section 1, Introduction

The problem we are concerned with deals with independent and identically distributed Rademacher-$\left(\frac{1}{2}\right)$ random varibles X_i, $i \in \mathbb{N}$, with values in $\{-1,+1\}$. For fixed $n \in \mathbb{N}$ we can understand a realization of the random vector $\mathbf{X}_n := (X_1, \ldots, X_n)$ as a lattice path on $\mathbb{N}_0 \times \mathbb{Z}$. We are interested in the joint distribution of the following three random variables:

$S_n := \sum_{i=1}^{n} X_i$

$A_n := \sum_{i=1}^{n} S_i$

$\tau(-1) := \min\{i \in \mathbb{N}: S_i = -1\}$

$(S_0 = A_0 = 0)$

So we fix a probability space $(\Omega_n, \mathcal{P}(\Omega_n), \mathbb{P})$ as follows:

$\Omega_n := \{\{z_1, \ldots, z_n\}: \forall i \in \{1, \ldots, n\} \text{ it holds } z_i \in \{-1,+1\}\}$, $\mathcal{P}(\Omega_n)$ is the power-set of Ω_n, and $\mathbb{P}$ the uniform probability measure. The term in question now is:

$\mathbb{P}(A_n \geqslant x, \tau(-1) > n | S_n = y)$ for $n \in \mathbb{N}$, $y \in \mathbb{R}_0^+$ and $x \in \mathbb{R}^+$.

Graphicly speaking, we are interested in how many paths, with fixed start- and endpoint, and the further condition to stay non-negative always, lead to the same area locked up between the path and the $x-$axis. We use the following decomposition of our term in question:

$$\mathbb{P}(A_n = t, \tau(-1) > n | S_n = y) = \mathbb{P}(A_n = t | S_n = y) - \mathbb{P}(A_n = t, \tau(-1) \leqslant n | S_n = y) =$$

$$= \mathbb{P}(A_n = t | S_n = y) - \sum_{i=1}^{n} \mathbb{P}(A_n = t, \tau(-1) = i | S_n = y) =$$

$$= \mathbb{P}(A_n = t | S_n = y) - \frac{1}{\mathbb{P}(S_n = y)} \sum_{i=1}^{n} \sum_{\varphi \in \Phi(i)} \mathbb{P}(A_i = \varphi, \tau(-1) = i)\, \mathbb{P}\left(\tilde{A}_{n-i} = t - \varphi, \tilde{S}_{n-i} = y\right), \quad (\maltese)$$

with $t \in \Psi(n)$ and $\tilde{S}_0 = -1$. While we use this notation:

$x_n^+ := \max\{\varsigma \in \mathbb{N}: \exists \omega \in \Omega_n \text{ mit } A_n(\omega) = \varsigma, S_n(\omega) = y\} = \frac{1}{4}(n^2 - y^2 + 2y + 2n\, y)$

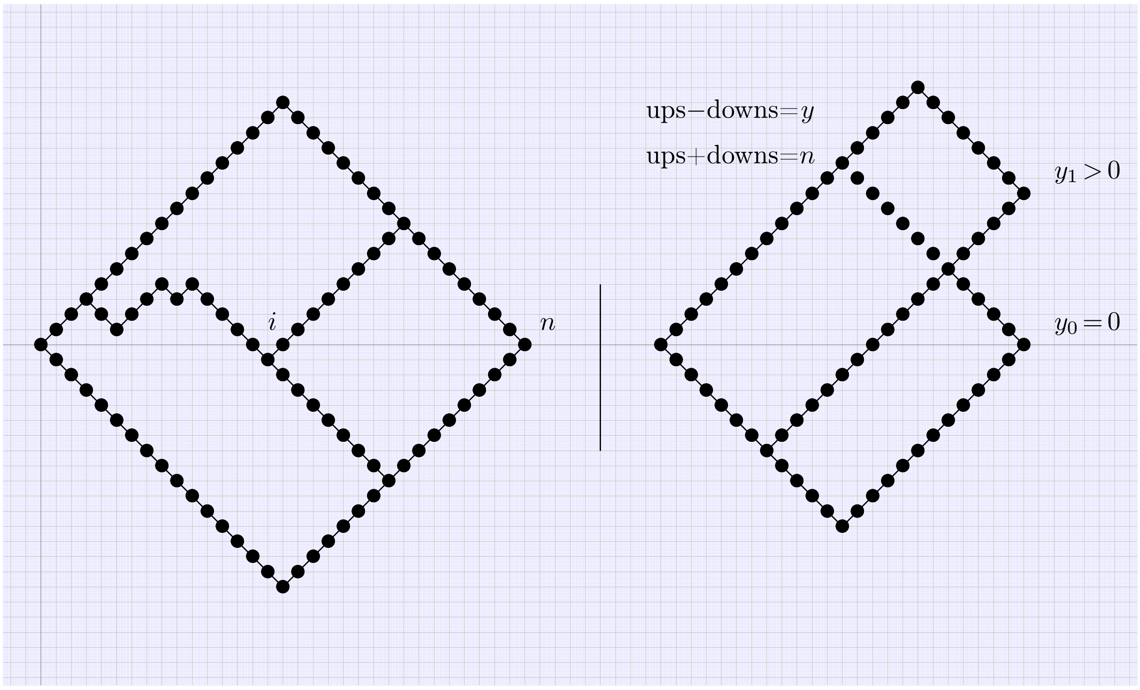
$x_n^- := \min\{\varsigma \in \mathbb{N}: \exists \omega \in \Omega_n \text{ mit } A_n(\omega) = \varsigma, S_n(\omega) = y\} = \frac{1}{4}(y^2 - n^2 + 2y + 2n\, y)$,

$\theta_i := \frac{i-1}{2}$, where the interesting values of i are only the uneven numbers smaller than n.

$z_i^+ := \max\{A_i : \tau(-1) = i\} = \theta_i^2 - 1$, and $z_i^- := \min\{A_i : \tau(-1) = i\} = \theta_i - 1$,

$\Phi(i) := \left\{ z_i^+ - 2\vartheta : \vartheta = 0, 1, 2, ..., \binom{\theta_i}{2} \right\}$, and $\Psi(n) := \{x_n^+ - 2\vartheta : \vartheta = 0, 1, 2, ..., x_n^+\}$.

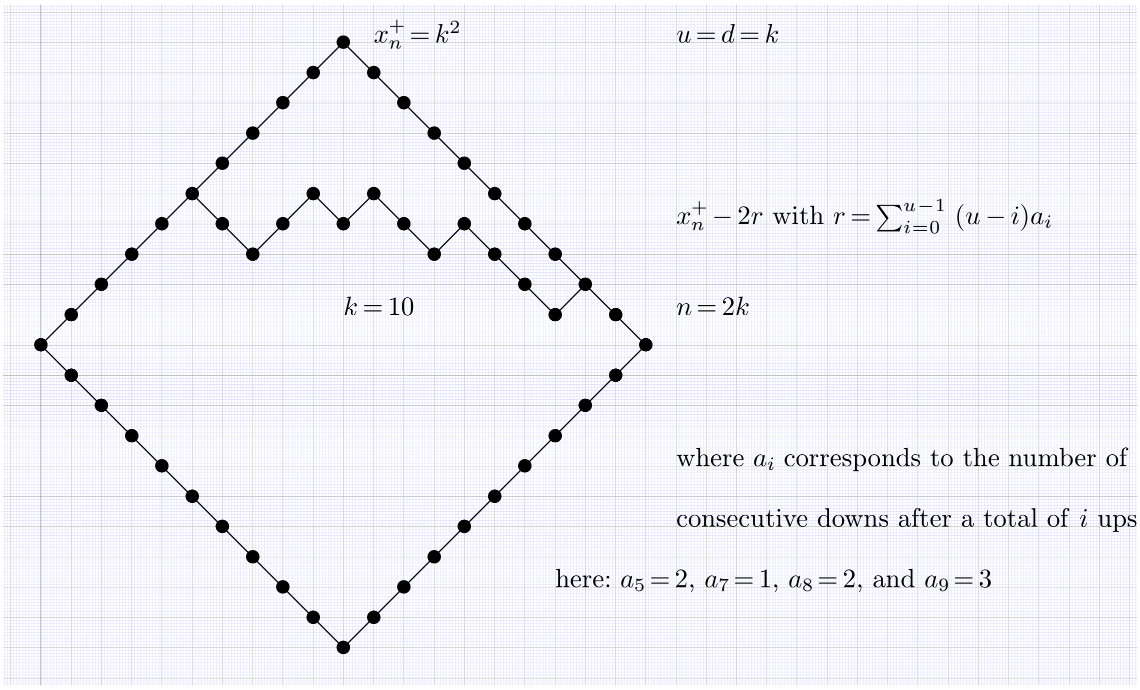
The following picture visualizes the situation.



In the terms of (✠), without the random variable $\tau(-1)$, we find as common ground that the possible paths have the same start- and endpoint without further restrictions, and thus their distribution only depends on the pathlength and the ups and downs, meaning ups $= u := \frac{n+y}{2}$ and downs $= d := \frac{n-y}{2}$. We will investigate this distrbution generally in the following section, depending on the values n and y. The remaining terms get treated separately in section 3.

section 2, Integer Partitions

For all $n \in \mathbb{N}$ we have an unique $\omega \in \Omega_n$ with $A_n(\omega) = x_n^+$. Since we have for all $\omega \in \Omega_n$, that $A_n(\omega) \in \Psi(n)$, we need to understand how the deviation of $A_n(\omega)$ from x_n^+ can be used to determining the probability $\mathbb{P}(A_n = x_n^+ - 2r | S_n = y)$. Due to brevity and clarity we will focus on the special case, where $y = 0$, and try to convince the reader that the general case is also covered by the presented strategy, which gets illustrated by the following picture.



By renaming ($\#^j \equiv a_{u-j}$) we arrive at the following representation of r:

$r = \sum_{i=0}^{u-1} (k-i)a_i = \sum_{j=1}^{u} j\,\#^j = \#^1 + 2\,\#^2 + \ldots + u\,\#^u$

Since each value of r corresponds to a number of paths $\omega \in \Omega_n$ with $S_n(\omega) = y$, we have generally to fulfill the extra condition $\sum_{j=1}^{u} \#^j \leqslant d$.

$\mathrm{IP}(r) := \left\{ (\#^1, \ldots \#^r) \in \{0, \ldots, r\}^r : r = \sum_{j=1}^{r} j\,\#^j \right\}$ is the set of integer partitions of an integer r.

Generally we are interested in the following function: (for the special case $y = 0$, we have $u = d = k$)

$\xi: \mathbb{N}^3 \rightarrow \mathbb{N}$ with $\xi(r,u,d) = \left| \left\{ (\#^1, \ldots, \#^u) \in \{0, \ldots, d\}^u : r = \sum_{j=1}^{u} j\,\#^j, \sum_{j=1}^{u} \#^j \leqslant d \right\} \right|$

Also, $|\mathrm{IP}(r)| = \xi(r,r,r)$, and asymptotically we have $|\mathrm{IP}(r)| \sim \frac{e^{\sqrt{\frac{2\pi}{3}}}}{4\,r\,\sqrt{3}}$. [1]

An exact formula for $|\mathrm{IP}(r)|$ is rather involved and was presented by Hans Rademacher in 1937.[1]

We arrive at $\xi(r,u,d)$ through splitting the cases:

subsection 2.1, $\xi(r,u,r)$

We consider the generating function R for integer partions which is defined for $|x| < 1$. Furthermore we use this notation:

$a_u(r) := \xi(r,u,r)$ and $q(j,i) := \mathbb{1}_{\{i=0\} \cup \left\{ \frac{j}{i} \in \mathbb{N} \right\}}$.

Of course we have for $p, q \in \mathbb{N}$ with $p \leqslant q$,that $\xi(p,q,p) = \xi(p,p,p)$. Furthermore we have $\forall m \in \mathbb{N}$ $a_m(0) = 1$, but $\forall l \in \mathbb{N}$ it is $a_0(l) = 0$, and follow the convention $a_0(0) = 1$.

$R(u) := \prod_{m=1}^{u} \frac{1}{(1-x^m)} = \prod_{m=1}^{u} \sum_{j=0}^{\infty} q(j,m)\, x^j = \sum_{j=0}^{\infty} a_u(j)\, x^j$

So, for $t \in \mathbb{N}$

$\sum_{j=0}^{\infty} q(j,1) x^j = \sum_{j=0}^{\infty} a_1(j) x^j$, and

$\sum_{j=0}^{\infty} a_t(j) x^j = \sum_{j=0}^{\infty} a_{t-1}(j) x^j \sum_{j=0}^{\infty} q(j,t) x^j$, and therefore

$a_t(i) = \sum_{l=0}^{i} q(l,t) a_{t-1}(i-l)$

Hence we can conclude, that for all $t \in \mathbb{N}$ and $i \in \mathbb{N}_0$ there are numbers $a_t^j, \alpha_t^i \in \mathbb{Q}_0^+$, such that

$a_t(i) := a_t^{t-1} i^{t-1} + a_t^{t-2} i^{t-2} + \ldots + a_t^0 - \alpha_t^i$, e.g.

$a_1(i) \equiv 1 = a_1^0$

$a_2(i) = \frac{i}{2} + 1 - \sigma_{i,2} = a_2^1 i + a_2^0 - \alpha_2^i$

$a_3(i) = \frac{i^2}{12} + \frac{7i}{12} + 1 - \left(\sigma_{i,3}/4 + (3\sigma_{i,3}^2)/4 + \sum_{t=0}^{[i/3]} \sigma_{i-3t,2}\right) = a_3^2 i^2 + a_3^1 i + a_3^0 - \alpha_3^i$, where we use the notation $\sigma_{i,t} := \frac{1}{t}\,(i \bmod t)$,for which holds $\left[\frac{i}{t}\right] = \frac{i}{t} - \sigma_{i,t}$,with $[\bullet]$ being the an integers floor function.

One can derive, that $a_t^{t-1} = \frac{1}{t!\,(t-1)!}$ and $a_t^{t-2} = \frac{(1+t)}{4\,(t-1)!\,(t-2)!}$.

Theorem 1:

For all $r, t \in \mathbb{N}$ with $t < r$ it holds:

1, $a_r(r) = \sum_{j=1}^{r} a_j(r-j)$

2, $a_t(r) = a_r(r) - \sum_{m=0}^{r-t-1} a_{r-1-m}(m)$.

Corollary:

1, For $t < r \leqslant 2t+1$ we have:

$a_t(r) = a_r(r) - \sum_{m=0}^{r-t-1} a_m(m)$

2, For $2t+2 \leqslant r \leqslant 3t+2$ we have:

$a_t(r) = a_r(r) - \sum_{m=0}^{\left[\frac{r-1}{2}\right]} a_m(m) - \sum_{m=\left[\frac{r-1}{2}\right]+1}^{r-t-1} \left(a_m(m) - \sum_{j=0}^{2m-r} a_j(j)\right)$.

subsection 2.2, $\xi(r,u,d)$

$\xi(r,u,d) = \xi(r,u,r) - \left|\left\{(\#^1, \ldots, \#^u) \in \{0, \ldots, r\}^u : r = \sum_{j=1}^{u} j\#^j, \sum_{j=1}^{u} \#^j > d\right\}\right| =: \xi(r,u,r) - g(r,u,d)$

Theorem 2:

a, Generally for $r \geqslant 2t+1$:

$\xi(r,t,t) = \xi(r,t,r) - g(r,t,t) = \xi(r,t,r) - \sum_{l=0}^{t-1} a_l(l) - \sum_{l=k}^{r-t-1} \xi(l, t-1, r-l)$

b, In particular:

1, for $r \in \{t+1, \ldots, 2t\}$ we have:

$\xi(r,t,t) = \xi(r,t,r) - g(r,t,t) = a_t(r) - \sum_{m=0}^{r-t-1} a_m(m)$

2, for $r \in \{2t+1, ..., 3t+1\}$ we have:
$\xi(r,t,t) = a_t(r) - \sum_{l=0}^{t-1} a_l(l) - \sum_{l=k}^{r-t-1} \left(a_{t-1}(l) - \sum_{m=0}^{2l-r-1} a_{t-2}(m)\right).$
The most general setting with $r > d$ and $u > d$ can be treated with the same technique used in the proof of *Theorem 2*, but the case distinction then becomes a bit longish as do the formulas.

section 3, non-negative paths

We consider the Index i in (✠) fixed as well as the value of $j \in \left\{0, 1, 2, ..., \binom{\theta_i}{2}\right\}$, and work with $\theta \equiv \theta_i$. The number of possible paths here is $C_\theta := \binom{2\theta}{\theta}\frac{1}{\theta+1}$, while these numbers are known as Catalan numbers, for $\theta \in \mathbb{N}$. The ones leading to $A_i = z_i^+ - 2j$ is captured by the term $F_\theta(\theta + 2j)$, for which we introduce its generating generating function: $\boldsymbol{\varphi}_\theta(z) = \sum_{j=0}^{\binom{\theta}{2}} F_\theta(\theta + 2j) z^j$.

Theorem 3: [2]
It holds, that $\boldsymbol{\varphi}_m(z) = \sum_{l=1}^{m} \boldsymbol{\varphi}_{m-l}(z)\boldsymbol{\varphi}_{l-1}(z) z^{l-1}$ für $m = 1, 2, ...$ und $\boldsymbol{\varphi}_0(z) = 1$.

section 4, consolidation

Theorem 4:
$\mathbb{P}(A_n = x_n^+ - 2r, \tau(-1) > n | S_n = y) = \frac{\xi(r,u,d)}{\binom{n}{u}} - \sum_{i=1}^{n-1} \mathbb{1}_{\left\{\frac{i}{2} \notin \mathbb{N}\right\}} \sum_{\varphi \in \Phi(i)} \left(\frac{F_{\theta_i}(\varphi)\xi(\bar{r}(\varphi), \bar{u}, \bar{d})\mathbb{1}_{\{\bar{r} \in \overline{G}_n\}}}{\binom{n}{u}}\right)$

using this notation:
1, $n - i = \bar{u} + \bar{d}$ and $y + 1 = \bar{u} - \bar{d}$, hence $\bar{u} = \frac{n-i+y+1}{2}$ and $\bar{d} = \frac{n-i-y-1}{2}$
2, $\bar{x}_n^+ = \bar{u}(\bar{u} - 2) + \frac{1}{2}(y - y^2) + 1$ and $\bar{x}_n^- = \bar{d}(\bar{d} - 2) + \frac{1}{2}(y + y^2)$
3, $\bar{G}_n := \left\{0, 1, ..., \frac{\bar{x}_n^+ - \bar{x}_n^-}{2}\right\}$, and 4, $\bar{r}(\varphi) = \frac{\bar{x}_n^+ - x_n^+ + 2r + \varphi}{2}$.

When we have $y = 0$, we can assume without loss of generality, that $r \in \left\{0, 1, ..., \frac{k^2 - k}{2}\right\}$, since otherwise (✠) is a priori 0.

Theorem 5:
For $r = \mu k^2$ and $\mu < \left(\ln(4)\frac{1}{\pi\sqrt{\frac{2}{3}}}\right)^2 \approx 0.292$ we have:
$\mathbb{P}(A_n \geqslant x_n^+ - 2r, \tau(-1) > n | S_n = 0) \to 0 \quad \text{as} \quad n \to \infty.$

References:
[1] Philippe Flajolet and Robert Sedgewick; Analytic Combinatorics; Cambridge University Press (2009), pp. 39-47 and 574-579.
[2] Lajos Takács; A Bernoulli Excursion and Its Various Applications; Advances in Applied Probability, Vol. 23, No. 3 (Sep., 1991), pp. 557-585.