Residence Time Near an Absorbing Set

J. Randon-Furling

SAMM (EA 4543), Université Paris-1 Panthéon-Sorbonne, 90 rue de Tolbiac, 75013 Paris, France

S. Redner

Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, New Mexico 87501, USA

Abstract.
    We determine how long a diffusing particle spends in a given spatial range before it dies at an absorbing boundary. In one dimension, for a particle that starts at x_0 and is absorbed at $x = 0$, the average residence time in the range $[x, x + dx]$ is $T(x) = \frac{x}{D} dx$ for $x < x_0$ and $\frac{x_0}{D} dx$ for $x > x_0$, where D is the diffusion coefficient. We extend our approach to biased diffusion, to a particle confined to a finite interval, and to general spatial dimensions. We use the generating function technique to derive parallel results for the average residence time of the one-dimensional symmetric nearest-neighbor random walk that starts at $x_0 = 1$ and is absorbed at $x = 0$. We also determine the distribution of times at which the random walk first revisits $x = 1$ before being absorbed.

1. Introduction

Suppose that a diffusing particle in one dimension starts at $x_0 > 0$ and is absorbed, or equivalently, dies, when $x = 0$ is reached. One classic property of diffusion is that the particle is sure to eventually reach the origin, but the average time for this event to occur is infinite [1–4]. This dichotomy between certain return and an infinite return time is the source of rich phenomenology and counter-intuitive phenomena about the statistical properties of diffusion. Another important feature of diffusion is the shape of its trajectory in space time (Fig. 1). In the interesting case of $T \gg x_0^2/D$, the particle wanders over a large spatial range before its eventual demise at $x = 0$ at time T. A trajectory that starts stays above $x = 0$ until absorption at time T is known as a *Brownian excursion* when x_0 is also equal to zero [5].

    Two basic questions about an excursion are: (i) What is its shape [6,7]? (ii) How much time does the excursion spend in the range $[x, x + dx]$ before being absorbed? We term this quantity as the residence time. The latter question has been addressed in the mathematics literature by local-time theorems [4,8–10] that specify the time that a Brownian particle spends in the region $[x, x+dx]$ before being absorbed when the origin is reached. When $x < x_0$, the distribution of this residence time was shown to be related
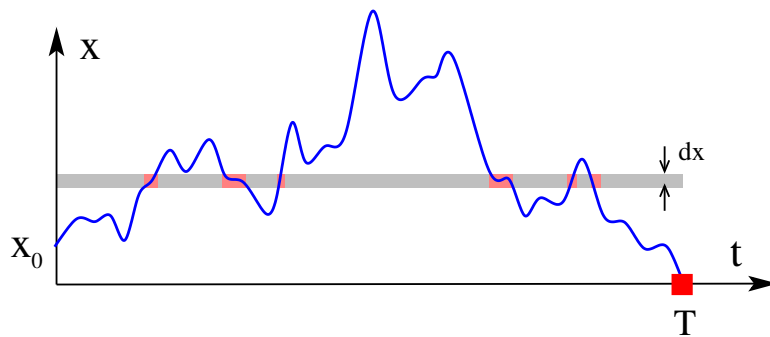
Figure 1. Schematic trajectory of a diffusing particle in space time that starts at x_0 and is absorbed at time T (square). The time that the particle spends in $[x, x + dx]$ is indicated by the colored segments.

to the distribution of the radial distance of a two-dimensional Brownian motion [9, 10]. If the particle wanders in a finite domain with reflection at the domain boundary and absorption at a given point (or points) within the domain, the residence time at each site is related to the first-passage time to the absorbing set [11–13]. This general formalism allows one to compute both the residence time and the distribution of residence times at a given location.

While the consequences of local-time theorems are profound, the mathematical literature is sometimes presented in a style that is not readily accessible to the community of physicists who study random walks, and some of the results derived in Refs. [11–13] are extremely general in their formulation. In this work, we investigate residence-time phenomena for both continuum diffusion and the discrete random walk by using ideas and approaches from first-passage processes. We focus on cases where the particle is eventually absorbed at a specified boundary (*e.g.*, one specific side of an interval) and/or starts close to this boundary. Such cases do not seem to have been given explicit solutions in the literature.

In Sec. 2.1, we first derive the residence time within the interval $[x, x + dx]$ in continuum diffusion, by solving the relevant diffusion equation. We extend this approach to: (a) biased diffusion on the semi-infinite line (Sec. 2.2), and (b) unbiased diffusion in a finite domain $[0, L]$, with the condition that the particle is eventually absorbed at $x = 0$ (Sec. 2.3). We then determine the residence time in general spatial dimension in the domain exterior to a small absorbing sphere of radius a (Sec. 2.4).

We then turn to the corresponding discrete system of a nearest-neighbor symmetric random walk that starts at $x_0 = 1$ and is absorbed when $x = 0$ is reached. The analog of the residence time is the number of times that the walk revisits a given point x before it dies at $x = 0$. We write the total number of steps of this random walk—which is necessarily odd—as $2n + 1$. In Sec. 3, we use generating function methods to derive the number of visits to a given site for a symmetric random walk on the semi-infinite line. For fixed n, we will show that the number of times that $x = 1$ is revisited equals $3n/(n + 2)$. By averaging this quantity over the number of steps of the walk, the

average number of revisits to $x = 1$ equals 2. Moreover, the average number of times that the random walk visits a site at $x > 1$ equals 4 for any $x > 1$. These results match those found in continuum diffusion in the analogous geometry. Finally, in Sec. 4, we determine the time when a walk first revisits $x = 1$, when it starts at $x = 1$ and is eventually absorbed at $x = 0$. We give some concluding comments in Sec. 5.

2. Residence Time for Diffusion

2.1. Isotropic diffusion on the semi-infinite line

Consider a particle with diffusion coefficient D that starts at x_0 and is absorbed when $x = 0$ is reached. For such a particle, the image method gives the probability density of the particle in the region $x > 0$ as [14, 15]

$$P(x, t) = \frac{1}{\sqrt{4\pi Dt}} \left[e^{-(x-x_0)^2/4Dt} - e^{-(x+x_0)^2/4Dt} \right]. \tag{1}$$

The time $T(x)$ that the particle spends in the range $[x, x + dx]$ before being absorbed at $x = 0$ is simply the integral of the probability density over all time times dx (see Refs. [2, 3, 16–18] for related approaches). Performing this integral, with $P(x, t)$ from Eq. (1), the residence time $T(x)$ is given by

$$T(x) = dx \int_0^\infty dt\, P(x, t) = \begin{cases} \dfrac{x}{D}\, dx & x < x_0\,, \\[2mm] \dfrac{x_0}{D}\, dx & x > x_0\,. \end{cases} \tag{2}$$

To appreciate this result, we present simulations for a nearest-neighbor random walk that starts at $x_0 = 1$ and at $x_0 = 10$ in Fig. 2. As a function of the number of walks in the ensemble, $T(x)$ slowly converges to the asymptotic time-independent value in Eq. (2). A curious feature of this residence-time data is that it becomes erratic for large x, as shown in Figs. 2(c) and (d). We can understand the origin of these large fluctuations by the following rough argument: for a diffusing particle that starts at x_0, the probability $S(t)$ that it survives until time t is $S(t) = \mathrm{erf}(x_0/\sqrt{4Dt}) \simeq x_0/\sqrt{4Dt}$ for $t \to \infty$ [15]. For $\mathcal{M}$ random walks, we estimate the longest lived of them by the extreme-statistics criterion $S(t_{\max}) \simeq 1/\mathcal{M}$ [19, 20], which states that one of out $\mathcal{M}$ walks survives until at least time $t_{\max}$. This criterion gives $t_{\max} \simeq (\mathcal{M}x_0)^2/4D$. Correspondingly, the maximal range reached by an ensemble of $\mathcal{M}$ random walks is, roughly, $x_{\max} \sim \sqrt{Dt_{\max}} \sim \mathcal{M}x_0$.

We now use this estimate to determine the nature of the large-x fluctuations in Figs. 2(c) and (d). To obtain an accuracy of, say, 10%, in $N(x)$, the number of times that the lattice site at x is visited by a random walk, we need roughly 100 walks to reach this value of x. Since $x_{\max}$ scales linearly in the number of realizations, roughly $m = 100$ walks will reach a distance $x = \mathcal{M}x_0/m$. For example, for 25000 walks starting at $x_0 = 1$, roughly 100 of them will reach $x = 250$. Thus up to $x \approx 250$, the variation in $N(x)$ should be smaller than 10%, and beyond this point fluctuations should become
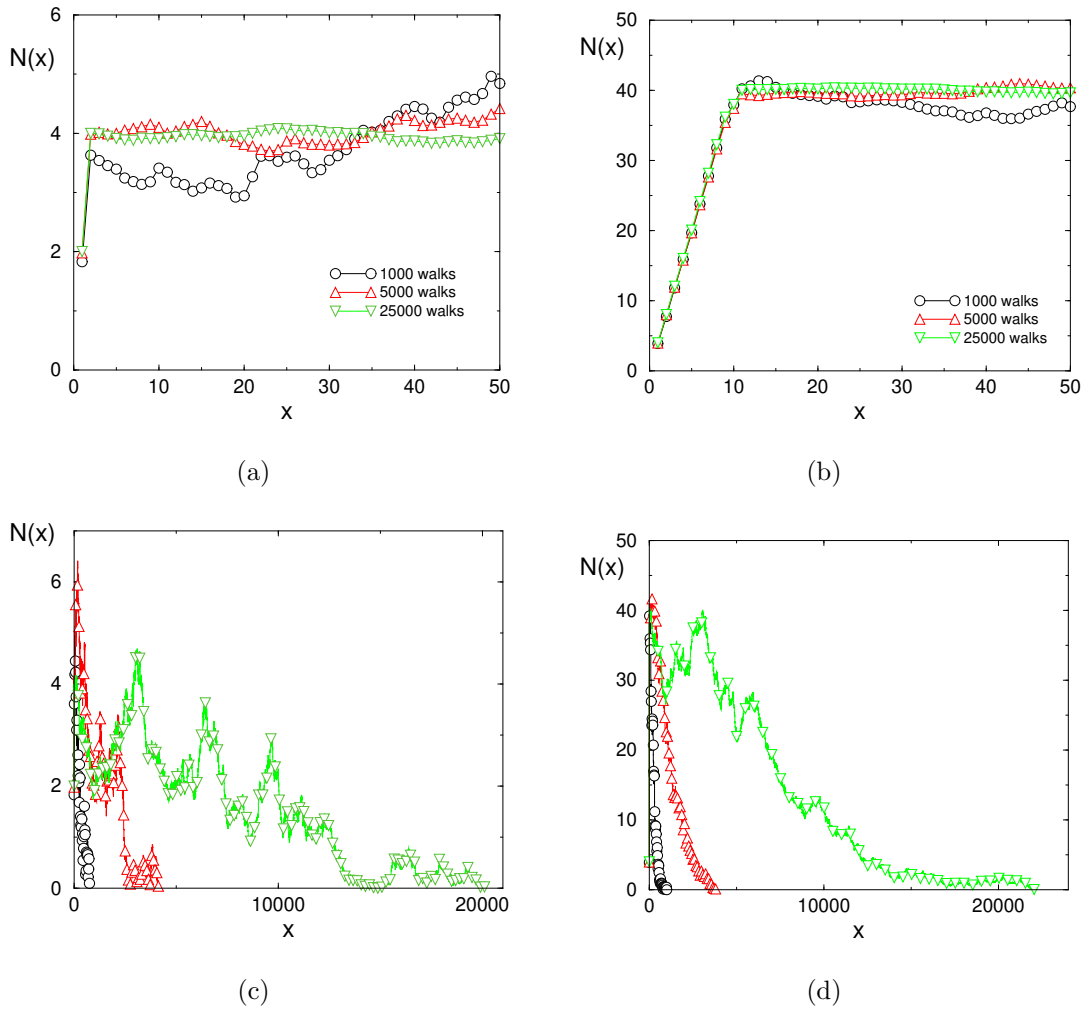
Figure 2. Simulation results for $N(x)$, the average number of times that a random walk visits x when it starts at: (a) $x_0 = 1$, and (b) $x_0 = 10$. (c) & (d): The same data as in (a) and (b) over the full range of x.

progressively more pronounced. This estimate is consistent with the data of Figs. 2(c) and (d).

The approach given here can be readily extended to any situation where the spatial probability distribution can be computed explicitly. We now investigate three such cases: (a) biased diffusion, (b) diffusion constrained to remain in the interval $[0, L]$, and (c) diffusion exterior to an absorbing sphere in general spatial dimension d.

2.2. Biased diffusion on the semi-infinite line

Suppose that a diffusing particle also experiences a constant bias velocity $-v$ that systematically pushes the particle towards the origin, so that the average time for the particle to reach the origin is finite. For a diffusing particle that starts at x_0,

its probability density can be obtained by the image method [14, 15], and is given by

$$P(x, t) = \frac{1}{\sqrt{4\pi Dt}} \left[e^{-(x-x_0+vt)^2/4Dt} - e^{-vx_0/D} e^{-(x+x_0+vt)^2/4Dt} \right]. \tag{3}$$

Notice that the magnitude of the image particle is different from that of the initial particle, while the velocities of the initial and image particles are the same.

We again integrate this expression over all time and obtain, for the time that the particle spends in $[x, x + dx]$ before it dies:

$$T(x) = \begin{cases} \dfrac{dx}{v} \left[1 - e^{-vx/D} \right] & x < x_0 \,, \\[2ex] \dfrac{dx}{v} e^{-vx/D} \left[e^{vx_0/D} - 1 \right] & x > x_0 \,. \end{cases} \tag{4a}$$

For $v \to 0$, Eq. (2) is recovered, while in the opposite limit of $v \to \infty$, (4a) reduces to

$$T(x) \to \begin{cases} \dfrac{dx}{v} & x < x_0 \,, \\[2ex] \dfrac{dx}{v} e^{-v(x-x_0)/D} & x > x_0 \,. \end{cases} \tag{4b}$$

As one might expect, the time spent in $[x, x + dx]$ with $x < x_0$ is just that of a ballistic particle, while it is exponentially unlikely for the particle to reach the classically forbidden region $x > x_0$ for large Péclet number, vx/D.

2.3. Diffusion in a finite interval

Suppose that an isotropically diffusing particle is constrained to remain within the interval $[0, L]$ and is eventually absorbed at $x = 0$. We again want the time $T(x)$ that the particle spends in $[x, x + dx]$ before it dies. As in the previous two subsections, we need the spatial probability distribution for a diffusing particle with absorbing boundary conditions at 0 and at L. A straightforward computation of this distribution is unwieldy, as it involves either an infinite Fourier series or an infinite number of images.

However, we can avoid this complication by noticing that we only want the integral of the probability distribution over all time, which corresponds to its Laplace transform at Laplace variable $s = 0$. The Laplace transform satisfies $s\widetilde{P} - \delta(x - x_0) = D\widetilde{P}_{xx}$, where $\widetilde{P}$ denotes the Laplace transform and the subscript denotes partial differentiation. For $s = 0$, this reduces to the Laplace equation

$$D\widetilde{P}_{xx} = -\delta(x - x_0) \,.$$

We solve this equation separately in the subdomains $x < x_0$ and $x > x_0$, impose the boundary conditions, continuity of the solution at $x = x_0$, and the joining condition $D\left(\widetilde{P}_x|_> - \widetilde{P}_x|_< \right) = -1$ to give, after standard steps,

$$\widetilde{P} = \frac{x_<}{D} \left(1 - \frac{x_>}{L} \right), \tag{5}$$
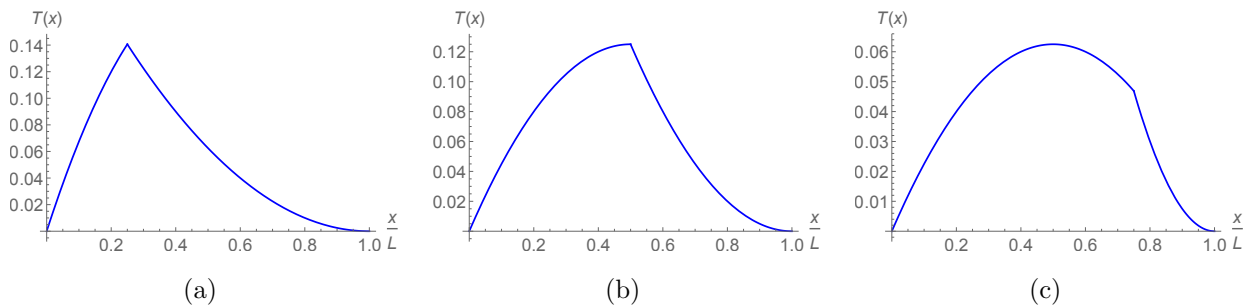
(a)   (b)   (c)

Figure 3. The residence time $T(x)$ for a diffusing particle that is constrained to remain within the interval $[0, L]$ until it exits at $x = 0$. Shown are the cases: (a) $x_0 = 0.25L$, (b) $x_0 = 0.5L$ and (c) $x_0 = 0.75L$.

where $\widetilde{P}_x|_>$ is the derivative just to the right of x_0 (and similarly for $\widetilde{P}_x|_<$), and $x_< = \min(x, x_0)$, $x_> = \max(x, x_0)$.

Finally, to obtain $T(x)$, we need to multiply the above distribution by the probability that the particle eventually exits the strip at $x = 0$, which is simply $1 - \frac{x}{L}$. Thus we have

$$T(x) = dx \int_0^\infty dt\, P(x, t) \left(1 - \frac{x}{L}\right) = \widetilde{P}(x, s = 0) \left(1 - \frac{x}{L}\right) dx,$$

$$= \begin{cases} \dfrac{x}{D}\left(1 - \dfrac{x_0}{L}\right)\left(1 - \dfrac{x}{L}\right) dx & x < x_0, \\[2mm] \dfrac{x_0}{D}\left(1 - \dfrac{x}{L}\right)^2 dx & x > x_0. \end{cases} \tag{6}$$

The maximum residence time occurs at $x = x_0$ for $x_0 < L/2$ and then "sticks" at $x = L/2$ for $x_0 \geq L/2$, with a cusp always occurring at $x = x_0$ (Fig. 3). In the limit $L \to \infty$, we recover the result (2) for diffusion on the semi-infinite line.

2.4. Diffusion exterior to a small sphere in dimension $d > 2$

We now determine the residence time for a diffusing particle that wanders in the region exterior to an absorbing sphere of radius a, a geometry that is the analog of the semi-infinite system in one dimension. Without loss of generality, we take the initial condition to be a spherical shell of unit total probability at radius r_0. We first treat the case of spatial dimensions $d > 2$ and then the special case of $d = 2$.

For general d, we need to solve the Laplace equation

$$D\nabla^2 \widetilde{P} = -\frac{1}{\Omega_d\, r_0^{d-1}}\, \delta(r - r_0), \tag{7}$$

where Ω_d is the surface area of a d-dimensional unit sphere and r_0 is the radial coordinate of the starting point. Because of the spherically symmetric source term, angular coordinates are irrelevant. We therefore separately solve $\widetilde{P}'' + \frac{d-1}{r}\widetilde{P}' = 0$ in the subdomains $a < r < r_0$ and $r_0 < r$, and then impose the absorbing boundary condition

at $r = a$ and the joining condition by integrating (7) over an infinitesimal interval than includes x_0. The result of these standard manipulations is

$$\widetilde{P}(r) = \frac{1}{(2-d)\Omega_d} \left[\left(\frac{r_<}{a} \right)^{2-d} - 1 \right] r_>^{2-d} . \tag{8}$$
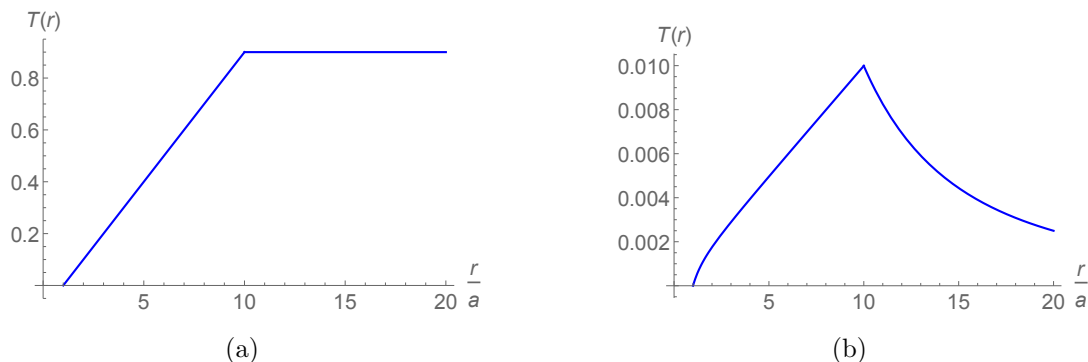

(a)

(b)

Figure 4. The residence time $T(r)$ for a diffusing particle that starts at $r_0 = 10$ exterior to a sphere of radius 1 in: (a) $d = 3$ and (b) $d = 5$.

To obtain $T(r)$, the residence time in a shell of radius r and thickness dr, we again need to multiply the above expression by the probability that the particle eventually hits the sphere, which, for $d > 2$, is simply $(a/r)^{d-2}$ [15]. Thus we have

$$
\begin{aligned}
T(r) &= \Omega_d \, r^{d-1} \widetilde{P}(r) \left(\frac{a}{r} \right)^{d-2} dr , \\
&= \begin{cases} \dfrac{dr}{D(d-2)} \left(\dfrac{a}{r} \right)^{d-2} \left[1 - \left(\dfrac{a}{r} \right)^{d-2} \right] \dfrac{r^{d-1}}{r_0^{d-2}} & r < r_0 , \\[2ex] \dfrac{dr}{D(d-2)} \left(\dfrac{a}{r} \right)^{d-2} \left[1 - \left(\dfrac{a}{r_0} \right)^{d-2} \right] r & r > r_0 . \end{cases}
\end{aligned} \tag{9}
$$

Two representative results are shown in Fig. 4. For large spatial dimension, a particle that eventually hits the sphere of radius a must do so quickly. Thus the residence time in the domain $r > r_0$ must necessarily be small, as shown in Fig. 4(b) for $d = 5$.

In spatial dimension $d = 2$, the result analogous to Eq. (8) is

$$\widetilde{P}(r) = \frac{1}{2\pi D} \ln \frac{r_<}{a} . \tag{10}$$

Since a diffusing particle always reaches the absorbing sphere in $d = 2$, we immediately have $T(r) = 2\pi r \widetilde{P}(r) \, dr$.

3. Visitation by a Discrete Random Walk

We now investigate the corresponding residence time for a symmetric random walk in the semi-infinite one-dimensional domain $[0, \infty]$. The walk starts at lattice site x_0 and

is absorbed when it first reaches $x = 0$. The analog of the residence time is $N(x)$, the number of times that the random walk visits site x (excluding the initial visit if $x = x_0$) before the walk dies. We use the generating function approach to derive this quantity for the case of $x_0 = 1$.

3.1. Average number of revisits to $x = 1$

For a random walk that starts at $x_0 = 1$ and is absorbed at $x = 0$, the number of steps in the walk is necessarily odd. For convenience, we write this number as $2n + 1$, with n an arbitrary non-negative integer. We define $A(n, k)$ as the number of random-walk paths that start at $x_0 = 1$, take the first step to the right (thus upward in the space-time representation of Fig. 5), and make k revisits to $x = 1$, before dying at the $(2n + 1)$st step. The number of such paths was found in [21] and is given by

$$A(n, k) = \frac{k\,(2n - k - 1)!}{(n - k)!\,n!}\,, \tag{11}$$

which happens to be directly related to the triangular Catalan numbers [22, 23]. To compute the average number of revisits to $x = 1$, we will need $\mathbf{P}(k\,|\,n)$, the conditional probability for a path to make exactly k revisits to $x = 1$ before dying at step $2n + 1$. This probability is

$$\mathbf{P}(k\,|\,n) = A(n, k)/C_n\,, \tag{12}$$

where $C_n = \frac{1}{n+1}\binom{2n}{n}$ is the nth Catalan number [24], which counts the total number of random walks of $2n$ steps that start at $x = 0$, remain in the region $x \geq 0$, and return to $x = 0$ at step $2n$. For what follows, we will also need

$$P(n) = C_n/2^{2n}\,, \tag{13}$$

the probability that a random walk first returns to its starting point at step $2n$.
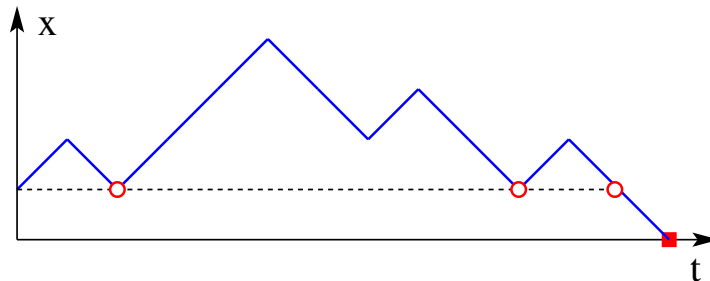

Figure 5. Space-time trajectory of a one-dimensional random walk of $2n + 1 = 13$ steps that starts at $x = 1$ and makes 3 revisits to $x = 1$ (red circles) before being absorbed at $x = 0$ (square). This path contributes to $A(6, 3)$.

From Eqs. (11) and (12), we have

$$\mathbf{P}(k\,|\,n) = \frac{k\,(2n - k - 1)!\,(n + 1)!}{(n - k)!\,(2n)!} = k\,\binom{n + 1}{k + 1}\bigg/\binom{2n}{k + 1}. \tag{14}$$

Thus the number of revisits to $x = 1$, averaged over all walks of $2n+1$ steps is given by

$$\langle k \rangle_n = \sum_{k=1}^{\infty} k \, \mathbf{P}(k \,|\, n) \,. \tag{15}$$

Using expression (14) for $\mathbf{P}(k \,|\, n)$ in the above average, we obtain the remarkably simple result

$$\langle k \rangle_n = \sum_{k=1}^{\infty} k^2 \binom{n+1}{k+1} \bigg/ \binom{2n}{k+1} = \frac{3n}{n+2} \,. \tag{16}$$

For long paths of $2n+1$ steps, there are, on average, 3 revisits to $x = 1$, after which the walk immediately dies.

We now determine the number of revisits to $x = 1$ upon also averaging over all n. This double average is

$$\langle\!\langle k \rangle\!\rangle = \sum_{n,k \geq 1} k \, \mathcal{P}(n, k) \,. \tag{17}$$

Here $\mathcal{P}(n, k)$ is the joint probability that the walk first reaches $x = 0$ at step $2n+1$, *and* the walk makes k revisits to $x = 1$ within $2n+1$ steps. This joint probability is

$$\mathcal{P}(n, k) = \mathbf{P}(k \,|\, n) \, P(n) = \frac{A(n, k)}{C_n} \frac{C_n}{2^{2n}} = \frac{A(n, k)}{2^{2n}} \,. \tag{18}$$

The average in (17) may be now be expressed in terms of the generating function for $\mathcal{P}(n, k)$:

$$\begin{aligned} g(x, y) &= \sum_{n,k \geq 1} \mathcal{P}(n, k) \, x^n \, y^k = \sum_{n,k \geq 1} \frac{1}{2^{2n}} A(n, k) \, x^n \, y^k \,, \\ &= \sum_{n,k \geq 1} \frac{1}{2^{2n}} \frac{(2n - k - 1)! \, k}{(n - k)! \, n!} x^n \, y^k \,, \\ &= \frac{xy}{2 - xy + 2\sqrt{1 - x}} \,, \end{aligned} \tag{19}$$

Which was derived in [21] (see also [25]). In terms of the generating function, we immediately obtain the remarkably simple result

$$\langle\!\langle k \rangle\!\rangle = \sum_{n,k \geq 1} k \, \mathcal{P}(n, k) \, x^n \, y^k \bigg|_{(1,1)} = y \, \frac{\partial g}{\partial y} \bigg|_{(1,1)} = 2 \,. \tag{20}$$

There are, on average, 2 revisits to $x = 1$ in the ensemble of all random walks that start at $x = 1$, take their first step to the right, and are eventually absorbed at $x = 0$.

We can extend Eq. (16) to higher integer moments of the average number of revisits to $x = 1$ for walks of $2n + 1$ steps. The first few of these fixed-n moments are:

$$\langle k^2 \rangle_n = \frac{n(13n - 1)}{(2 + n)(3 + n)},$$

$$\langle k^3 \rangle_n = \frac{15n^2(5n - 1)}{(2 + n)(3 + n)(4 + n)},$$

$$\langle k^4 \rangle_n = \frac{541n^2 - 196n^2 + 11n^2 + 4n}{(2 + n)(3 + n)(4 + n)(5 + n)}, \tag{21a}$$

etc. We can similarly compute the higher integer moments of the number of revisits to $x = 1$, averaged over all walk lengths, and the first few are:

$$\langle\langle k^2 \rangle\rangle = 6 \qquad \langle\langle k^3 \rangle\rangle = 26 \qquad \langle\langle k^4 \rangle\rangle = 150 \qquad \langle\langle k^5 \rangle\rangle = 1082, \tag{21b}$$

etc. Parenthetically, these numbers are also sequence A000629 in the On-Line Encyclopedia of Integer Sequences [26]

In the next section, we will also need the generating function when the first step of the walk can equiprobably be to the right or to the left. This leads to the possibility that the total number of steps $2n + 1 = 1$, i.e., $n = 0$, for which the number of revisits to 1 equals zero. The generating function for the joint probability $\mathcal{P}(n, k)$ for this ensemble of random walks therefore is

$$G(x, y) = \sum_{n,k \geq 0} \mathcal{P}(n, k)\, x^n\, y^k = \tfrac{1}{2}\big[1 + g(x, y)\big],$$

$$= \frac{1}{2}\left(1 + \frac{xy}{2 - xy + 2\sqrt{1 - x}}\right), \tag{22}$$

where the term $\frac{1}{2}$ comes from the walk that initially steps to the left and is immediately absorbed. Notice that $y\frac{\partial G}{\partial y}\big|_{(1,1)} = 1$, which is consistent with (19): half of all paths die immediately upon the first step, and thus never return to 1, while the other half return twice, on average, as derived in (20).

3.2. Average number of visits to $x = 2$

We now extend the above approach to a walk that starts at $x = 1$ and is constrained to take its first step to the right, to determine the number of revisits to $x = 2$. For this purpose, we define three random variables that characterize this set of walks:

- $2n + 1$, the total number of steps in the walk when it dies;
- k, the number of visits to $x = 2$ (*including* the first visit);
- ℓ, the number of excursions that lie above the level $x = 1$.

Since an excursion is a path that lies between two successive returns to $x = 1$ (and thus always remains above $x = 1$), the minimal length excursion is the path $1 \to 2 \to 1$.

We want the ensemble average of the number of revisits to $x = 2$. To facilitate this calculation, it is useful to define the three-variable generating function

$$\mathcal{G}(x, y, z) = \sum_{n \geq 1} \sum_{1 \leq \ell \leq n} \sum_{\ell \leq k \leq n} \mathcal{P}(n, k, \ell) \, x^n \, y^k \, z^\ell, \tag{23}$$

which encodes all paths according to (n, k, ℓ). We also label each successive excursion of the path above $x = 1$ by the index $1 \leq i \leq \ell$, and we introduce the variables $2m_i$ and j_i, respectively, for the number of steps in the ith such excursion, and the number of returns to $x = 2$ in this excursion (Fig. 6). As shown in this figure, $2m_i$ counts the number of steps that lie above $x = 1$. Thus for an excursion that goes from $x = 1$ to $x = 2$ and immediately returns to $x = 1$, $m_i = 0$. In addition, j_i counts the number of *revisits* to 2, so that the total number of visits to $x = 2$ in the ith excursion above $x = 1$ is $j_i + 1$. The variables j_i, m_i, and ℓ must satisfy the geometric constraints (see Fig. 6):

$$j_1 + j_2 + \cdots + j_\ell + \ell = k,$$
$$m_1 + m_2 + \cdots + m_l + \ell = n, \tag{24}$$
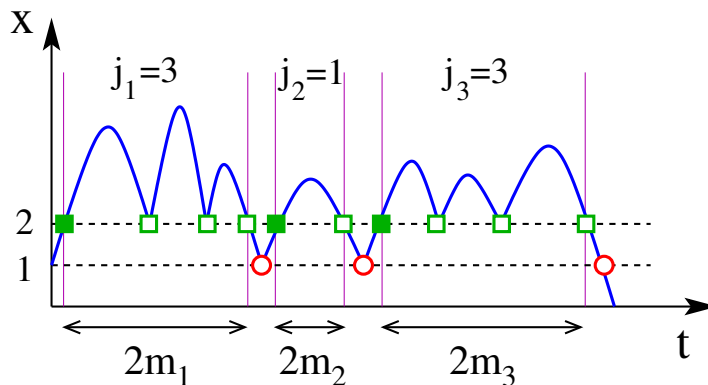$$j_i \leq m_i.$$

Figure 6. Schematic space-time trajectory of a random walk that starts at $x = 1$ and has 3 revisits to $x = 1$ (red circles) and $k = 10$ revisits to $x = 2$ (green squares). There are $\ell = 3$ excursions above $x = 1$. Immediately after a revisit to $x = 1$, the next revisit to $x = 2$ is shown as a solid green square.

Using these definitions, the three-variable generating function $\mathcal{G}(x, y, z)$ can be functionally expressed in terms of $G(x, y)$ defined in Eq. (22) as (see the Appendix for details of this derivation)

$$\mathcal{G}(x, y, z) = \sum_{\ell \geq 1} P(\ell) \, [x \, y \, z \, G(x, y)]^\ell, \tag{25}$$

where $P(\ell)$ is the probability that there are ℓ excursions above $x = 1$ averaged over walks of any length, which is also the distribution of the number of returns to $x = 1$.

Thus one may compute $P(\ell)$ as the marginal of the joint distribution of the number of steps and the number of excursions:

$$P(\ell) = \sum_{n \geq 1} \mathcal{P}(n, \ell) = \sum_{n \geq 1} \frac{1}{2^{2n}} A(n, \ell) \,. \tag{26}$$

The above sum starts at $n = 1$ because we are imposing the condition that the first step of the walk is to the right. Consequently the three-variable generating function in (25) will ultimately be expressed in terms of the restricted generating function g. Comparing the above formula with the first line of (19), we obtain

$$\begin{aligned}
\mathcal{G}(x, y, z) &= \sum_{\ell \geq 1} P(\ell) \left[x\, y\, z\, G(x, y) \right]^{\ell} \\
&= \sum_{\ell \geq 1} \sum_{n \geq 1} \frac{1}{2^{2n}} A(n, \ell) \left[x\, y\, z\, G(x, y) \right]^{\ell} \\
&= g\big(1, x\, y\, z\, G(x, y)\big) \,.
\end{aligned} \tag{27}$$

It is now straightforward to calculate of $\langle k \rangle$. From the definition of the generating function (23), we have

$$\begin{aligned}
\langle k \rangle &= y \left. \frac{\partial \mathcal{G}}{\partial y} \right|_{x=y=z=1} , \\
&= y \left. \frac{\partial g}{\partial y} \right|_{x=y=1} \left[G(1,1) + y \left. \frac{\partial G}{\partial y} \right|_{x=y=1} \right] , \\
&= 2 \times (1 + 1) \,, \\
&= 4 \,.
\end{aligned} \tag{28}$$

A random walk thus visits $x = 2$ twice as often as $x = 1$, as already predicted by the continuum solution (2).

3.3. Average number of visits to $x > 2$

The ensemble average of the number of visits to a given level $x \geq 2$ may be readily computed by induction. We start by calculating the average number of visits to $x = 3$, and it will become apparent that this approach applies for any $x > 2$. Each time a random walk reaches $x = 2$, there are two possibilities at the next step: the walk may step forward to $x = 3$ or step back to $x = 1$, each with probability $\frac{1}{2}$. Let us first assume that the walk goes to $x = 3$, which occurs with probability $\frac{1}{2}$. Each time this event occurs, we now ask: what is the average number of visits to $x = 3$ (including this first visit) before the walk returns to $x = 2$?

With probability $\frac{1}{2}$, the walk may immediately return to $x = 2$, in which case, there is one visit to $x = 3$. On the other hand, if the walk steps to $x = 4$, we have the same situation as that discussed in Sec. 3.1. Namely, if we view $x = 3$ as the starting point,

we know that there are 2 revisits to $x = 3$ and thus 3 visits to $x = 3$, on average, before the walk steps back to $x = 2$. Thus each time $x = 3$ is reached, there are

$$\left(\frac{1}{2} \times 1\right) + \left(\frac{1}{2} \times 3\right) = 2$$

two visits, on average, to $x = 3$.

For a walk that reaches $x = 2$, the average number of visits to $x = 3$ for this visit to $x = 2$ therefore is

$$\left(\frac{1}{2} \times 0\right) + \left(\frac{1}{2} \times 2\right) = 1\,.$$

The first term corresponds to the contribution from a walk that steps from $x = 2$ to $x = 1$ without hitting $x = 3$, and the second term is the contribution when the walk steps from $x = 2$ to $x = 3$.

To summarize, each time the walk visits $x = 2$, there is, on average, one visit to $x = 3$, before the walk is at $x = 2$ again. Clearly, this reasoning that determines the number of visits to $x + 1$ for each visit to x applies inductively for any level $x > 2$. Thus we conclude that the average number of times that a random walk visits a given level $x > 2$, equals 4, in agreement with the simulation results in Fig. 2(a). Clearly, our argument also applies for any starting point of the walk x_0, as long as we restrict to coordinates with $x > x_0 + 1$.

4. Time of the First Revisit

In addition to the *number* of revisits to $x = 1$ by a random walk excursion that starts at $x = 1$ and is eventually absorbed, we are interested in the *time* at which the first revisit occurs. This time characterizes the shape of the space-time trajectory of a random walk. Since the walk starts at $x = 1$ and ends at $x = 0$, its space-time shape is essentially that of a Brownian excursion — a Brownian trajectory that starts at $x = 0$, remains above $x = 0$ for all $0 < t < T$, and returns to $x = 0$ for the first time at $t = T$. The average shape of a Brownian excursion has been shown to be semi-circular [6,7]. From this shape, we might anticipate that the first revisit to $x = 1$ is unlikely to occur for t near $T/2$ because such a revisit involves a large fluctuation from the average trajectory. Instead, it seems more likely that the first revisit to $x = 1$ will occur either near the beginning or the end of the excursion, a feature that evokes the famous arcsine laws [1,4]. We now show that this expectation is correct.

4.1. Time of first return to 1 for fixed walk length n

Consider a random walk that starts at $x = 1$, takes its first step to the right, and is absorbed at $x = 0$ after $T = 2n + 1$ steps. What is the probability $P(2m \,|\, T)$ that such a walk revisits $x = 1$ *for the first time* at step $\tau_1 = 2m$? Since the walk necessarily revisits $x = 1$ at step $2n$ by definition, and the walk could revisit $x = 1$ immediately after 2

steps, m satisfies the constraint $1 \le m \le n$. The number of walks that revisit $x = 1$ after $2m$ steps may be obtained by decomposing the full path into two constituents (Fig. 7):

- Excursions of $(2m - 2)$ steps that wander in the domain $x \ge 2$ — the number of such paths is C_{m-1};
- Excursions of $(2n - 2m)$ steps that wander in the domain $x \ge 1$ — the number of such paths is C_{n-m}.

The first part accounts for the first return to $x = 1$ at step $2m$ and the second part accounts for the remaining path of $2n - 2m$ steps.
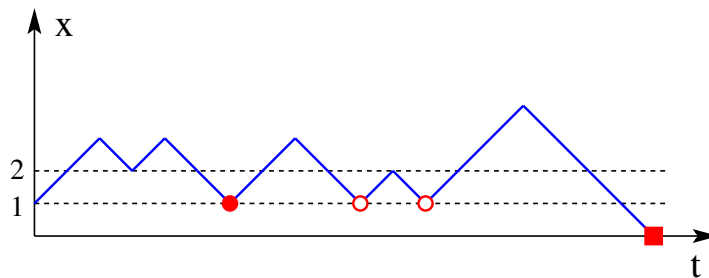

Figure 7. Space-time trajectory of a one-dimensional random walk that starts at $x = 1$ and first revisits $x = 1$ at step 6 (solid circle). Subsequent revisits to $x = 1$ are indicated by open circles and the walk is absorbed when it first reaches $x = 0$ (square).

The required probability is then simply the product of these two numbers divided by the total number of walks that start at $x = 1$ and are absorbed after $2n + 1$ steps, which is C_n. Therefore

$$\mathbf{P}\left(2m \,|\, T\right) = \frac{C_{m-1} C_{n-m}}{C_n} = \frac{n+1}{m\left(n - m + 1\right)} \frac{\binom{2m-2}{m-1}\binom{2n-2m}{n-m}}{\binom{2n}{n}}. \tag{29}$$

Conditioned on $T = 2n + 1$, the average value of τ_1, the time of the first revisit, can be immediately seen to be

$$\langle \tau_1 \rangle_n = n + 1, \tag{30}$$

because $\mathbf{P}\left(\tau_1 = 2m \,|\, T = 2n + 1\right)$ is symmetric under $m \to n + 1 - m$. The above result may also be obtained by direct calculation. Because of the bimodal nature of the underlying probability distribution, the average value is very different from the typical value. The average corresponds to the minimum of the probability distribution (Fig. 8(a)), just as in the arcsine law for the time of the last zero of a Brownian motion.
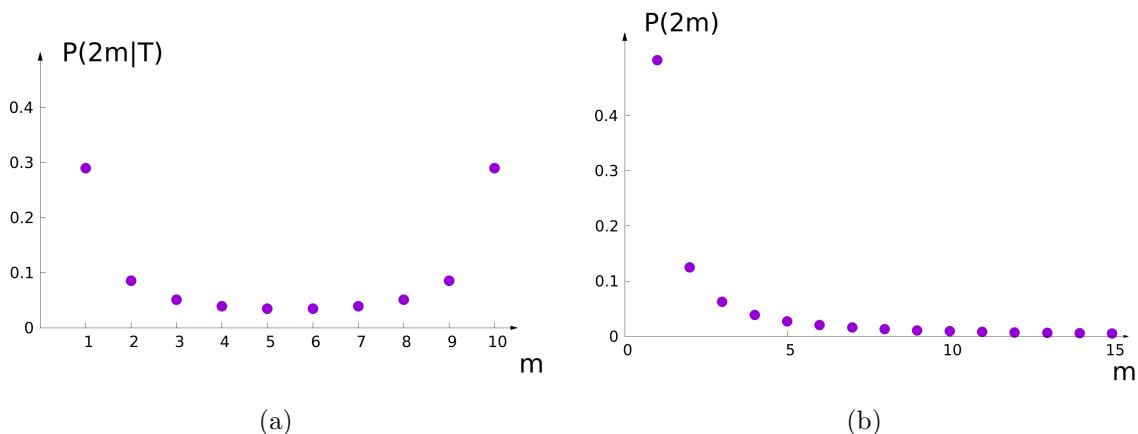
Figure 8. (a) Conditional distribution $P\left(\tau_1 = 2m \mid T = 2n+1\right)$ for $n = 10$. Note that typical (i.e. most likely) values of τ_1 are 2 and $2n$, while the average value is $\langle\tau_1\rangle = n+1$. (b) Distribution $P\left(\tau_1 = 2m\right)$ of the first revisit time to 1.

4.2. Time of first return to 1 for any n

From the conditional probability $\mathbf{P}\left(2m \mid T\right)$, we may now compute the joint probability $\mathcal{P}\left(2m, T\right)$:

$$
\begin{aligned}
\mathcal{P}\left(2m, T\right) &= \mathbf{P}\left(2m \mid T\right)\, P(T) \\
&= \frac{C_{m-1}\, C_{n-m}}{C_n} \times \frac{C_n}{2^{2n}} = \frac{C_{m-1}\, C_{n-m}}{2^{2n}}.
\end{aligned}
\tag{31}
$$

With this result, we can readily obtain the distribution $P(2m)$, the probability for a path of any length to perform an excursion of $2m$ steps that lies above $x = 2$ between steps 1 and $2m - 1$ (with the first step constrained to go from $x = 1$ to $x = 2$):

$$
\begin{aligned}
P(2m) &= \sum_{n \geq m} \mathcal{P}\left(2m, T\right) \\
&= C_{m-1} \sum_{n \geq m} \frac{C_{n-m}}{2^{2n}} = \frac{C_{m-1}}{2^{2m-1}}.
\end{aligned}
\tag{32}
$$

This distribution is normalized, because $\sum_{m \geq 1} C_{m-1}/2^{2m-1} = 1$. The Markovian nature of the random walk means that there is no memory between what happens after step $2m$ and the probability that the walk first revisits $x = 1$ at step $2m$. Hence $P(2m)$ is simply the probability that a symmetric random walk first returns to its starting point at step $2m$, which asymptotically scales as $m^{-3/2}$ [1, 4, 15]. Because of this scaling, the average time for the first return, $\langle\!\langle\tau_1\rangle\!\rangle = \sum_m 2m\, P(2m)$, is infinite, even though $P(2m)$ is peaked at $m = 1$ (Fig. 8(b)).

5. Summary

We showed how straightforward methods from first-passage processes could be used to determine the time that a diffusing particle spends in a given spatial range when the

particle starts at some point x_0 and dies when it reaches an absorbing point or set. We also derived corresponding results for the discrete random walk, where the analog of the residence time is the number of times that a given point is visited. For continuum diffusion, the residence time at a given point is simply the integral of the probability distribution over all time at this same point, with the given initial condition and the absorbing boundary condition. This perspective allowed us to also treat in a relatively simple manner biased diffusion, diffusion in a finite interval (conditioned on absorption at a given side of the interval), and diffusion in general spatial dimensions. It is also worth emphasizing the time integral of the probability distribution at a given point is essentially just the electrostatic potential at this point. This correspondence provides a simple way to calculate residence times and to understand the dependence of the residence time on basic parameters.

The main qualitative feature of the residence time at a given point is that it vanishes (often linearly) as the distance between this point and the absorber. That is, a diffusing particle simply does not linger when it is close to an absorbing point. Another interesting feature, almost intuitive from an analogy with electrostatics, is the fact that, in low dimensions ($d \leq 3$), the average residence time at any point beyond the starting point is constant: it is simply equal to the average residence time at the starting point. This is no longer the case for $d \geq 4$, with an abrupt transition between $d = 3$ and $d = 4$. It would be of interest to understand why this differs from the well known transition between recurrence and transience, which happens between $d = 2$ and $d = 3$.

For the discrete random walk, we exploited the generating function method to derive parallel results for the number of times that a given lattice site is visited before the walk dies at the absorbing point. For a walk that starts at $x_0 = 1$, there are, upon averaging over walks of all possible lengths, two subsequent visits to $x = 1$ and four subsequent visits to $x = 2$ before the random walk dies. We also showed that the random walk makes four subsequent visits to any point $x \geq 2$, on average. We also found that the first revisit to $x = 1$ occurs near the start or the end of the path. This means that it is very unlikely that there will be a large deviation toward the boundary and away from the average position of the path near the middle of an excursion. This suggests that an individual Brownian excursion always remains close to its average shape. We hope to investigate this behavior in future work, with a view to obtaining a full characterization of the fluctuations around an excursion's average semi-circular shape.

Acknowledgments

Appendix A. Relation Between Generating Functions

Using the geometric constraints (24) in Eq. (23), the generating function $\mathcal{G}(x, y, z)$ can be re-expressed as

$$\mathcal{G}(x, y, z) = \sum_{n=1}^{\infty} \sum_{\ell=1}^{n} \sum_{k=\ell}^{n} \mathcal{P}(n, k, \ell)\, x^n\, y^k\, z^\ell$$
$$= \sum_{n=1}^{\infty} \sum_{\ell=1}^{n} \sum_{k=\ell}^{n} P(\ell)\, \mathbf{P}(n, k|\ell)\, x^n\, y^k\, z^\ell, \tag{A.1}$$

where we use $\mathbf{P}(\ldots|\ldots)$ to denote the conditional joint probability.

Next we write $\mathbf{P}(n, k|\ell)$ in terms of the variables m_i and j_i (see Fig. 6):

$$\mathcal{G}(x, y, z) = \sum_{n=1}^{\infty} \sum_{\ell=1}^{n} \sum_{k=\ell}^{n} \sum_{\substack{j_1+\cdots+j_\ell=k-\ell \\ m_1+\cdots+m_\ell=n-\ell \\ j_i \leq m_i}} P(\ell)\, \mathbf{P}(m_1, \ldots, m_\ell, j_1, \ldots, j_\ell|\ell)\, x^{m_1+\ldots m_\ell+\ell}\, y^{j_1+\ldots j_\ell+\ell}\, z^\ell$$

$$= \sum_{n=1}^{\infty} \sum_{\ell=1}^{n} \sum_{\substack{m_1+\cdots+m_\ell=n-\ell \\ 0 \leq j_i \leq m_i}} P(\ell)\, \mathbf{P}(m_1, \ldots, m_\ell, j_1, \ldots, j_\ell|\ell)\, x^{m_1+\ldots m_\ell+\ell}\, y^{j_1+\ldots j_\ell+\ell}\, z^\ell$$

$$= \sum_{\ell \geq 1} P(\ell) \sum_{n \geq \ell} \sum_{\substack{m_1+\cdots+m_\ell=n-\ell \\ 0 \leq j_i \leq m_i}} \mathbf{P}(m_1, \ldots, m_\ell, j_1, \ldots, j_\ell|\ell)\, x^{m_1+\ldots m_\ell+\ell}\, y^{j_1+\ldots j_\ell+\ell}\, z^\ell$$

$$= \sum_{\ell \geq 1} P(\ell) \sum_{m_1, \ldots, m_\ell \geq 0} \sum_{j_1, \ldots, j_\ell=0}^{m_1, \ldots, m_\ell} \left[\prod_{i=1}^{\ell} \mathbf{P}(m_i, j_i|\ell)\, x^{m_i}\, y^{k_i} \right] (xyz)^\ell$$

$$= \sum_{\ell \geq 1} P(\ell)\, (xyz)^\ell \sum_{m_1, \ldots, m_\ell \geq 0} \sum_{j_1, \ldots, j_\ell=0}^{m_1, \ldots, m_\ell} \left[\prod_{i=1}^{\ell} \mathbf{P}(m_i, j_i|\ell)\, x^{m_i}\, y^{j_i} \right]$$

$$= \sum_{\ell \geq 1} P(\ell)\, (xyz)^\ell \prod_{i=1}^{\ell} \left[\sum_{m_i \geq 0} \sum_{0 \leq j_i \leq m_i} \mathbf{P}(m_i, j_i|\ell)\, x^{m_i}\, y^{j_i} \right]. \tag{A.2}$$

We now use the fact that the random walk is a Markov process, which implies that $\mathbf{P}(m_i, j_i|\ell) = \mathcal{P}(m_i, j_i) = A(m_i, j_i)/2^{2m_i}$. Therefore

$$\sum_{m_i \geq 0} \sum_{0 \leq k_i \leq m_i} \mathbf{P}(m_i, j_i|\ell)\, x^{m_i}\, y^{j_i} = G(x, y). \tag{A.3}$$

Note that $G(x, y)$, as defined in Eq. (22), appears here and not $g(x, y)$ because upon starting from $x = 2$, when coming from $x = 1$, the path is not conditioned to immediately move to $x = 3$. In fact, the path is allowed to return immediately to $x = 1$, as reflected in the fact that, for the ith excursion, m_i may be 0.

Eq. (A.2) becomes

$$\mathcal{G}(x, y, z) = \sum_{\ell \geq 1} P(\ell)\, (xyz)^\ell \prod_{1 \leq i \leq \ell} [G(x, y)] = \sum_{\ell \geq 1} P(\ell)\, [x\, y\, z\, G(x, y)]^\ell. \tag{A.4}$$

This is Eq. (25) in the main text.

References

[1] W. Feller, *Introduction to Probability Theory and its Applications, 3rd ed.*, (J Wiley & Sons, Inc. New York, 1964).

[2] F. Spitzer, *Principles of Random Walk*, (Van Nostrand, Princeton NJ, 1964)

[3] E. W. Montroll and G. H. Weiss, J. Math. Phys. **6**, 167 (1965).

[4] P. Mörters and Y. Peres, *Brownian Motion*, (Cambridge University Press, Cambridge, UK, 2010).

[5] K. L. Chung, Ark. Mat. **14**, 155 (1976).

[6] A. Baldassarri, F. Colaiori, and C. Castellano, Phys. Rev. Lett. **90**, 060601 (2003).

[7] F. Colaiori, A. Baldassarri, and C. Castellano, Phys. Rev. E **69**, 041105 (2004).

[8] P. Levy, *Processus stochastique et mouvement Brownien*, (Gauthier-Villars, Paris, 1948).

[9] D. Ray, Illinois Journal of Mathematics **7**, 615 (1963)

[10] F. B. Knight, Trans. Am. Math, Soc. **109**, 56 (1963).

[11] S. Condamin, V. Tejedor, and O. Bénichou, Phys. Rev. E **76**, 050102 (2007).

[12] O. Bénichou and J. Desbois, J. Phys. A: Math. Theor. **42**, 015004 (2009).

[13] O. Bénichou and R. Voituriez, Phys. Repts. **539**, 225 (2014).

[14] H. E. Daniels, Ann. Statist. **10**, 394 (1982).

[15] S. Redner, *A Guide to First-Passage Processes*, (Cambridge University Press, Cambridge, UK, 2001).

[16] M. Ferraro and L. Zaninetti, Phys. Rev. E **64**, 056107 (2001).

[17] M. Ferraro and L. Zaninetti, Physica A **338**, 307 (2004).

[18] M. Ferraro and L. Zaninetti, Phys. Rev. E **73**, 057102 (2006).

[19] J. Galambos, *The Asymptotic Theory of Extreme Order Statistics*, (Krieger, Malabar, FL, 1987).

[20] P. L. Krapivsky, S. Redner, and E. Ben-Naim, *A Kinetic View of Statistical Physics*, (Cambridge University Press, Cambridge UK, 2010).

[21] U. Bhat, S. Redner and O. Bénichou, J. Stat. Mech. 073213 (2017).

[22] D. F. Bailey, Math. Mag. **69**, 128 (1996) (http://www.jstor.org/stable/269067).

[23] K.-H. Lee and S.-J. Oh, arXiv:1601.06685.

[24] R. P. Stanley, *Catalan Numbers*, (Cambridge University Press, Cambridge UK, 2015).

[25] R. P. Stanley, *Enumerative Combinatorics vol. 2*, (Cambridge University Press, Cambridge UK, 1999).

[26] *The On-Line Encyclopedia of Integer Sequences*, https://oeis.org.