Some surprising properties of a conditioned simple random walk in two dimensions

by **Serguei Popov***

Abstract.—We define the two-dimensional conditioned simple random walk as the Doob's h-transform of the simple random walk with respect to its potential kernel, and discuss some of its properties. This is a very brief exposition of some of the topics presented in [12].

1 Introduction: simple random walks on integer lattices

This note is about the simple random walk[1] $(S_n, n \geq 0)$ on the integer lattice $\mathbb{Z}^d$ and we will pay special attention to the case $d = 2$. SRW is a discrete-time stochastic process defined as follows: if at a given time the walker is at $x \in \mathbb{Z}^d$, then at the next time moment it will be at one of x's $2d$ neighbours chosen uniformly at random, as shown on Figure 1. In other words, the probability that the walk follows a fixed length-n path of nearest-neighbour sites equals $(2d)^{-n}$. As a general fact, a random walk may be recurrent (i.e., almost surely it returns infinitely many times to its starting location) or transient (i.e., with positive probability it never returns to its starting location). An important result about SRWs on integer lattices is Pólya's classical theorem:

Theorem 1 ([10]).— Simple random walk in dimension d is recurrent for $d = 1, 2$ and transient for $d \geq 3$.

A well-known interpretation of this fact, attributed to Shizuo Kakutani, is: "a drunken man always returns home, but a drunken bird will eventually be lost". Still, despite recurrence, the drunken man's life is not so easy either: as we will see, it may take him *quite* some time to return home.

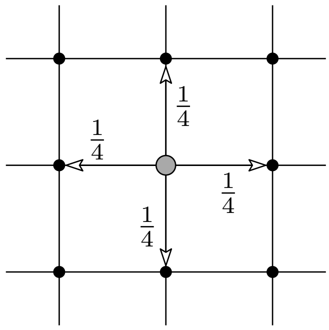

Figure 1.— Simple random walk in two dimensions.

Indeed, it is possible to obtain (see (10) and (6) below) that the probability that two-dimensional SRW gets more than distance n away from its starting position before revisiting it is approximately $(1.02937 + \frac{2}{\pi} \ln n)^{-1}$. While this probability does converge to zero as $n \to \infty$, it is important to notice how slow this convergence is. Here is a concrete example. Imagine a (two-dimensional) SRW taking place on the *galactic plane* of our galaxy, with the size of the walker's step being equal to 1m. What is the probability of reaching the galaxy's boundary before returning to the initial location? Since the walk is recurrent and the galaxy is enormous, one would expect this probability to be *extremely* small, correct? Now, let us do the calculations. The radius of the Milky Way galaxy is around 10^{21}m, and $(1.02937 + \frac{2}{\pi} \ln 10^{21})^{-1} \approx 0.031$, which is *surprisingly* large. Indeed, this means that the walker

[1] Also abbreviated as SRW.

* Centro de Matemática, University of Porto, Porto, Portugal.
Email: serguei.popov@fc.up.pt

would revisit the origin only around 30 times on average, before leaving the galaxy; this is not something one would normally expect from a *recurrent* process.

Incidentally, these sorts of facts explain why it is difficult to verify conjectures about two-dimensional SRW using computer simulations. (For example, imagine that one needs to estimate via simulations how long we will wait until the walk returns to the origin, say, a hundred times.)

As we will see shortly, the recurrence of d-dimensional SRW is related to the divergence of the series $\sum_{n=1}^{\infty} n^{-d/2}$. Notice that this series diverges if and only if $d \leq 2$, and for $d = 2$ it is the harmonic series that diverges quite slowly. This might explain why the two-dimensional case is, in some sense, *really* critical. It is always interesting to study critical cases – they frequently exhibit behaviours not observable away from criticality.

It is not our intention to present the proof of Theorem 1 in this note (one can find a modern proof of that result e.g. in [7]), but let us at least give a heuristic explanation of why it should be true. First, let us show that the number of visits to the origin is a.s. finite if and only if the *expected* number of visits to the origin is finite (note that this is something which is *not* true for general random variables). This is a useful fact, because, as it frequently happens, it is easier to control the expectation than the random variable itself.

Let $p_m(x, y) = \mathbb{P}_x[S_m = y]$ be the transition probability from x to y in m steps for the simple random walk in d dimensions. Let q_d be the probability that, starting at the origin, the walk eventually returns to the origin. If $q_d < 1$, then the total number of visits (counting the initial instance $S_0 = 0$ as a visit) is a geometric random variable with success probability $1 - q_d$, which has expectation $(1 - q_d)^{-1} < \infty$. If $q_d = 1$, then, clearly, the walk visits the origin infinitely many times a.s.. So, random walk is transient (i.e., $q_d < 1$) if and only if the *expected* number of visits to the origin is finite. This expected number equals[2]

$$\mathbb{E}_0 \sum_{k=0}^{\infty} \mathbf{1}\{S_k = 0\} = \sum_{k=0}^{\infty} \mathbb{E}_0 \mathbf{1}\{S_k = 0\} = \sum_{n=0}^{\infty} p_{2n}(0,0)$$

(observe that the walk can be at the starting point only after an even number of steps). We thus obtain that the recurrence of the walk is *equivalent* to

$$\sum_{n=0}^{\infty} p_{2n}(0,0) = \infty. \tag{1}$$

So, let us try to *understand* why Theorem 1 should hold. One can represent the d-dimensional simple random walk S as

$$S_n = X_1 + \cdots + X_n,$$

where $(X_k, k \geq 1)$ are i.i.d. random vectors, uniformly distributed on the set $\{\pm e_j, j = 1, \ldots, d\}$, where $e_1, \ldots, e_d$ is the canonical basis of $\mathbb{R}^d$. Since these random vectors are centered (expectation is equal to 0, component-wise), one can apply the (multivariate) Central Limit Theorem (CLT) to obtain that $S_n/\sqrt{n}$ converges in distribution to a (multivariate) centered Normal random vector with a diagonal covariance matrix. That is, it is reasonable to expect that S_n should be at distance of order $\sqrt{n}$ from the origin.

So, what about $p_{2n}(0,0)$? If $x, y \in \mathbb{Z}^d$ are two *even* sites[3] at distance of order at most $\sqrt{n}$ from the origin, then our CLT intuition tell us that $p_{2n}(0, x)$ and $p_{2n}(0, y)$ should be *comparable*, i.e., their ratio should be bounded away from 0 and ∞. In fact, this statement can be made rigorous by using the *local* Central Limit Theorem (e.g., Theorem 2.1.1 of [7]). Now, if there are $O(n^{d/2})$ sites where $p_{2n}(0, \cdot)$ are comparable, then the value of these probabilities (including $p_{2n}(0,0)$) should be of order $n^{-d/2}$. It remains only to observe that the series $\sum_{n=1}^{\infty} n^{-d/2}$ diverges only for $d = 1$ and 2 to convince oneself that Pólya's theorem indeed holds.

1.1 Potential kernel

Before starting the discussion on conditioned random walks, we need some technical preparations. Let us denote by

$$\tau_A = \min\{n \geq 0 : S_n \in A\}, \tag{2}$$

and

$$\tau_A^+ = \min\{n \geq 1 : S_n \in A\} \tag{3}$$

the entrance and the hitting times of a set A. Let $\partial A = \{x \in A : \exists y \in A^{\complement} \text{ such that } x \sim y\}$ be the boundary of $A \subset \mathbb{Z}^2$, and $\partial_e A = \partial A^{\complement}$ be its external boundary. Denote by $\mathsf{B}(x, r) = \{y : \|y - x\| \leq r\} \subset \mathbb{Z}^2$; $\mathsf{B}(r)$

[2] We can put the expectation inside the sum because of the Monotone Convergence Theorem.

[3] A site is called even if the sum of its coordinates is even; observe that the origin is even.

stands for $\mathsf{B}(0, r)$.

For transient random walks, a very important object is the Green's function, defined by $G(x, y) = \sum_{m=0}^{\infty} p_m(x, y)$, so that $G(x, y)$ is the expected number of visits to y starting from x. Its usefulness stems from the fact that $G(x, S_{n\wedge\tau_x})$ is a martingale, and martingales are really effective (see [9] for some interesting examples). However, for recurrent random walks that definition does not work since, as we know, the mean visit count equals infinity in that case. Fortunately, there is a way to amend that, essentially by considering the *difference* between mean visit counts starting from two different sites (of course, defining it properly). So, in two dimensions, let us define the *potential kernel* $a(\cdot)$ by

$$a(x) = \sum_{k=0}^{\infty} \big(\mathbb{P}_0[S_k = 0] - \mathbb{P}_x[S_k = 0]\big), \qquad (4)$$

where $x \in \mathbb{Z}^2$. By definition, it holds that $a(0) = 0$, and one can show that the above series converges and that the resulting value is strictly positive for all $x \neq 0$ (here and in the sequel we refer to Section 4.4 of [7]). Also, the function a is harmonic outside the origin, i.e.,

$$a(x) = \frac{1}{4}\sum_{y\sim x} a(y) \quad \text{for all } x \neq 0. \qquad (5)$$

It is possible to prove that, as $x \to \infty$,

$$a(x) = \frac{2}{\pi}\ln\|x\| + \gamma' + O(\|x\|^{-2}), \qquad (6)$$

where $\gamma' = \pi^{-1}(2\gamma + \ln 8)$, with $\gamma = 0.57721\ldots$ the Euler-Mascheroni constant, cf. Theorem 4.4.4 of [7]. Another important observation is that $a(x) = 1$ if x is a neighbor of the origin.

Observe that the harmonicity of a outside the origin immediately implies that the following result holds:

Proposition 2.— *The process $a(S_{k\wedge\tau_0})$ is a martingale.*

Besides, note that, due to (6),

$$a(x + y) - a(x) = O\big(\tfrac{\|y\|}{\|x\|}\big) \qquad (7)$$

for all $x, y \in \mathbb{Z}^2$ such that (say) $\|x\| > 2\|y\|$.

With some (slight) abuse of notation, we also consider the function

$$a(r) = \frac{2}{\pi}\ln r + \gamma'$$

of a *real* argument $r \geq 1$. Note that, in general, $a(x)$ need not be equal to $a(\|x\|)$, although they are of course quite close for large x. The advantage of using this notation is e.g. that, due to (6) and (7), we may write (for fixed x or at least x such that $2\|x\| \leq r$)

$$\sum_{y\in\partial\mathsf{B}(x,r)} \nu(y)a(y) = a(r) + O\big(\tfrac{\|x\|\vee 1}{r}\big) \qquad (8)$$

as $r \to \infty$, for *any* probability measure ν on $\partial\mathsf{B}(x, r)$.

We need the following result for the probability of going a long distance before revisiting the origin:

Lemma 3.— *Assume that $x \in \mathsf{B}(r)$ and $x \neq 0$. Then*

$$\mathbb{P}_x\big[\tau_{\partial\mathsf{B}(r)} < \tau_0^+\big] = \frac{a(x)}{a(r) + O(r^{-1})}, \qquad (9)$$

as $r \to \infty$.

Proof.— Indeed, use Proposition 2, and the optional stopping theorem to write (recall that $a(0) = 0$)

$$a(x) = \mathbb{P}_x\big[\tau_{\partial\mathsf{B}(r)} < \tau_0^+\big]\mathbb{E}_x\big(a(S_{\tau_{\partial\mathsf{B}(r)}}) \mid \tau_{\partial\mathsf{B}(r)} < \tau_0^+\big),$$

and then use (8). ∎

Note that Lemma 3 implies that (since, from the origin, on the next step the walk will go to a neighbor of the origin where the potential kernel equals 1)

$$\mathbb{P}_0[\tau_{\partial\mathsf{B}(r)} < \tau_0^+] = \frac{1}{a(r) + O(r^{-1})}. \qquad (10)$$

With the technical facts established above, we are now ready to pass to the main subject of this note.

2 Random walk conditioned on never hitting the origin

2.1 Doob's h-transforms

Let us start with a one-dimensional example. Let $(S_n, n \geq 0)$ be the simple random walk in dimension 1. It is well known that for any $0 < x < R$

$$\mathbb{P}_x[\tau_R < \tau_0] = \frac{x}{R} \qquad (11)$$

— this is the solution of Gambler's Ruin Problem for players of equal strength, and note also that it is straightforward to obtain it from the optional stopping theorem using the fact that S_n is a martingale. Now, how will the walk behave if we *condition* it to reach R before reaching the origin? Using (11), we write

$$\begin{aligned}\mathbb{P}_x[S_1 = x+1 \mid \tau_R < \tau_0] &= \frac{\mathbb{P}_x[S_1 = x+1, \tau_R < \tau_0]}{\mathbb{P}_x[\tau_R < \tau_0]}\\ &= \frac{\mathbb{P}_x[S_1 = x+1]\mathbb{P}_{x+1}[\tau_R < \tau_0]}{\mathbb{P}_x[\tau_R < \tau_0]}\end{aligned}$$

$$= \frac{\frac{1}{2}\times\frac{x+1}{R}}{\frac{x}{R}} = \frac{1}{2}\times\frac{x+1}{x},$$

which also implies that

$$\mathbb{P}_x[S_1 = x-1 \mid \tau_R < \tau_0] = \frac{1}{2}\times\frac{x-1}{x}.$$

The above calculation does not yet formally show that the conditioned walk is a Markov process (we would have needed to condition on the whole history), but let us forget about that for now, and examine the new transition probabilities we just obtained,

$$\hat{p}(x, x-1) = \frac{1}{2}\times\frac{x-1}{x}$$

and

$$\hat{p}(x, x+1) = \frac{1}{2}\times\frac{x+1}{x}.$$

First, it is remarkable that they do not depend on R, which *suggests* that we can just send R to infinity and obtain "the random walk conditioned on never returning to the origin". Secondly, look at the arguments of $\hat{p}$'s and the second fraction in the right-hand sides: these new transition probabilities are related to the old ones (which are $p(x,y) = 1/2$ for $x \sim y$) in a special way:

$$\hat{p}(x,y) = p(x,y)\times\frac{h(y)}{h(x)} \qquad (12)$$

with $h(x) = |x|$ (soon it will be clear why do we prefer to keep the function nonnegative). What is special about this function h is that it is harmonic outside the origin, so that $h(S_{n\wedge\tau_0})$ is a martingale. It is *precisely* this fact that permitted us to obtain (11) with the help of the optional stopping theorem.

Keeping the above discussion in mind, we pass to a more general setup. Consider a countable Markov chain on a state space Σ, and let $A \subset \Sigma$ be finite. Let $h : \Sigma \to \mathbb{R}_+$ be a nonnegative function which is zero on A and strictly positive and harmonic outside A, i.e., $h(x) = \sum_y p(x,y)h(y)$ for all $x \notin A$. We assume also that $h(x) \to \infty$ as $x \to \infty$; this implies that the Markov chain is recurrent (this follows e.g. from Theorem 2.4 of [12]).

Definition 4.— The new Markov chain with transition probabilities defined as in (12) is called *Doob's h-transform* of the original Markov chain with respect to h.

Observe that the harmonicity of h implies that $\hat{p}$'s are transition probabilities indeed:

$$\sum_y \hat{p}(x,y) = \frac{1}{h(x)}\sum_y p(x,y)h(y) = \frac{1}{h(x)}\times h(x) = 1.$$

To the best of the author's knowledge, this kind of object first appeared in [4], in the continuous-space-and-time context. Further information can be found e.g. in [1, 8, 14], and the book [5] provides a systematic treatment of the subject in full generality.

Note the following simple calculation: for any $x \notin A \cup \partial_e A$, we have (note that $h(y) \neq 0$ for all $y \sim x$ then)

$$\begin{aligned}\mathbb{E}_x \frac{1}{h(\hat{X}_1)} &= \sum_{y\sim x}\hat{p}(x,y)\frac{1}{h(y)} \\ &= \sum_{y\sim x} p(x,y)\frac{h(y)}{h(x)}\frac{1}{h(y)} \\ &= \frac{1}{h(x)}\sum_{y\sim x} p(x,y) \\ &= \frac{1}{h(x)},\end{aligned}$$

which implies the following:

Proposition 5.— The process $1/h(\hat{X}_{n\wedge\tau_{A\cup\partial_e A}})$ is a martingale and the Markov chain $\hat{X}$ is transient.

(The last statement follows from Theorem 2.5 of [12] since $1/h(x) \to 0$ as $x \to \infty$.)

Now let us try to get an idea about what the h-transformed chain really does. For technical reasons, let us make another assumption: there exists $c > 0$ such that $|h(x) - h(y)| \le c$ for all $x \sim y$ (for general Markov chains, $x \sim y$ means $p(x,y) + p(y,x) > 0$).

For $R > 0$, let us define

$$\Lambda_R = \{x \in \Sigma : h(x) \le R\};$$

under the previous assumptions, Λ_R is finite for any R. Note that the optional stopping theorem implies that, for $x_0 \in \Lambda_R \setminus A$

$$h(x_0) = \mathbb{P}_{x_0}[\tau_{\Lambda_R^\complement} < \tau_A]\mathbb{E}_{x_0}\big(h(X_{\tau_{\Lambda_R^\complement}}) \mid \tau_{\Lambda_R^\complement} < \tau_A\big),$$

(recall that $\mathbb{E}_{x_0}(h(X_{\tau_A}) \mid \tau_A < \tau_{\Lambda_R^\complement}) = 0$ because h vanishes on A) and, since the second factor in the preceding display is in $[R, R+c]$, we have

$$\mathbb{P}_{x_0}[\tau_{\Lambda_R^\complement} < \tau_A] = \frac{h(x_0)}{R}\big(1 + O(R^{-1})\big). \qquad (13)$$

Then, we consider another countable Markov chain $\hat{X}$ on the state space $\Sigma \setminus A$ with transition probabilities $\hat{p}(\cdot,\cdot)$ defined as in (12) for $x \notin A$.

Now, consider a *path* $\wp = (x_0, \ldots, x_{n-1}, x_n)$, where $x_0, \ldots, x_{n-1} \in \Lambda_R \setminus A$ and $x_n \in \Sigma \setminus \Lambda_R$ (see Figure 2 (next page); here, *path* is simply a sequence of neighbouring sites; in particular, it need not be self-avoiding). The original weight of that path (i.e., the probability that the Markov chain X follows it starting from x_0) is

$$P_{\wp} = p(x_0, x_1)p(x_1, x_2) \dots p(x_{n-1}, x_n),$$

and the weight of the path for the new Markov chain $\widehat{X}$ will be

$$\begin{aligned}\widehat{P}_{\wp} &= p(x_0, x_1)\frac{h(x_1)}{h(x_0)} \dots p(x_{n-1}, x_n)\frac{h(x_n)}{h(x_{n-1})} \\ &= p(x_0, x_1) \dots p(x_{n-1}, x_n)\frac{h(x_n)}{h(x_0)} \\ &= P_{\wp}\frac{h(x_n)}{h(x_0)}. \end{aligned} \tag{14}$$

Here comes the key observation: the last term in (2.1) actually equals $(1 + O(R^{-1}))R/h(x_0)$, that is, it is *almost* inverse of the expression in the right-hand side of (13). So, we have

$$\widehat{P}_{\wp} = \frac{P_{\wp}}{\mathbb{P}_{x_0}[\tau_{\Lambda_R^{\complement}} < \tau_A]}\big(1 + O(R^{-1})\big),$$

that is, the probability that the $\widehat{X}$ chain follows a path is almost the *conditional* probability that that the original chain X follows that path, under the condition that it goes out of Λ_R before reaching A (and the relative error goes to 0 as $R \to \infty$). Now, the (decreasing) sequence of events $\{\tau_{\Lambda_R^{\complement}} < \tau_A\}$ converges to $\{\tau_A = \infty\}$ as $R \to \infty$. Therefore, we can rightfully call $\widehat{X}$ the Markov chain *conditioned* on never reaching A, even though the probability of the latter event equals zero.

2.2 The conditioned SRW in two dimensions

Now, we will consider the two-dimensional SRW conditioned on never entering the origin, which is the Doob's h-transform of (unconditional) two-dimensional SRW with respect to its potential kernel a. It turns out that the conditioned walk $\widehat{S}$ is quite an interesting object on its own — some of its surprising properties are described later in this section.

By (5), the potential kernel a can play the role of the h (the one of the previous section), so let us define another random walk $(\widehat{S}_n, n \geq 0)$ on $\mathbb{Z}^2 \setminus \{0\}$ in the following way: its transition probability matrix equals

$$\hat{p}(x, y) = \begin{cases} \dfrac{a(y)}{4a(x)}, & \text{if } x \sim y, x \neq 0, \\ 0, & \text{otherwise.} \end{cases} \tag{15}$$

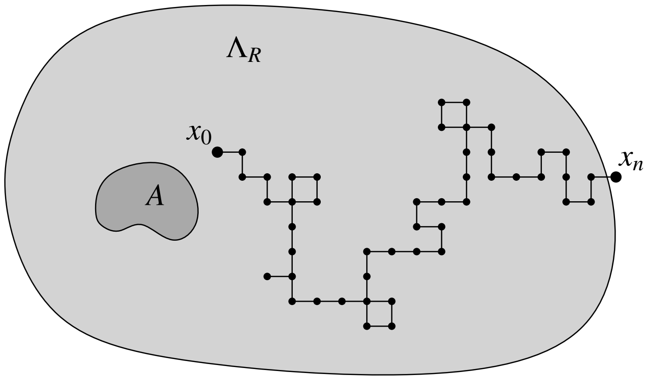

Figure 2.— Comparing the weights of the path..

The discussion of the previous section then means that the random walk $\widehat{S}$ is the Doob h-transform of the simple random walk, under the condition of not hitting the origin. Let $\hat{\tau}$ and $\hat{\tau}^+$ be the entrance and the hitting times for $\widehat{S}$; they are defined as in (1.1) and (1.1), only with $\widehat{S}$. We summarize the basic properties of the random walk $\widehat{S}$ in the following:

Proposition 6.— The following statements hold:

(i) The walk $\widehat{S}$ is reversible, with the reversible measure $\mu(x) = a^2(x)$.

(ii) In fact, it can be represented as a random walk on the two-dimensional lattice with the set of conductances $\big(a(x)a(y), x, y \in \mathbb{Z}^2, x \sim y\big)$.

(iii) The process $1/a(\widehat{S}_{n\wedge\hat{\tau}_{\mathcal{N}}})$ is a martingale.[4]

(iv) The walk $\widehat{S}$ is transient.

Proof.— Indeed, for (i) and (ii) note that

$$a^2(x)\hat{p}(x, y) = \frac{a(x)a(y)}{4} = a^2(y)\hat{p}(y, x)$$

for all adjacent $x, y \in \mathbb{Z}^2 \setminus \{0\}$, and, since a is harmonic outside the origin,

$$\frac{a(x)a(y)}{\sum_{z\sim x} a(x)a(z)} = \frac{a(y)}{4\sum_{z\sim x}\frac{1}{4}a(z)} = \frac{a(y)}{4a(x)} = \hat{p}(x, y).$$

Items (iii) and (iv) are Proposition 5. ∎

The Green's function of the conditioned walk (which is transient) is defined in the usual way: for $x, y \in \mathbb{Z}^2 \setminus \{0\}$

$$\widehat{G}(x, y) = \mathbb{E}_x \sum_{k=0}^{\infty} \mathbf{1}\{\widehat{S}_k = y\}. \tag{16}$$

[4] $\mathcal{N}$ is the set of the four neighbours of the origin.

One can calculate this function in terms of the potential kernel a (this is Theorem 1.1 of [11]): for all $x, y \in \mathbb{Z}^2 \setminus \{0\}$ it holds that

$$\widehat{G}(x,y) = \frac{a(y)}{a(x)}\big(a(x) + a(y) - a(x-y)\big). \qquad (17)$$

Now, we are able to obtain exact expressions (in terms of Green's function) for one-site escape probabilities, and probabilities of (not) hitting a given site. Indeed, since, under $\mathbb{P}_x$, the number of visits (counting the one at time 0) to x is geometric with success probability $\mathbb{P}_x[\hat{\tau}_x = \infty]$, we have

$$\mathbb{P}_x[\hat{\tau}_x^+ < \infty] = 1 - \frac{1}{\widehat{G}(x,x)} = 1 - \frac{1}{2a(x)} \qquad (18)$$

for $x \neq 0$. Also, since

$$\widehat{G}(x,y) = \mathbb{P}_x[\hat{\tau}_y^+ < \infty]\widehat{G}(y,y) \quad \text{for } x \neq y,\ x, y \neq 0$$

(one needs to go to y first, to start counting visits there), we have

$$\mathbb{P}_x[\hat{\tau}_y < \infty] = \frac{\widehat{G}(x,y)}{\widehat{G}(y,y)} = \frac{a(x) + a(y) - a(x-y)}{2a(x)}. \qquad (19)$$

Let us also observe that (19) (together with (6)) implies the following surprising[5] fact: for any $x \neq 0$,

$$\lim_{y\to\infty} \mathbb{P}_x[\hat{\tau}_y < \infty] = \frac{1}{2}. \qquad (20)$$

It is interesting to note that this fact permits us to obtain a criterion for recurrence of a set with respect to the conditioned walk. We say that *a set* is recurrent with respect to a (transient) Markov chain, if it is visited infinitely many times almost surely; a set is called *transient*, if it is visited only finitely many times almost surely (note that, trivially, for a transient Markov chain every finite set is transient). Recall that, for SRW in dimensions $d \geq 3$, the characterization of recurrent/transient sets is provided by *Wiener's criterion* (see e.g. Corollary 6.5.9 of [7]) formulated in terms of capacities of intersections of the set with exponentially growing annuli. Although this result does provide a complete classification, it may be difficult to apply it in practice, because it is not always trivial to calculate (even to *estimate*) capacities. Now, it turns out that for the conditioned two-dimensional walk $\widehat{S}$ the characterization of recurrent and transient sets is particularly simple:

Theorem 7 ([6]).— A set $A \subset \mathbb{Z}^2$ is recurrent with respect to $\widehat{S}$ if and only if A is infinite.

Proof.— We only need to prove that every infinite subset of $\mathbb{Z}^d$ is recurrent for $\widehat{S}$. As mentioned before, this is basically a consequence of (20). Indeed, let $\widehat{S}_0 = x_0$; since A is infinite, by (20) one can find $y_0 \in A$ and R_0 such that $\{x_0, y_0\} \subset \mathsf{B}(R_0)$ and

$$\mathbb{P}_{x_0}\big[\hat{\tau}_{y_0} < \hat{\tau}_{\partial\mathsf{B}(R_0)}\big] \geq \frac{1}{3}.$$

Then, for any $x_1 \in \partial\mathsf{B}(R_0)$, we can find $y_1 \in A$ and $R_1 > R_0$ such that $y_1 \in \mathsf{B}(R_1) \setminus \mathsf{B}(R_0)$ and

$$\mathbb{P}_{x_1}\big[\hat{\tau}_{y_1} < \hat{\tau}_{\partial\mathsf{B}(R_1)}\big] \geq \frac{1}{3}.$$

Continuing in this way, we can construct a sequence $R_0 < R_1 < R_2 < \ldots$ (depending on the set A) such that, for each $k \geq 0$, the walk $\widehat{S}$ hits A on its way from $\partial\mathsf{B}(R_k)$ to $\partial\mathsf{B}(R_{k+1})$ with probability at least $\frac{1}{3}$, regardless of the past. This clearly implies that A is a recurrent set. ∎

Next, we state an even more surprising result, which attests the "fractal" behaviour of $\widehat{S}$'s trajectories. For a set $T \subset \mathbb{Z}_+$ (thought of as a set of time moments) let

$$\widehat{S}_T = \bigcup_{m\in T}\{\widehat{S}_m\}$$

be the *range* of the walk $\widehat{S}$ with respect to that set, i.e., it is made of sites that are visited by $\widehat{S}$ over T. For simplicity, we assume in the following that the walk $\widehat{S}$ starts at a fixed neighbour x_0 of the origin, and we write $\mathbb{P}$ for $\mathbb{P}_{x_0}$. For a nonempty and finite set $A \subset \mathbb{Z}^2$, let us consider the random variable

$$\mathscr{R}(A) = \frac{\big|A \cap \widehat{S}_{[0,\infty)}\big|}{|A|};$$

that is, $\mathscr{R}(A)$ is the proportion of visited sites of A by the walk $\widehat{S}$ (and, therefore, $1 - \mathscr{R}(A)$ is the proportion of unvisited sites of A). A natural question is: how should $\mathscr{R}(A)$ behave for "large" sets? By (20), in average approximately half of A should be covered, i.e., $\mathbb{E}\mathscr{R}(A)$ should be close to 1/2. Surprisingly, for a "typical" large set (e.g., a disk, a rectangle, a segment) the random variable $\mathscr{R}(A)$ does not concentrate, and instead the following holds: the proportion of visited sites is a *random variable* which is close in distribution to Uniform[0, 1]. The paper [6] contains the corresponding results in greater generality, but here we content ourselves in stating the result for a particular case of a large disk which does not "touch" the origin:

[5] We know that the conditioned walk is transient and there is "a lot of space" in $\mathbb{Z}^2$, so one would rather expect that the probability to eventually hit a very distant site would go to zero.

THEOREM 8.— Let $D \subset \mathbb{R}^2$ be a closed disk such that $0 \notin D$, and denote $D_n = nD \cap \mathbb{Z}^2$. Then, for all $s \in [0,1]$, we have, with positive constant c_1 depending only on D,

$$\left|\mathbb{P}[\mathscr{R}(D_n) \le s] - s\right| \le c_1 \left(\frac{\ln \ln n}{\ln n}\right)^{1/3}. \qquad (21)$$

The last result we mention here is a quantitative assessment of how fast the transience of $\widehat{S}$ happens. Let us define the *future minima process*

$$M_n := \min_{m \ge n} |\widehat{S}_m|;$$

so far, we only know that $M_n \to \infty$ a.s. by transience. It is possible to obtain some finer asymptotic properties of M_n:

THEOREM 9 ([13]).— For every $0 < \delta < \frac{1}{2}$ we have, almost surely,

$$M_n \le n^\delta \text{ i.o.} \quad \text{but} \quad M_n \ge \frac{\sqrt{n}}{\ln^\delta n} \text{ i.o.}$$

The above result means that the transience of the conditioned SRW is "very irregular": sometimes it goes to infinity in the usual "diffusive" way, but sometimes slows down quite dramatically.

As a concluding remark, we also mention that in [2] even finer results were obtained for the continuous analogue of the conditioned SRW (which is the Brownian motion conditioned on never hitting the unit disk — one is then able to use the fact that it is radially symmetric, something that does not hold in the discrete setting). We are now working on extending the results of [2] to the discrete case.

References

[1] CHUNG KAI LAI, J.B. WALSH (2005) *Markov Processes, Brownian Motion, and Time Symmetry*. Springer, New York.

[2] O. COLLIN, F. COMETS (2022) Rate of escape of conditioned Brownian motion. *Electr. J. Probab.* **27** (31).

[3] F. COMETS, S. POPOV, M. VACHKOVSKAIA (2016) Two-dimensional random interlacements and late points for random walks. *Commun. Math. Phys.* **343**, 129–164.

[4] J.L. DOOB (1957) Conditional Brownian motion and the boundary limits of harmonic functions. *Bull. Soc. Math. France*, **85**, 431–458.

[5] J.L. DOOB (1984) *Classical Potential Theory and its Probabilistic Counterpart.* Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 262. Springer-Verlag, New York.

[6] N. GANTERT, S. POPOV, M. VACHKOVSKAIA (2019) On the range of a two-dimensional conditioned simple random walk. *Ann. Henri Lebesgue*, **2**, 349–368.

[7] G. LAWLER, V. LIMIC (2010) *Random Walk: A Modern Introduction.* Cambridge Studies in Advanced Mathematics, **123**. Cambridge University Press, Cambridge.

[8] D.A. LEVIN, Y. PERES (2017) *Markov Chains and Mixing Times*. American Mathematical Society, Providence.

[9] Y. PERES (2009) The unreasonable effectiveness of martingales. In: *Proceedings of the Twentieth Annual ACM-SIAM Symposium on Discrete Algorithms*. SIAM, Philadelphia.

[10] G. PÓLYA (1921) Über eine Aufgabe der Wahrscheinlichkeitsrechnung betreffend die Irrfahrt im Straßennetz. *Math. Ann.*, **84** (1–2), 149–160.

[11] S. POPOV (2021) Conditioned two-dimensional simple random walk: Green's function and harmonic measure. *J. Theor. Probab.* **34**, 418–437.

[12] S. POPOV (2021) *Two-dimensional Random Walk: From Path Counting to Random Interlacements.* Cambridge University Press, Cambridge.

[13] S. POPOV, L. ROLLA, D. UNGARETTI (2020) Transience of conditioned walks on the plane: encounters and speed of escape. *Electr. J. Probab.* **25** (52).

[14] W. WOESS (2009) *Denumerable Markov chains*. EMS Textbooks in Mathematics. European Mathematical Society (EMS), Zürich.