Prime numbers and Gaussian random walks

K. Bruce Erickson

Department of Mathematics
University of Washington
Seattle, WA 98195-4350

March 24, 2015

1 Introduction

Consider a symmetric aperiodic random walk $W = \{W_r\} = (W_r^1, W_s^2, W_r^3)$, $r = 0, 1, \ldots$, on the integer lattice $\mathbf{Z}^3$ starting at $W_0 = 0$. This means that the steps or increments $\{W_{r+1} - W_r\}_{r \geq 0}$ are independent with a common distribution that is invariant under reflection through the origin and that the support of this distribution generates the entire lattice $\mathbf{Z}^3$. We shall also assume that its one-step displacement has a finite second moment:

$$\mathbf{E}\,\|W_{r+1} - W_r\|^2 < \infty \tag{1}$$

The symmetry and (1) implies $\mathbf{E}\,(W_{r+1} - W_r) = 0$ (the vector). (We call describe any walk on $\mathbf{Z}^3$ satisfying all of the preceeding as "Gaussian".)

The walk induces a non-defective walk on the z-axis. More precisely put

$$\begin{aligned} T_1 &= \min\{r > 0 : W_r^1 = W_r^2 = 0\} \\ T_n &= \min(r > T_{n-1} : W_r^1 = W_r^2 = 0\},\ n \geq 2, \end{aligned}$$

$S_0 = 0$, and $S_n = W^3_{T(n)}, n \geq 1$. The two dimensional walk $\{(W_r^1, W_r^2)\}_{r \geq 0}$ returns to the origin infinitely often, a.s., so that $\mathbf{P}\{T_n < \infty\,\forall\, n\} = 1$. The Markov property and the linear structure of the z-axis imply that the sequence of increments $\{X_n\} \equiv \{S_n - S_{n-1}\}$ is stationary and mutually totally independent. Therefore the sequence $S = \{S_n\}_{n \geq 0}$, defines a (symmetric) random walk [1]on the one-dimensional lattice $\mathbf{Z}$. (With a slight abuse of terminology, we will call this sequence the walk induced by W on the z-axis.) The well known transience of W implies the transience of S which means that S cannot visit any *finite* set of integer points more than a finite number of times if at all.

Many infinite subsets of lattice points also have this transience property. For example, S will not hit the set of perfect squares, $\{1, 4, 9, \dots\}$, infinitely often. Let us denote the set of prime numbers by Q, and by $\overline{Q}$ the set $\{0\} \times \{0\} \times Q$. P. Erdös, H. P. Mckean, C. Stone, and B. Kochen established the interesting result that if W is the simple random walk, then, W returns to the set $\overline{Q}$ infinitely often. In terms of the induced walk S, $\mathbf{P}\{S_n \in Q \text{ i. o.}\} = 1$. From this result and an important invariance principle due to F. Spitzer, [6], one then gets the recurrence of $\overline{Q}$ for any aperiodic walk in $\mathbf{Z}^3$ with mean 0 displacement and finite second moments (wether or not the step distribution is symmetric).

Although these Gaussian walks W do hit $\overline{Q}$ infinitely often, it comes as a surprise that for some walks the set $\overline{Q}$ is transient for the sequence of successive high points on the z-axis. More precisely, let $M_n = \max\{S_j : j \leq n\}$, then we prove,under various additional assumptions, that $\mathbf{P}\{M_n \in Q \text{ i.o.}\} = 0$. For these random walks, with probability 1, there is a finite time such that the knowledge that W has reached a new prime point on the z-axis after that time implies that S has already hit a (non-prime) point higher up. (For the exact statement, see Theorem 1 in section 4.) One can certainly conjecture that this assertion remains valid for all Gaussian walks.

2 Characteristic function of the induced walk

Let $\{p_n\}$ denote the common distribution of the steps of S and let $\varphi(t)$ denote the characteristic function of $X_1 = S_1$, i.e.,

$$\varphi(t) = \mathbf{E}\, e^{itS_1} = \sum_n p_n e^{int}.$$

Our first item is the asymptotic formula

$$1 - \varphi(t) \simeq \frac{1}{\log(1/|t|)}, \qquad t \to 0,^{2} \tag{2}$$

[1]The common distribution of the X_k is very "heavy tailed"; $\mathbf{E}\,|X_1|^q = \infty$ for any $q > 0$.

I do not have a complete proof of the transience of Q for the z-axis ladders for general Gaussian walks. In order to make progress I need to impose some additional restrictions. The first one is not as serious as the second. (See the statement of Theorem 1 below.)

Our first additional condition is to assume that the two processes

$$\{(W_n^1, W_n^2)\}_{r\geq 0} \text{ and } \{W_s^3\}_{s\geq 0} \text{ are independent.} \tag{3}$$

Let $\gamma(t) = \mathbf{E}\,\exp\left(itW_1^3\right)$. Then

$$\begin{aligned} \varphi(t) = \mathbf{E}\, e^{itS_1} &= \sum_j \mathbf{P}\{T_1 = j\}\mathbf{E}\, e^{itW_j^3} \\ &= \mathbf{E}\,\gamma(t)^{T_1} \end{aligned} \tag{4}$$

because T_1 and W^3 are independent by (3). However, it is well known that

$$\mathbf{P}\{T_1 > n\} \simeq \frac{1}{\log n}, \qquad n \to \infty \tag{5}$$

(For the simple walk case, see [6], page 167. This result does not depend on the asumption (3). A proof will be given in Appendix 2.)

The formula (5) and an Abelian theorem, [2], page 447, $\rho = 1$ in (5.24)/(5.26), imply that

$$\frac{1 - \mathbf{E}\, s^{T_1}}{1-s} \simeq \frac{1}{(1-s)\log\left(\dfrac{1}{1-s}\right)}, \qquad s \uparrow 1. \tag{6}$$

By Taylor's formula

$$\gamma(t) = 1 - \frac{1}{2}m_2 t^2\left[1 + \mathrm{o}(1)\right], \qquad \text{for } |t| \to 0. \tag{7}$$

This and (6) and (4) (and the reality of γ by symmetry) imply that

$$\begin{aligned} 1 - \varphi(t) &= 1 - \big(1 - \gamma(t)\big)\frac{1 - \mathbf{E}\,\gamma(t)^{T_1}}{1 - \gamma(t)} \\ &\simeq \frac{1}{\log(1/|t|)}, \qquad |t| \to 0. \end{aligned}$$

This completes the proof of (2) in this case.

[2]We use the notation $a(t) \simeq b(t)$, $t \to t_0$, to mean that the $\lim_{t\to t_0} a(t)/b(t)$ is finite and strictly positive. The notation $a(t) \asymp b(t)$, $t \to t_0$, means the ratio of $a(t)$ to $b(t)$ is bounded strictly away from 0 and ∞ in a neighborhood of t_0.

3 Renewal function for the ladders

Let N, N^-, S_N, and S_{N^-} denote the first strict ascending/descending ladder epochs and positions. For example, $N^+ = \min\{n : S_n > 0\}$. Then, see [2], Chapter XVIII,

$$1 - s\varphi(t) = c(s)\big[1 - \mathbf{E}\{s^N e^{it S_N}\}\big]\big[1 - \mathbf{E}\{s^{N^-} e^{it S_{N^-}}\}\big] \tag{8}$$

where $\log c(s) = -\sum (s^n/n)\mathbf{P}(S_n = 0)$. But $-S_{N^-}$ has the same distribution as S_N so that on setting $s = 1$ we find that

$$1 - \varphi(t) = c(1)|1 - \chi(t)|^2$$

where χ is the characteristic funcion of S_N. From (2) it follows that

$$|1 - \chi(t)| \simeq \frac{1}{\sqrt{\log(1/|t|)}} \qquad t \to 0. \tag{9}$$

Now let

$$f(z) = \chi(-i \log z) = \sum_{n=1}^{\infty} f_n z^n, \qquad (f_n = \mathbf{P}\{S_N = n\}),$$

so that $\chi(t) = f(e^{it})$, and consider the function

$$D(z) = \begin{cases} \log\left\{[1 - f(z)]^2 \dfrac{\log(1-z)^{-1}}{z}\right\}, & \text{for } |z| \le 1,\ z \ne 1 \\ \log C_2^2 & \text{for } z = 1. \end{cases} \tag{10}$$

The distribution $\{f_n\}$ is aperiodic, that is, $\chi(t) = 1$ if and only if t is a multiple of 2π. This implies that $f(z) = 1$ and $|z| \le 1$ if and only if $z = 1$. (Note that the coefficients of f are non-negative and sum to 1.) The function

$$\frac{\log(1-z)^{-1}}{z} = \sum_{n=0}^{\infty} \frac{z^n}{n+1},$$

is also analytic in $|z| < 1$ and has positive derivatives at 0 of all orders. It is also continuous and non-zero at the boundary $|z| = 1$ except of course at the point $z = 1$. Finally, it never vanishes in $|z| \le 1$. It follows that D is analytic in $|z| < 1$, is continuous on $|z| \le 1$, and is real valued for z real and positive. Moreover

$$\Re\{D(z)\} = \log \left| [1 - f(z)]^2 \frac{\log(1-z)^{-1}}{z} \right| \tag{11}$$

Hence

$$D(z) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{e^{it} + z}{e^{it} - z} \log \left(|1 - \chi(t)|^2 \left| \frac{\log(1-e^{it})^{-1}}{e^{it}} \right| \right) dt, \quad |z| < 1. \tag{12}$$

It follows from this formula, an integral splitting argument, and the asymptotic formula $|(e^{it} \log(1-e^{it})^{-1}| \sim \log(1/|t|)$ for $t \to 0$, that

$$\lim_{s \uparrow 1} [1 - f(s)]^2 \log(1-s)^{-1} = \lim_{s \uparrow 1} e^{D(s)} = C_2^2 \tag{13}$$

and consequently

$$\frac{1 - f(s)}{1 - s} \sim \frac{C_2}{(1-s)\{\log(1-s)^{-1}\}^{1/2}}, \qquad s \uparrow 1. \tag{14}$$

Karmata's Tauberian theorem, [2], equation (5.26) on page 447 (with $\rho = 1$), shows that (14) implies

$$\begin{aligned} q_n = {} & \text{ coefficient on } s^n \text{ in LHS (14) } = \sum_{k>n} f_k \\ & = \mathbf{P}\{L_1 = S_N > n\} \simeq \frac{1}{\sqrt{\log n}}, \qquad n \to \infty^3 \end{aligned} \tag{15}$$

Now consider the sequence $\{N_k\}$ of strict ascending ladder times of the walk S defined inductively by

$$N_{k+1} = \min\{n > N_k : S_n > S_{N_k}\} \tag{16}$$

This coincides with the sequence of times of occurrence of new maximums of S. Anway if we put $L_0 = 0$ and $L_k = S_{N(k)}$, we get the strict ascending ladder process. Its increments $\{L_k - L_{k-1}\}_{k \ge 1}$ are mutually independent and have the same distribution $\{f_n\}$ defined earlier. Let

$$\psi_n = \sum_{j=0}^{\infty} \mathbf{P}\{L_j = n\} = \mathbf{P}\big[n \in \{L_0, L_1, \ldots, \}\big] \tag{17}$$

[3] I want to thank Professor Don Marshall for suggesting the possibility of going from the formula (9) to the formula (13) via the integral (12).

This is the renewal function for $\{L_k\}$ and its generating function satisfies:

$$\sum_{n\geq 0}\psi_n s^n = \frac{1}{1-f(s)} \tag{18}$$

Therefore, by (15) and Karamata Tauberians,

$$\psi_0+\psi_1+\cdots+\psi_n \simeq \sqrt{\log n} \tag{19}$$

Conjecture 1 *For the sequence $\{\psi_n\}$ as defined above in connection with a given Gaussian walk, we have*

$$\psi_n = \mathrm{O}\Big(\frac{1}{n\sqrt{\log n}}\Big) \tag{20}$$

Unfortunately, neither (15) nor (19) is sufficient to establish (20). On the other hand the renewal sequence $\{\psi_n\}$ has additional useful properties. In particular let $g(w)$ be the expected number of hits of W at the point $w = (w^1, w^2, w^3)$ and let v_n be the expected number of hits by S at n. Then $v_n = g(0,0,n)$. But for general aperiodic random walks on $\mathbf{Z}^3$ with mean 0 and finite second moments we know that $g(w) \asymp 1/\|w\|$ for large $\|w\|$. See [6], **P1**, page 308. Therefore

$$v_n \asymp \frac{1}{n}, \quad n\to\infty \tag{21}$$

A sequence $\{\psi_n\}$ is *asymptotically decreasing* if there are finite positive constants C, n_0 such that

$$\psi_m \leq C\psi_n \quad \text{for all } n \text{ and } m \text{ with } m > n \geq n_0 \tag{22}$$

(A bounded increasing sequence is asymptotically decreasing by this definition!)

Lemma 1. *If the renewal sequence $\{\psi_n\}$ is asymptotically decreasing, then* (20) *holds.*

Proof. For $n \geq 0$ we will prove in an appendix that for some number $c > 0$,

$$v_n = c\sum_{k=0}^{\infty}\psi_{k+n}\psi_k. \tag{23}$$

It follows from (23) and (21) and (19) that for some constants C_1, C_2, etc., and all k sufficiently large, we have

$$\begin{aligned}\frac{C_1}{k} \geq v_k &\geq C_2 \sum_{j=0}^{k} \psi_{j+k}\psi_j \\ &\geq C_3\psi_{2k} \sum_{j=0}^{k} \psi_j \geq C_4\psi_{2k}\sqrt{\log k}\end{aligned}$$

And this imples (20) for even n because $\rho_n = n\sqrt{\log n}$ is a regularly varying sequence (with exponent 1) so that $\rho_{2k} \simeq 2\rho_k$ as $k \to \infty$. Applying (22) again gives $\psi_{2k+1} \leq C\psi_{2k} = \mathrm{O}(k\sqrt{\log k})^{-1})$. Thus (20) holds for odd n also. □

Conjecture 2 *The sequence $\{\psi_n\}$ as defined above in connection with a given Gaussian walk is asymptotically decreasing.*

I have no example for which I can actually verify this.

4 Transience of the primes for the maxima

Clearly the random point sets determined by the sequences $\{L_k\}$ and $\{M_n\}$ coincide; the transience of the primes for $\{M_n\}$ is equivalent to the transience of the primes for the ascending ladder height process $\{L_n\}$.

Theorem 1. *Let $W = \{W_r\}$ be a Gaussian walk which satisfies* (3) *and for which the renewal function for the ladder process $\{L_k\}$ of the induced walk $\{S_n\}$ satisfies* (20). *If $\{M_n\}$ denotes the sequence of succesive maximum values of $\{S_n\}$, then $\mathbf{P}\{M_n$ is prime i. o. $\} = 0$.*

The proof is almost trivial. Let $x_k = x(k)$ denote the k-th prime. Then

$$x_k \sim k \log k$$

as $k \to \infty$. Hence if (20) is correct, then

$$\begin{aligned}\sum_{k=2}^{\infty} \psi_{x(k)} &= \sum_{k=2}^{\infty} \mathrm{O}\Big(\frac{1}{k\log k\,\sqrt{k\log(k\log k)}}\Big) \\ &= \sum_{k=2}^{\infty} \mathrm{O}\Big(\frac{1}{k\,(\log k)^{3/2}}\Big) < \infty \qquad (24)\end{aligned}$$

Clearly (24) is more than enough to imply $\mathbf{P}\{L_n \text{ is prime i. o.}\} = 0$ which, as noted above, is equivalent to the assertion of the theorem.

5 Appendix 1: Proof of (23)

Write

$$\chi(s,t) = \mathbf{E}\left(s^N e^{itS_N}\right)$$
$$\chi^-(s,t) = \mathbf{E}\left(s^{N^-} e^{itS_{N^-}}\right)$$

where N^- is the epoch of the first strict descending ladder height (i.e., the first entrance time to $(-\infty, 0)$ of $\{S_n\}$, when $S_0 = 0$). Then because S is symmetric, $(N^-, -S_{N^-})$ has the same distribution as (N, S_N), so that $\chi^-(s,t) = \chi(s,-t) = \overline{\chi(s,t)}$. The Wiener-Hopf facorization (8) then gives

$$\frac{1}{1 - s\varphi(t)} = \frac{c(s)}{[1-\chi(s,t)][1-\chi^-(s,t)]} = \frac{c(s)}{|1-\chi(s,t)|^2} \tag{25}$$

where $c(s) = \exp[\sum_{n\geq 1}(s^n/n)\mathbf{P}\{S_n = 0\}]$. See [2] page 605, formula (3.7). All terms in (25) are continuous 2π-periodic functions of t for $|s| < 1$. Denote the Fourier coefficients of $\{1 - s\varphi(t)\}^{-1}$ by $v_n^{(s)}$ and those of $\chi(s,t)$ and $\chi^-(s,t)$ by $\psi_k^{(s)}$ and $\psi_k^{(s)-}$, and by N_j the j-th strict ascending ladder epoch with $N_0 = S_0 = 0$. Recall that $\{(N_j - N_{j-1}, L_j - L_{j-1})\}_{j=1}^{\infty}$ are i.i.d. random vectors, where $L_j = S_{N(j)}$.

$$\begin{aligned}
v_n^{(s)} &= \sum_{k=0}^{\infty} s^k \mathbf{P}\{S_k = n\}, \quad n = 0 \pm 1 \pm 2, \ldots \\
\psi_k^{(s)} &= \begin{cases} 0 & \text{if } k \leq 0 \\ \sum_{j=0}^{\infty} \mathbf{E}\left\{s^{N_j};\, L_j = k\right\} & \text{for } k > 0 \end{cases} \\
\psi_k^{(s),-} &= \psi_{-k}^{(s)}, \quad \text{for all } k.
\end{aligned}$$

From this we see that all of these coefficients are non-negative and form absolutely convergent Fourier series (for fixed $0 \leq s < 1$). Expanding both sides of (25) in Fourier series and equating coeffiecents on like powers of e^{it}

and using symmetry of v_n, we get, for $n > 0$,

$$\begin{aligned} v_n^{(s)} = v_{-n}^{(s)} &= c(s) \sum_{k=-\infty}^{\infty} \psi_k^{(s)} \psi_{-n-k}^{(s),-} \\ &= c(s) \sum_{k=0}^{\infty} \psi_k^{(s)} \psi_{k+n}^{(s)} \end{aligned}$$

Also from the formulas for the coefficients, it is clear that $\psi_k^{(s)} \uparrow \psi_k$ and $v_n^{(s)} \uparrow v_n$ as $s \uparrow 1$. Therefore, from the last formula and ordinary monotone convergence for sums, we obtain

$$\begin{aligned} v_n = \lim_{s \uparrow 1} v_n^{(s)} &= \lim_{s \uparrow 1} c(s) \sum_{k=0}^{\infty} \psi_k^{(s)} \psi_{n+k}^{(s)} \\ &= c \sum_{k=0}^{\infty} \psi_k \psi_{n+k} \end{aligned}$$

where $c = c(1) = \exp(\sum_{n \geq 1} (1/n) \mathbf{P}\{S_n = 0\})$ which is non-zero and finite. This completes the proof of (23).

The formula (23) does not appear explicitly in either [2] or [6], nor in one other book I consulted. But it must be well known to workers in the field considering the elementary nature of the proof.

6 Appendix 2: Proof of (5)

Let $h = \min[r \geq 1 : \mathbf{P}\{W_r^1 = W_r^2 = 0\} > 0]$. Then for each k the values of T_k (the epoch of the k-th return to the z-axis of W) must be integer multiples of h. Moreover the sub-walk $\{(W_{rh}^1, W_{rh}^2)\}_{r \geq 0}$ is *strongly* aperiodic on $\mathbf{Z}^2$ and has mean 0 and finite second moments. See [6], **P1**, page 42.

From **P9** page 75 of [6] it follows that

$$u_n \equiv \mathbf{P}\{T_k = nh \text{ for some k }\} = \mathbf{P}\{W_{nh}^1 = W_{nh}^2 = 0\} \simeq \frac{1}{n}$$

so $\sum_{n=0}^{\infty} u_n s^n \simeq \log(1-s)^{-1}$ as $s \uparrow 1$. On the other hand

$$\sum_{n=0}^{\infty} u_n s^n = \frac{1}{1 - \mathbf{E}\, s^{T_1/h}}$$

Hence

$$\sum_{n=0}^{\infty} \mathbf{P}\{T_1/h > n\} = \frac{1 - \mathbf{E}\, s^{T_1/h}}{1-s} \simeq \frac{1}{(1-s)\log(1-s)^{-1}} \quad \text{as } s \uparrow 1$$

Karamata's Tauberian theorem now yields $\mathbf{P}\{T_1 > nh\} \simeq (\log n)^{-1}$ which certainly implies (5). See [2] equation (5.26), page 447.

References

[1] Benjamini, I., Pemantle, R., and Peres, Y. (1995) Martin capacity for markov chains. *Annals of Probab.* **23** 1332–1346.

[2] Feller, William. (1971). *An Introduction to Probability Theory and Its Applications*, II, 2nd ed. John Wiley & Sons, New York.

[3] Kochen, S. and Stone, C. (1964). A note on the Borel-Cantelli lemma. *Illinois J. Math.* **8** 248–251.

[4] McKean, H. P. Jr. (1961). A problem about prime numbers and the random walk I. *Illinois J. Math.* **5** 151.

[5] Ruzsa, I. Z. and Székely, G. J. (1982). Intersections of traces of random walks with fixed sets. *Annals of Probab.*, **10** 132–136.

[6] Spitzer, Frank. (1976). *Principles of Random Walk*, 2nd ed. Graduate Texts in Mathematics 34. Springer-Verlag, New York.