RETURN STATISTICS OF SIMPLE RANDOM WALKS

Peter Kirschenhofer and Helmut Prodinger

Department of Algebra and Discrete Mathematics
Technical University of Vienna, Austria

1. Introduction

In this paper we present some further results concerning simple random walks

$$S_m = \sum_{k=1}^{m} X_k \qquad \text{with} \quad S_0 = 0$$

where X_k, $k = 1, 2, \dots$ are independent and identically distributed random variables with $\mathbb{P}\{X_k = 1\} = \mathbb{P}\{X_k = -1\} = \frac{1}{2}$.

Our investigations were motivated by some recent work of W. Katzenbeisser and W. Panny [5], [6] and A.W. Kemp [7] on the number of visits to the origin of a random walk that ends at level 0 resp. the number of times where such a random walk reaches its maximum. (Other references to this kind of problems are e.g. [1, 4].) In the short note [9] we sketched an alternative approach to the first problem, where we give "finite" formulæ for all moments of the corresponding random variables.

It is the aim of this paper to extend this approach to a larger class of walks and different problems. In particular we want to emphasize the methodological point of view, i.e. the use of combinatorial constructions as well as their associated generating functions. This approach not only yields explicit formulæ in a straightforward and elegant manner, but also allows to "transfer" the results immediately to asymptotic expansions. An excellent reference to these techniques in general is e.g. [3].

2. Return Statistics

Consider the simple random walk

$$S_m = \sum_{k=1}^{m} X_k, \quad 0 \le m \le N, \text{ with } S_0 = 0 \text{ and } S_N = i \ \ (N \equiv i \bmod 2),$$

i.e. a simple random walk starting at 0 and ending at level i after N steps (compare Figure 1). Let the variable T be the number of visits to the origin (the starting point is not counted).

Key words and phrases. random walks, generating functions, moments, singularity analysis.

Let $\mathcal{W}_i$ denote the family of random walks in question with arbitrary length. It is our intention to decompose this family combinatorically into substructures of similar or easier type obtaining in this way "symbolic equations".

A simple random walk starting at 0 and ending at level i

Fig. 1

We assume $i \neq 0$ at first. By symmetry, it is sufficient to study the case $i > 0$.

Let us first cut each random walk in $\mathcal{W}_i$ at his last visit to the origin. Then the first part is an element of $\mathcal{W}_0$, since it ends at level 0, whereas the second part is in the set $\mathcal{N}_i$ of random walks that start in the origin, end at level i, but do not visit the origin in between; symbolically this reads

$$\mathcal{W}_i = \mathcal{W}_0 \times \mathcal{N}_i. \tag{2.1}$$

Now we focus on $\mathcal{W}_0$: Let $\mathcal{W}_{0,+}$ denote the subset of walks in $\mathcal{W}_0$ that are strictly positive between the first and the last point, and $\mathcal{W}_{0,-}$ the corresponding set with a strictly negative sojourn. Noting that between any two consecutive returns to the 0-level a walk is either positive ($\in \mathcal{W}_{0,+}$) or negative ($\in \mathcal{W}_{0,-}$), we get that $\mathcal{W}_0$ can be decomposed as a *sequence* of elements in $\mathcal{W}_{0,+} \cup \mathcal{W}_{0,-}$; symbolically

$$\mathcal{W}_0 = \left(\mathcal{W}_{0,+} \cup \mathcal{W}_{0,-}\right)^*. \tag{2.2}$$

Altogether we have

$$\mathcal{W}_i = \left(\mathcal{W}_{0,+} \cup \mathcal{W}_{0,-}\right)^* \times \mathcal{N}_i. \tag{2.3}$$

Equation (2.3) can be translated immediately into an equation for the corresponding ordinary generating functions. If we count the steps by the variable z, the returns by the variable u and observe that disjoint unions, cartesian products and stars translate to sums, Cauchy products and geometric series, respectively, we get

$$W_i(z,u) = \frac{1}{1 - \left(W_{0,+}(z,u) + (W_{0,-}(z,u)\right)} N_i(z,u). \tag{2.4}$$

Of course, $W_{0,+}(z,u) = W_{0,-}(z,u)$ since the corresponding situations are symmetric. Furthermore we have

$$W_{0,+}(z,u) = u \cdot W_{0,+}(z,1), \tag{2.5}$$

since we have exactly one return. It is this property that makes the described decomposition especially useful.

$W_{0,+}(z,1)$ is now the counting function of the positive random walks, and it is well-known that this is $C(z^2)$, where

$$C(z) = \frac{1 - \sqrt{1-4z}}{2} = \sum_{n \geq 1} \frac{1}{n} \binom{2n-2}{n-1} z^n \tag{2.6}$$

is one version of the generating function of the CATALAN–numbers. At this stage we have proved

$$W_i(z,u) = \frac{1}{1 - 2uC(z^2)} \cdot N_i(z,u). \tag{2.7}$$

Since we have assumed $i > 0$, there is no return to the 0–level, and we have $N_i(z,u) = N_i(z,1)$. Furthermore each walk in $\mathcal{N}_i$ can be decomposed according to its last points on the levels $1, 2, \ldots, i$ $(i > 0)$ into an i–tuple of walks that start at level 0, are strictly positive and end at level 1 (compare Fig. 2).

Decomposition of a walk in $\mathcal{N}_i$

Fig. 2

It follows that

$$N_i(z,1) = z^i \left(\frac{C(z^2)}{z^2}\right)^i = \frac{C^i(z^2)}{z^i}. \tag{2.8}$$

Altogether

$$W_i(z,u) = \frac{C^i(z^2)}{z^i\big(1 - 2uC(z^2)\big)}. \tag{2.9}$$

Inserting (2.6) we get the explicit form

$$W_i(z,u) = \frac{1}{1 - u + u\sqrt{1-4z^2}} \left(\frac{1-\sqrt{1-4z^2}}{2z}\right)^i. \tag{2.10}$$

The generating function of the s–th factorial moments multiplied by the number of walks of length N in $\mathcal{W}_i$ is given by

$$M_i^{(s)}(z) = \frac{\partial^s}{\partial u^s} W_i(z,u)\bigg|_{u=1} = s! \frac{\big(1-\sqrt{1-4z^2}\,\big)^{s+i}}{2^i z^i \left(\sqrt{1-4z^2}\,\right)^{s+1}}. \tag{2.11}$$

The number of walks of length N in $\mathcal{W}_i$ is well-known and easy to obtain by the observation that a walk in $\mathcal{W}_i$ must have $\frac{N+i}{2}$ upward steps (and $\frac{N-i}{2}$ downward steps), so that the result is

$$\binom{N}{\frac{N+i}{2}}. \tag{2.12}$$

Combining the last formulæ we find for the factorial moments $m_i^{(s)}(N)$

$$\binom{N}{\frac{N+i}{2}} m_i^{(s)}(N) = [z^N] M_i^{(s)}(z) = \frac{s!}{2^i} [z^{N+i}] \frac{\big(1-\sqrt{1-4z^2}\,\big)^{s+i}}{\left(\sqrt{1-4z^2}\,\right)^{s+1}} \tag{2.13}$$

where we denote by $[z^N]F(z)$ the coefficient of z^N in $F(z)$. Applying the binomial theorem the last quantity equals

$$= \frac{s!}{2^i} [z^{N+i}] \sum_{j=0}^{s+i} \binom{s+i}{j} (-1)^j \left(\sqrt{1-4z^2}\right)^{j-s-1}. \tag{2.14}$$

Using the binomial series we finally get

$$m_i^{(s)}(N) = \frac{s!}{2^i} \sum_{j=0}^{s+i} \binom{s+i}{j} (-1)^j \binom{\frac{j-s-1}{2}}{\frac{N+i}{2}} (-4)^{\frac{N+i}{2}} \bigg/ \binom{N}{\frac{N+i}{2}}. \tag{2.15}$$

In order to get an asymptotic result we substitute $z^2 = x$ in $M_i^{(s)}(z)$ and obtain

$$z^i M_i^{(s)}(z) = \frac{s!}{2^i} \frac{\left(1-\sqrt{1-4x}\,\right)^{s+i}}{\left(\sqrt{1-4x}\,\right)^{s+1}}. \tag{2.16}$$

The local expansion of this function close to its singularity $x = \frac{1}{4}$ reads

$$\begin{aligned} z^i M_i^{(s)}(z) = \frac{s!}{2^i} \Bigg\{ \frac{1}{(1-4x)^{\frac{s+1}{2}}} - (s+i)\frac{1}{(1-4x)^{\frac{s}{2}}} + \binom{s+i}{2} \frac{1}{(1-4x)^{\frac{s-1}{2}}} \\ + O\left(\frac{1}{(1-4x)^{\frac{s-2}{2}}}\right) \Bigg\}. \end{aligned} \tag{2.17}$$

Applying a transfer Lemma (compare [2]) to the right-hand-side we find

$$\begin{aligned} [x^n] z^i M_i^{(s)}(z) = \frac{s! 4^n}{2^i} \frac{n^{\frac{s-1}{2}}}{\Gamma\left(\frac{s+1}{2}\right)} \Bigg\{ 1 - \frac{s+i}{n^{1/2}} \frac{\Gamma\left(\frac{s+1}{2}\right)}{\Gamma\left(\frac{s}{2}\right)} \\ + \frac{s-1}{2n}\left(\frac{s-5}{4} + \binom{s+i}{2}\right) + O(n^{-3/2}) \Bigg\}. \end{aligned} \tag{2.18}$$

Observe that the last expression with $n = \frac{N+i}{2}$ gives an asymptotic formula for $[z^n] M_i^{(s)}(z)$.

The asymptotics of $\binom{N}{\frac{N+i}{2}}$ reads with the same substitution $n = \frac{N+i}{2}$

$$\binom{N}{\frac{N+i}{2}} = \binom{2n-i}{n} = \frac{1}{2^i} \frac{4^n}{\sqrt{\pi n}} \left(1 - \frac{1}{2n}\left(\frac{1}{4} + \binom{i}{2}\right) + O(\frac{1}{n^2})\right). \tag{2.19}$$

Dividing the two expansions we get the final asymptotic result

$$\begin{aligned} m_i^{(s)}(N) = \frac{s!}{\sqrt{\pi}} \frac{n^{s/2}}{\Gamma\left(\frac{s+1}{2}\right)} \Bigg(1 - \frac{s+i}{n^{1/2}} \frac{\Gamma\left(\frac{s+1}{2}\right)}{\Gamma\left(\frac{s}{2}\right)} \\ + \frac{1}{2n}\left(\frac{(s-3)^2}{4} + (s-1)\binom{s+i}{2} + \binom{i}{2}\right) + O(\frac{1}{n^2}) \Bigg). \end{aligned} \tag{2.20}$$

for $n = \frac{N+i}{2} \to \infty$, i, s fixed.

In the sequel we analyze the number of visits to the origin *for an arbitrary random walk starting at* 0, i.e. we consider the set

$$\mathcal{W} = \bigcup_{i \in \mathbb{Z}} \mathcal{W}_i. \tag{2.21}$$

The corresponding bivariate generating function is

$$\begin{aligned} W(z,u) &= W_0(z,u) + 2 \sum_{i \geq 1} W_i(z,u) \\ &= \frac{1}{1-u+u\sqrt{1-4z^2}} \frac{1+2z}{\sqrt{1-4z^2}}, \end{aligned} \tag{2.22}$$

whence we see

$$M^{(s)}(z) = s!\frac{\left(1-\sqrt{1-4z^2}\right)^s}{\left(\sqrt{1-4z^2}\right)^{s+1}} \cdot \frac{1+2z}{\sqrt{1-4z^2}} \tag{2.23}$$

for the s–th factorial moments multiplied by the numbers of paths.

Now even and odd powers of z are easily distinguished,

$$[z^{2n}]M^{(s)}(z) = [x^n]s!\frac{\left(1-\sqrt{1-4x}\right)^s}{\left(\sqrt{1-4x}\right)^{s+2}} = \frac{1}{2}[z^{2n+1}]M^{(s)}(z), \tag{2.24}$$

and the machinery from before works perfectly. Since, in order to obtain the moments, these numbers have to be divided by 2^{-2n} and 2^{-2n-1}, respectively, we see the coincidence between an even number and the consecutive odd one. So we find

$$m^{(s)}(n) = s!\sum_{j=0}^{s}\binom{s}{j}(-1)^{s-j}\binom{\frac{j}{2}+\lfloor\frac{n}{2}\rfloor}{\lfloor\frac{n}{2}\rfloor}. \tag{2.25}$$

3. Maximum Statistics

In this section we analyze a third problem with the generating functions approach, namely the number of times where a simple random walk reaches its maximum. This was recently studied in [6] by different methods.

We start from the observation that each random walk leading from $(0,0)$ to $(2n,0)$ can be generated in the following way. We take any non-positive random walk from $(0,0)$ to $(2n,0)$ (compare Figure 3), cut the first sojourn at any point different from its right end and glue together the cut off part with the end. The maxima of the produced random walk at time >0 correspond obviously to the returns (at time >0) to the x–axis of the original non-positive walk.

A nonpositive random walk with 4 returns to the origin

Fig. 3

The corresponding walk with 4 points at maximum height

Fig. 4

Another walk that reaches its maximum 4 times and is equivalent to the walk in Figure 3

Fig. 5

We observe that there are

$$2j\frac{1}{j}\binom{2j-2}{j-1} = 2\binom{2j-2}{j-1} \tag{3.1}$$

"marked" negative pathes between the origin and a first return point at time $2j$. The corresponding generating function is

$$\frac{2z^2u}{\sqrt{1-4z^2}}, \tag{3.2}$$

if the variable z counts the steps in the path and u counts again the number of returns. According to our combinatorial construction we have to multiply this term by the generating function of the set $\left(\mathcal{W}_{0,-}\right)^*$, i.e. by

$$\frac{1}{1 - W_{0,-}(z,u)} = \frac{1}{1 - uC(z^2)},$$

to get the generating function

$$1 + \frac{2z^2 u}{\sqrt{1-4z^2}} \frac{1}{1 - uC(z^2)}. \tag{3.3}$$

(The 1 counts the "empty" path of length 0.)

Therefore the corresponding generating function $M^{(s)}(z)$ reads

$$\begin{aligned} &\frac{s!\left(1-\sqrt{1-4z^2}\right)^{2s+1}}{4^s z^{2s}\sqrt{1-4z^2}} + s\frac{(s-1)!\left(1-\sqrt{1-4z^2}\right)^{2s-1}}{4^{s-1}z^{2s-2}\sqrt{1-4z^2}} \\ &\qquad = \frac{s!\left(1-\sqrt{1-4z^2}\right)^{2s}}{2^{2s-1}z^{2s}\sqrt{1-4z^2}}. \end{aligned} \tag{3.4}$$

From (3.4) the moments are found to be

$$\begin{aligned} m^{(s)}(n) &= \frac{s!}{2^{2s-1}} \sum_{j=0}^{2s} \binom{2s}{j} (-1)^{2s-j} \binom{\frac{2s-j-1}{2}}{n+s} (-4)^{n+s} \Big/ \binom{2n}{n} \\ &= 2s!4^n \sum_{j=0}^{2s} \binom{2s}{j} (-1)^j \binom{n + \frac{j-1}{2}}{n+s} \Big/ \binom{2n}{n}. \end{aligned} \tag{3.5}$$

We note that the corresponding maximum problem for non-negative paths is considerably harder, and was studied in [8].

References

1. K.G. Aneja and Kanwar Sen, *Maxima in a Random Walk and Related Rank Order Statistics*, Stud. Sci. Math. Hung. **7** (1972), 425–429.
2. P. Flajolet and A. Odlyzko, *Singularity Analysis of Generating Functions*, SIAM J. Disc. Math. **3** (1990), 216–240.
3. P. Flajolet and J. Vitter, *Average-Case Analysis of Algorithms and Data Structures*, Handbook of Theoretical Computer Science Vol. A "Algorithms and Complexity" (J. van Leeuwen, ed.), North-Holland, 1990, pp. 431–524.
4. A. Gupta and Kanwar Sen, *Distributions of Indices in Random Walk*, Mathematica Balkanica **7** (1977), 69–90.
5. W. Katzenbeisser and W. Panny, *A note on the higher moments of the random variable t associated with the number of returns of a simple random walk*, Advances in Applied Probability **18** (1986), 279–282.
6. W. Katzenbeisser and W. Panny, *On the number of times where a simple random walk reaches its maximum*, Journal of Applied Probability **29** (1992), 305–312.

7. A.W. Kemp, *The moments of the random variable for the number of returns of a simple random walk*, Advances in Applied Probability **19** (1987), 505–507.
8. R. Kemp, *On the number of deepest nodes in ordered trees*, Discrete Mathematics **81** (1990), 247–258.
9. P. Kirschenhofer and H. Prodinger, *Some Comments on the Higher Moments of the Number of Returns of a Simple Random Walk*, Advances in Applied Probability **26** (1994), 561–563.

TU Vienna
Department of Algebra and Discrete Mathematics
Wiedner Hauptstrasse 8–10
A-1040 Vienna
Austria
E-mail address: kirsch@rsmb.tuwien.ac.at, proding@rsmb.tuwien.ac.at