathcal{E}:\\ k \leqslant i \leqslant n-2}} \Bigg(\sum_{\substack{j \in \mathcal{E}:\\ j \leqslant i-k}} \alpha_j \alpha_{i-j} \Bigg) &= \sum_{\substack{j \in \mathcal{E}:\\ j \leqslant n-k-2}} \alpha_j \Bigg(\sum_{\substack{i \in \mathcal{E}:\\ j+k \leqslant i \leqslant n-2}} \alpha_{i-j} \Bigg) = \sum_{\substack{j \in \mathcal{E}:\\ j \leqslant n-k-2}} \alpha_j \Bigg(\sum_{\substack{i \in \mathcal{E}:\\ k \leqslant i \leqslant n-j-2}} \alpha_i \Bigg) \\
&= \sum_{\substack{i \in \mathcal{E}:\\ k \leqslant i \leqslant n-2}} \alpha_i \Bigg(\sum_{\substack{j \in \mathcal{E}:\\ j \leqslant n-i-2}} \alpha_j \Bigg) = \sum_{\substack{i \in \mathcal{E}:\\ k \leqslant i \leqslant n-2}} \alpha_i \Bigg(\sum_{j \in \mathcal{E}} \alpha_j - \sum_{\substack{j \in \mathcal{E}:\\ j \geqslant n-i-1}} \alpha_j \Bigg) \\
&= \frac{1}{2(p \vee q)} \sum_{\substack{i \in \mathcal{E}:\\ k \leqslant i \leqslant n-2}} \alpha_i - \sum_{\substack{i \in \mathcal{E}:\\ k \leqslant i \leqslant n-2}} \alpha_i \Bigg(\sum_{\substack{j \in \mathcal{E}:\\ j \geqslant n-i-1}} \alpha_j \Bigg).
\end{aligned}$$

Therefore,

$$r^-_{k,n} = \mathbb{1}_{\mathcal{E}}(k) \Bigg[2q \mathbb{1}_{\mathcal{O}}(n) \alpha_{n-1} + 2(q-p)^+ \sum_{\substack{i \in \mathcal{E}:\\ k \leqslant i \leqslant n-2}} \alpha_i + 4pq \sum_{\substack{i \in \mathcal{E}:\\ k \leqslant i \leqslant n-2}} \alpha_i \Bigg(\sum_{\substack{j \in \mathcal{E}:\\ j \geqslant n-i-1}} \alpha_j \Bigg) \Bigg].$$

Recall that $\alpha_i^N = a_i (p_N q_N)^{i/2}$. We now search an asymptotic for the following probability:

$$\mathbb{P}\{s_1 < \frac{1}{N} T^N_{[Nt]} \leqslant s_2, B^N_t < 0\} = \sum_{k=[Ns_1]+1}^{[Ns_2]} \mathbb{P}\{T^N_{[Nt]} = k, S_{[Nt]} < 0\}$$

where

$$\begin{aligned}
\mathbb{P}\{T^N_{[Nt]} = k, S_{[Nt]} < 0\} = \mathbb{1}_{\mathcal{E}}(k) \Bigg[& 2q_N \mathbb{1}_{\mathcal{O}}([Nt]) \alpha^N_{[Nt]-1} + 2(q_N - p_N)^+ \sum_{\substack{i \in \mathcal{E}:\\ k \leqslant i \leqslant [Nt]-2}} \alpha^N_i \\
& + 4p_N q_N \sum_{\substack{i \in \mathcal{E}:\\ k \leqslant i \leqslant [Nt]-2}} \alpha^N_i \Bigg(\sum_{\substack{j \in \mathcal{E}:\\ j \geqslant [Nt]-i-1}} \alpha^N_j \Bigg) \Bigg].
\end{aligned}$$

Observing that, as $N \to \infty$,

$$(q_{{}_N} - p_{{}_N})^+ \sim \frac{2\rho^-}{\sqrt{N}}, \quad 4p_{{}_N} q_{{}_N} \sim 1, \quad 2q_{{}_N} \alpha^{N}_{[Nt]-1} = \mathcal{O}(N^{-3/2}),$$

and invoking (8.4), we get, for $k \in \mathcal{E}$,

$$\mathbb{P}\{T^{N}_{[Nt]} = k, S_{[Nt]} < 0\} \sim 2\sqrt{\frac{2}{\pi}}\,\frac{\rho^-}{N^2} \sum_{\substack{i\in\mathcal{E}:\\ k\leqslant i\leqslant [Nt]-2}} \varphi\Big(\frac{i}{N}\Big) + \frac{2}{\pi N^3} \sum_{\substack{i\in\mathcal{E}:\\ k\leqslant i\leqslant [Nt]-2}} \varphi\Big(\frac{i}{N}\Big) \sum_{\substack{j\in\mathcal{E}:\\ j\geqslant [Nt]-i-1}} \varphi\Big(\frac{j}{N}\Big). \tag{8.5}$$

We can easily see that

$$\sum_{\substack{i\in\mathcal{E}:\\ k\leqslant i\leqslant [Nt]-2}} \varphi\Big(\frac{i}{N}\Big) \sim \frac{N}{2}\int_{k/N}^{t} \varphi(z)\,\mathrm{d}z \tag{8.6}$$

and

$$\sum_{\substack{j\in\mathcal{E}:\\ j\geqslant [Nt]-i-1}} \varphi\Big(\frac{j}{N}\Big) \sim \frac{N}{2}\int_{([Nt]-i-1)/N}^{\infty} \varphi(v)\,\mathrm{d}v \sim \frac{N}{2}\int_{t-i/N}^{\infty} \varphi(v)\,\mathrm{d}v.$$

Moreover,

$$\begin{aligned}\sum_{\substack{i\in\mathcal{E}:\\ k\leqslant i\leqslant [Nt]-2}} \varphi\Big(\frac{i}{N}\Big) \sum_{\substack{j\in\mathcal{E}:\\ j\geqslant [Nt]-i-1}} \varphi\Big(\frac{j}{N}\Big) &\sim \frac{N}{2} \sum_{\substack{i\in\mathcal{E}:\\ k\leqslant i\leqslant [Nt]-2}} \varphi\Big(\frac{i}{N}\Big) \int_{t-i/N}^{\infty} \varphi(v)\,\mathrm{d}v \\ &\sim \frac{N^2}{4}\int_{k/N}^{t} \varphi(u)\,\mathrm{d}u \int_{t-u}^{\infty} \varphi(v)\,\mathrm{d}v.\end{aligned} \tag{8.7}$$

Put now

$$\chi^{\pm}(s,t) = \frac{2\rho^{\pm}}{\sqrt{2\pi}}\int_{s}^{t} \varphi(z)\,\mathrm{d}z + \frac{1}{2\pi}\int_{s}^{t} \varphi(u)\,\mathrm{d}u \int_{t-u}^{\infty} \varphi(v)\,\mathrm{d}v.$$

In the light of (8.5), (8.6) and (8.7), we derive

$$\mathbb{P}\{T^{N}_{[Nt]} = k, S_{[Nt]} < 0\} \underset{N\to\infty}{\sim} \frac{1}{N}\,\chi^-(k/N,t)$$

and finally

$$\mathbb{P}\{s_1 < \frac{1}{N}\,T^{N}_{[Nt]} \leqslant s_2, B^{N}_t < 0\} \underset{N\to\infty}{\sim} \frac{1}{N} \sum_{\substack{k\in\mathcal{E}:\\ [Ns_1]<k\leqslant [Ns_2]}} \chi^-(k/N,t) \underset{N\to\infty}{\longrightarrow} \frac{1}{2}\int_{s_1}^{s_2} \chi^-(s,t)\,\mathrm{d}s.$$

Since $\frac{1}{N}\,T^{N}_{[Nt]}\mathbb{1}_{(-\infty,0)}(B^{N}_t) \underset{N\to\infty}{\longrightarrow} \mathbf{T}_t\mathbb{1}_{(-\infty,0)}(B_t)$, we deduce that

$$\mathbb{P}\{s_1 \leqslant \mathbf{T}_t \leqslant s_2, B_t < 0\} = \frac{1}{2}\int_{s_1}^{s_2} \chi^-(s,t)\,\mathrm{d}s \quad \text{and} \quad \mathbb{P}\{\mathbf{T}_t \in \mathrm{d}s, B_t < 0\}/\mathrm{d}s = \frac{1}{2}\,\chi^-(s,t).$$

We can prove in a quite similar way that

$$\mathbb{P}\{\mathbf{T}_t \in \mathrm{d}s, B_t > 0\}/\mathrm{d}s = \frac{1}{2}\,\chi^+(t-s,t).$$

This last result can be deduced also from the previous one by invoking a duality argument related to drifted Brownian motion. Indeed, setting more precisely $B_t = B^{\rho}_t$ and $\mathbf{T}_t = \mathbf{T}^{\rho}_t$, it can be easily seen that $(B^{\rho}_t, \mathbf{T}^{\rho}_t)$ and

$(-B_t^{-\rho}, t-\mathbf{T}_t^{-\rho})$ have the same distributions. This explains why the term $\chi^-(s,t)$ is changed into $\chi^+(t-s,t)$. The Proof of Theorem 8.4 is finished. □

For $\rho = 0$, we retrieve the well-known results for standard Brownian motion: for $s \in (0,t)$,

$$\mathbb{P}\{\mathbf{T}_t \in \mathrm{d}s, B_t < 0\}/\mathrm{d}s = \frac{1}{\pi t}\sqrt{\frac{t-s}{s}} \quad \text{and} \quad \mathbb{P}\{\mathbf{T}_t \in \mathrm{d}s, B_t > 0\}/\mathrm{d}s = \frac{1}{\pi t}\sqrt{\frac{s}{t-s}}.$$

Appendix A

Proof of Lemma 8.2. The main idea is roughly speaking that, referring to Stirling formula for a_i,

$$a_i = \frac{1}{i+2}\binom{i}{i/2} \underset{\substack{i\to\infty\\ i\in\mathcal{E}}}{\sim} \frac{2^{i+1/2}}{\sqrt{\pi}\, i^{3/2}} \quad \text{and} \quad (p_{\scriptscriptstyle N} q_{\scriptscriptstyle N})^{i/2} = \left[\frac{1}{4}\left(1-\frac{\rho^2}{N}\right)\right]^{i/2} \underset{\substack{i,N\to\infty\\ i/N^2\to 0}}{\sim} \frac{1}{2^i}\,\mathrm{e}^{-\rho^2 i/(2N)}.$$

Then, recalling that $\alpha_i^{\scriptscriptstyle N} = a_i(p_{\scriptscriptstyle N} q_{\scriptscriptstyle N})^{i/2}$,

$$\alpha_i^{\scriptscriptstyle N} \underset{\substack{i,N\to\infty\\ i/N^2\to 0}}{\sim} \sqrt{\frac{2}{\pi}}\,\frac{\mathrm{e}^{-\rho^2 i/(2N)}}{i^{3/2}}$$

and next

$$\sum_{\substack{i\in\mathcal{E}:\\ i\geqslant k}} \alpha_i^{\scriptscriptstyle N} \underset{\substack{k,N\to\infty\\ c_1 N\leqslant k\leqslant c_2 N}}{\sim} \sqrt{\frac{2}{\pi}} \sum_{\substack{i\in\mathcal{E}:\\ i\geqslant k}} \frac{\mathrm{e}^{-\rho^2 i/(2N)}}{i^{3/2}} = \frac{\sqrt{2}}{\sqrt{\pi}\,N^{3/2}} \sum_{\substack{i\in\mathcal{E}:\\ i\geqslant k}} \varphi(i/N) \underset{\substack{k,N\to\infty\\ c_1 N\leqslant k\leqslant c_2 N}}{\sim} \sqrt{\frac{2}{\pi N}} \int_{k/N}^{\infty} \varphi(z)\,\mathrm{d}z.$$

To make the argument more precise, we note that

$$\left(1-\frac{\rho^2}{N}\right)^{i/2} = \exp\left[\frac{i}{2}\,\ln\!\left(1-\frac{\rho^2}{N}\right)\right] = \mathrm{e}^{-\rho^2 i/(2N)}\,\exp\left[\frac{i}{2}\left(\ln\!\left(1-\frac{\rho^2}{N}\right)+\frac{\rho^2}{N}\right)\right]$$

from which we see that

$$\left(1-\frac{\rho^2}{N}\right)^{i/2} \underset{\substack{i,N\to\infty\\ i/N^2\to 0}}{\sim} \mathrm{e}^{-\rho^2 i/(2N)}.$$

Pick $\varepsilon > 0$. There exists an integer N_0 such that, for any $N \geqslant N_0$, for any $i,k \in \mathcal{E}$ such that $c_1 N \leqslant k \leqslant i \leqslant N^{3/2}$,

$$\alpha_i^{\scriptscriptstyle N} \geqslant (1-\varepsilon)\,\frac{\sqrt{2}}{\sqrt{\pi}\, i^{3/2}}\,\mathrm{e}^{-\rho^2 i/(2N)}$$

and for any $i,k \in \mathcal{E}$ such that $c_1 N \leqslant k \leqslant i \wedge (c_2 N)$,

$$\alpha_i^{\scriptscriptstyle N} \leqslant (1+\varepsilon)\,\frac{\sqrt{2}}{\sqrt{\pi}\, i^{3/2}}\,\mathrm{e}^{-\rho^2 i/(2N)}.$$

Then, for any $N \geqslant N_0$ and any k such that $c_1 N \leqslant k \leqslant c_2 N$,

$$(1-\varepsilon)\,\frac{\sqrt{2}}{\sqrt{\pi}\,N^{3/2}} \sum_{\substack{i\in\mathcal{E}:\\ k\leqslant i\leqslant N^{3/2}}} \varphi\!\left(\frac{i}{N}\right) \leqslant \sum_{\substack{i\in\mathcal{E}:\\ i\geqslant k}} \alpha_i^{\scriptscriptstyle N} \leqslant (1+\varepsilon)\,\frac{\sqrt{2}}{\sqrt{\pi}\,N^{3/2}} \sum_{\substack{i\in\mathcal{E}:\\ i\geqslant k}} \varphi\!\left(\frac{i}{N}\right). \tag{A.1}$$

Remarking that φ is decreasing, we have

$$\frac{N}{2}\int_{i/N}^{(i+2)/N}\varphi(z)\,\mathrm{d}z \leqslant \varphi\Big(\frac{i}{N}\Big) \leqslant \frac{N}{2}\int_{(i-2)/N}^{i/N}\varphi(z)\,\mathrm{d}z.$$

Then

$$\frac{N}{2}\int_{k/N}^{\infty}\varphi(z)\,\mathrm{d}z \leqslant \sum_{\substack{i\in\mathcal{E}:\\ i\geqslant k}}\varphi\Big(\frac{i}{N}\Big) \leqslant \varphi\Big(\frac{k}{N}\Big)+\frac{N}{2}\int_{k/N}^{\infty}\varphi(z)\,\mathrm{d}z$$

which shows, since we plainly have for $c_1N\leqslant k\leqslant c_2N$, $\varphi(k/N)=o(N\int_{k/N}^{\infty}\varphi(z)\,\mathrm{d}z)$, that

$$\sum_{\substack{i\in\mathcal{E}:\\ i\geqslant k}}\varphi\Big(\frac{i}{N}\Big) \underset{\substack{k,N\to\infty\\ c_1N\leqslant k\leqslant c_2N}}{\sim} \frac{N}{2}\int_{k/N}^{\infty}\varphi(z)\,\mathrm{d}z. \tag{A.2}$$

On the other hand,

$$\sum_{\substack{i\in\mathcal{E}:\\ k\leqslant i\leqslant N^{3/2}}}\varphi\Big(\frac{i}{N}\Big)=\sum_{\substack{i\in\mathcal{E}:\\ i\geqslant k}}\varphi\Big(\frac{i}{N}\Big)-\sum_{\substack{i\in\mathcal{E}:\\ i>N^{3/2}}}\varphi\Big(\frac{i}{N}\Big). \tag{A.3}$$

The last sum in (A.3) can be estimated as follows:

$$\sum_{\substack{i\in\mathcal{E}:\\ i>N^{3/2}}}\varphi\Big(\frac{i}{N}\Big)\leqslant\frac{N}{2}\int_{\sqrt{N}}^{\infty}\varphi(z)\,\mathrm{d}z+\varphi\Big(\frac{[N^{3/2}]}{N}\Big)+\varphi\Big(\frac{[N^{3/2}]+1}{N}\Big)\underset{N\to\infty}{\sim}\frac{N}{2}\int_{\sqrt{N}}^{\infty}\varphi(z)\,\mathrm{d}z\underset{N\to\infty}{=}o(N). \tag{A.4}$$

Putting (A.2) and (A.4) into (A.3), we obtain that

$$\sum_{\substack{i\in\mathcal{E}:\\ k\leqslant i\leqslant N^{3/2}}}\varphi\Big(\frac{i}{N}\Big)\underset{N\to\infty}{\sim}\frac{N}{2}\int_{k/N}^{\infty}\varphi(z)\,\mathrm{d}z \tag{A.5}$$

and next putting (A.2) and (A.5) into (A.1), we find, for N large enough and $c_1N\leqslant k\leqslant c_2N$, that

$$(1-2\varepsilon)\,\frac{\sqrt{2}}{\sqrt{\pi}\,N}\int_{k/N}^{\infty}\varphi(z)\,\mathrm{d}z\leqslant\sum_{\substack{i\in\mathcal{E}:\\ i\geqslant k}}\alpha_i^N\leqslant(1+2\varepsilon)\,\frac{\sqrt{2}}{\sqrt{\pi}\,N}\int_{k/N}^{\infty}\varphi(z)\,\mathrm{d}z$$

from which we finally deduce (8.4). □

Acknowledgements. The author thanks two anonymous referees for many comments. In particular, one of them suggested the elegant proof of Theorems 3.1 and 6.1 by using the excursions theory for random walk. The proofs in the first draft were based on a recursive relationship concerning the probabilities $r_{k,n}$ (as in [4]) and can be found in the report [7].

References

[1] P. Billingsley, *Convergence of probability measures.* John Wiley & Sons (1968).
[2] A.-N. Borodin and P. Salminen, Handbook of Brownian motion – facts and formulae, *Probability and its Applications.* Birkhäuser Verlag (1996).
[3] V. Cammarota, A. Lachal and E. Orsingher, Some Darling-Siegert relationships connected with random flights. *Stat. Probab. Lett.* **79** (2009) 243–254.

[4] K.-L. Chung and W. Feller, On fluctuations in coin-tossings. *Proc. Natl. Acad. Sci. USA* **35** (1949) 605–608.
[5] W. Feller, *An introduction to probability theory and its applications* I, 3rd edition. John Wiley & Sons (1968).
[6] P. Flajolet and R. Sedgewick, *Analytic combinatorics*. Cambridge University Press, Cambridge (2009).
[7] A. Lachal, `arXiv:1003.5009[math.PR]` (2010).
[8] A. Rényi, *Calcul des probabilités*. Dunod (1966).
[9] E. Sparre Andersen, On the number of positive sums of random variables. *Skand. Aktuarietidskrift* (1949) 27–36.
[10] E. Sparre Andersen, On the fluctuations of sums of random variables I-II. *Math. Scand.* **1** (1953) 263–285; **2** (1954) 195–223.
[11] F. Spitzer, Principles of random walk, 2nd edition. *Graduate Texts in Mathematics* **34** (1976).