Journal of Statistical Mechanics:
Theory and Experiment

PAPER: CLASSICAL STATISTICAL MECHANICS, EQUILIBRIUM AND NON-EQUILIBRIUM

Residence time near an absorbing set

Journal of Statistical Mechanics: Theory and Experiment
An IOP and SISSA journal

PAPER: Classical statistical mechanics, equilibrium and non-equilibrium

Residence time near an absorbing set

J Randon-Furling[1] **and S Redner**[2]

[1] SAMM (EA 4543), Université Paris-1 Panthéon-Sorbonne, 90 rue de Tolbiac, 75013 Paris, France
[2] Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, United States of America
E-mail: redner@santafe.edu

Received 27 June 2018
Accepted for publication 7 September 2018
Published 19 October 2018

CrossMark

Online at stacks.iop.org/JSTAT/2018/103205
https://doi.org/10.1088/1742-5468/aae02a

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 of the particle in the range $[x, x + \mathrm{d}x]$ is $T(x) = \frac{x}{D}\,\mathrm{d}x$ for $x < x_0$ and $\frac{x_0}{D}\,\mathrm{d}x$ 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 then use the generating function technique to derive parallel results for the average number of times that a one-dimensional symmetric nearest-neighbor random walk visits site x when the walk starts at $x_0 = 1$ and is absorbed at $x = 0$. We also determine the distribution of times when the random walk first revisits $x = 1$ before being absorbed.

Keywords: diffusion, residence time

J. Stat. Mech. (2018) 103205

1742-5468/18/103205+18$33.00

Contents

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 (figure 1). Consider a Brownian particle that starts at x_0 at $t = 0$ and first returns to $x = 0$ at time T_0. In the interesting case of $T_0 \gg x_0^2/D$, the particle wanders over a large spatial range before its eventual demise. A trajectory that stays strictly in the range $x > 0$ until absorption at time T_0 is known as a *Brownian excursion* when the starting point 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 $[x, x + \mathrm{d}x]$ before being absorbed? We term this latter quantity as the residence time. The properties of the residence time have 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 $[x, x + \mathrm{d}x]$ before being absorbed when the origin is reached. When $x < x_0$, the distribution of this

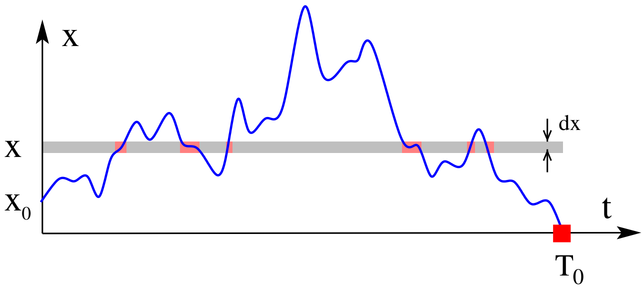

Figure 1. Schematic trajectory of a diffusing particle in space time that starts at x_0 and is absorbed at time T_0 (square). The time $T(x)$ that the particle spends in $[x, x + \mathrm{d}x]$ is indicated by the colored segments.

residence time was shown to be related 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 average 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 [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 simple 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. Some of these situations have been treated previously in [14], by a more formal approach than that presented here.

In section 2.1, we first derive the average residence time within the interval $[x, x + \mathrm{d}x]$ in continuum diffusion, by solving the relevant diffusion equation. We extend this approach to: (a) biased diffusion on the semi-infinite line (section 2.2), and (b) unbiased diffusion in a finite domain $[0, L]$, with the condition that the particle is eventually absorbed at $x = 0$ (section 2.3). We then determine the average residence time in general spatial dimension in the domain exterior to an absorbing sphere of radius a (section 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 section 3, we use generating function methods to derive the average 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 average number of times that $x = 1$ is revisited equals $3n/(n + 2)$. By averaging this quantity over the total number of steps of the walk, we will show that there are, on average, 2 revisits to $x = 1$. 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 section 4, we determine the time when a walk first revisits $x = 1$,

J. Stat. Mech. (2018) 103205

when it starts at $x = 1$ and is eventually absorbed at $x = 0$. We give some concluding comments in section 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 for the particle to be at position $x > 0$ as [15, 16]

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

The time $T(x)$ that the particle, which starts at x_0, spends in the range $[x, x+\mathrm{d}x]$ before being absorbed at $x = 0$ (figure 1) is simply the integral of the probability density over all time times $\mathrm{d}x$ (see [2, 3, 17–19] for related approaches). Performing this integral, with $P(x,t)$ from equation (1), the residence time $T(x)$ is given by

$$T(x) = \mathrm{d}x\int_0^\infty \mathrm{d}t\,P(x,t) = \begin{cases} \dfrac{x}{D}\,\mathrm{d}x & x < x_0, \\ \dfrac{x_0}{D}\,\mathrm{d}x & x > x_0. \end{cases} \tag{2}$$

To illustrate this result, we present simulations for a nearest-neighbor random walk that starts at: (a) $x_0 = 1$ and (b) at $x_0 = 10$ in figure 2. As a function of the number of walks in the ensemble, $T(x)$ slowly converges to the asymptotic time-independent value in equation (2). A curious feature of this residence-time data is that it becomes erratic for large x, as shown in figures 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 $S(t) = \mathrm{erf}\big(x_0/\sqrt{4Dt}\big) \simeq x_0/\sqrt{4Dt}$ for $t \to \infty$ [16]. For $\mathcal{M}$ random walks, we estimate the longest lived of them by the extreme-statistics criterion $S(t_{\max}) \simeq 1/\mathcal{M}$ [20, 21], which states that one out of $\mathcal{M}$ 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 large-x fluctuations in figures 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 25 000 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 progressively more pronounced. This estimate is consistent with the data of figures 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

https://doi.org/10.1088/1742-5468/aae02a

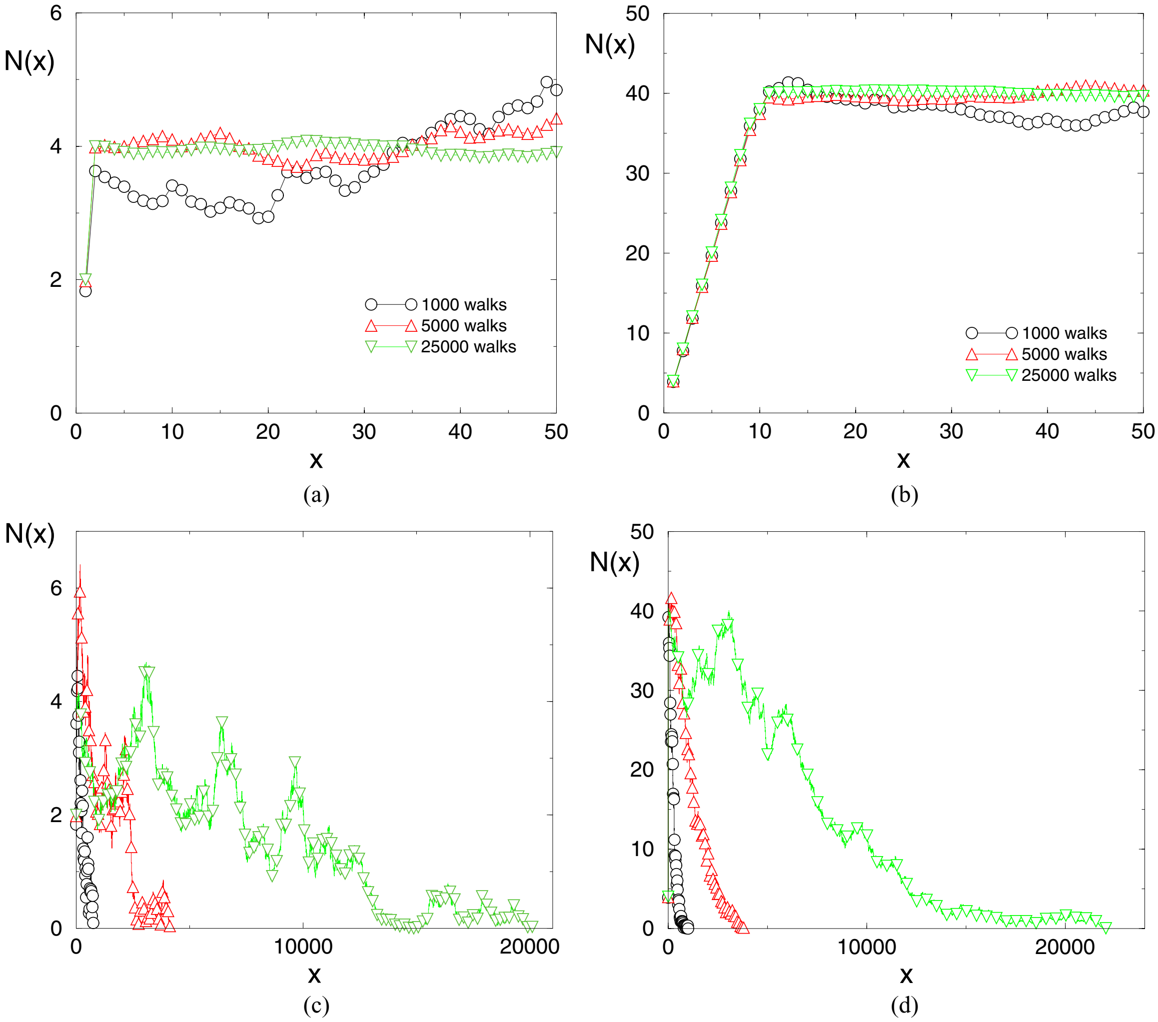

Figure 2. Simulation results for $N(x)$, the average number of times that a random walk visits x when it starts at: (a) $x_0 = 1$, and (b) $x_0 = 10$. Here $N(x)$ is the discrete analog of the residence time $T(x)$. (c) & (d): The same data as in (a) and (b) over the full range of x.

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 [15, 16], and is given by

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

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 $[x, x+\mathrm{d}x]$ before it dies:

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

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

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

As one might expect, the time spent in $[x, x+\mathrm{d}x]$ 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 $[x, x+\mathrm{d}x]$ 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 $D\big(\widetilde{P}_x|_{>}-\widetilde{P}_x|_{<}\big)=-1$ to give, after standard steps,

$$\widetilde{P}=\frac{x_<}{D}\Big(1-\frac{x_>}{L}\Big),\tag{5}$$

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

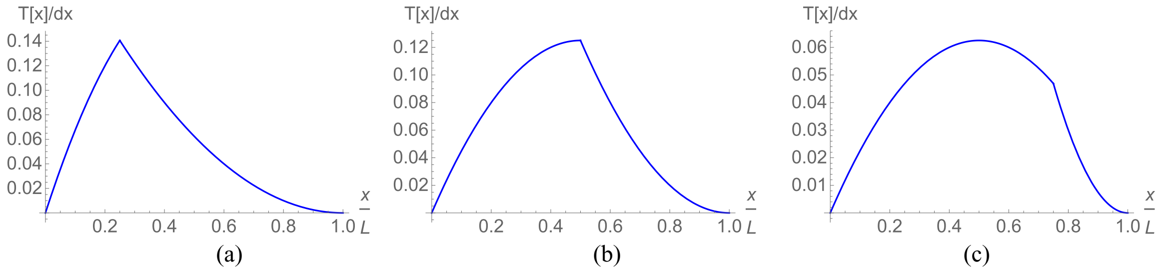

Figure 3. The residence time density $T(x)/\mathrm{d}x$ for a diffusing particle that is constrained to remain within the interval $[0, L]$ until it exits at $x = 0$. Shown are the cases: (a) $x_0 = 0.25L$, (b) $x_0 = 0.5L$ and (c) $x_0 = 0.75L$.

$$
\begin{aligned}
T(x) &= \mathrm{d}x \int_0^\infty \mathrm{d}t\, P(x,t)\left(1-\frac{x}{L}\right) = \widetilde{P}(x, s=0)\left(1-\frac{x}{L}\right)\mathrm{d}x,\\
&= \begin{cases} \dfrac{x}{D}\left(1-\dfrac{x_0}{L}\right)\left(1-\dfrac{x}{L}\right)\mathrm{d}x & x < x_0,\\[2mm] \dfrac{x_0}{D}\left(1-\dfrac{x}{L}\right)^2 \mathrm{d}x & x > x_0. \end{cases}
\end{aligned}
\tag{6}
$$

The maximum residence time occurs at $x = x_0$ for $x_0 < L/2$ and then 'sticks' at $x = L/2$ for $x_0 \geqslant L/2$, with a cusp always occurring at $x = x_0$ (figure 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 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 that includes x_0. The result of these standard manipulations is

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

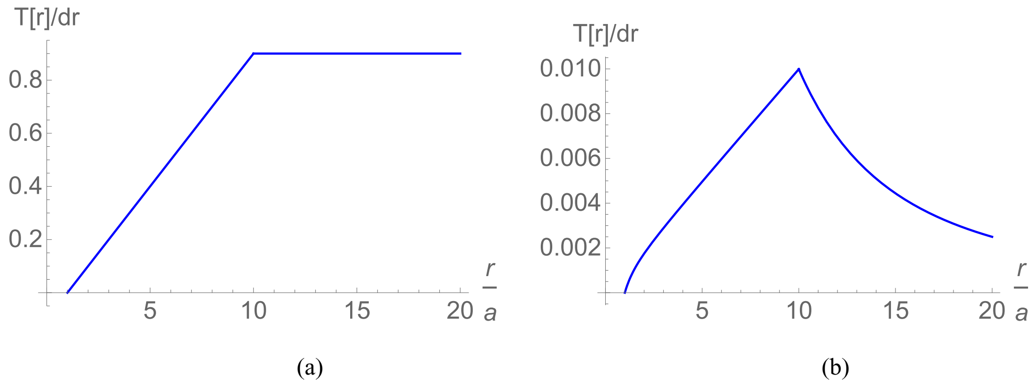

Figure 4. The residence time density $T(r)/\mathrm{d}r$ for a diffusing particle that starts at $r_0 = 10$ exterior to a sphere of radius 1 in: (a) $d = 3$ and (b) $d = 5$.

To obtain $T(r)$, the residence time in a shell of radius r and thickness $\mathrm{d}r$, we again need to multiply the above expression by the probability that the particle eventually hits the sphere, which, for $d > 2$, is simply $(a/r)^{d-2}$ [16]. Thus we have

$$T(r) = \Omega_d\, r^{d-1}\widetilde{P}(r)\left(\frac{a}{r}\right)^{d-2}\mathrm{d}r,$$

$$= \begin{cases} \dfrac{\mathrm{d}r}{D(d-2)}\left(\dfrac{a}{r}\right)^{d-2}\left[1-\left(\dfrac{a}{r}\right)^{d-2}\right]\dfrac{r^{d-1}}{r_0^{d-2}} & r < r_0, \\ \dfrac{\mathrm{d}r}{D(d-2)}\left(\dfrac{a}{r}\right)^{d-2}\left[1-\left(\dfrac{a}{r_0}\right)^{d-2}\right] r & r > r_0. \end{cases} \tag{9}$$

Two representative results are shown in figure 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 figure 4(b) for $d = 5$.

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

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

In distinction to the cases of $d = 1$ and $d > 2$, $\widetilde{P}(r)$ is constant for $r > r_0$. Since a diffusing particle always reaches the absorbing sphere in $d = 2$, we immediately have $T(r) = 2\pi r \widetilde{P}(r)\,\mathrm{d}r$.

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 $x_0 = 1$.

J. Stat. Mech. (2018) 103205

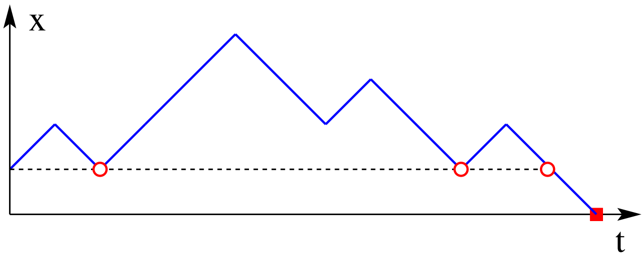

Figure 5. Space-time trajectory of a one-dimensional random walk of $2n+1=13$ steps that starts at $x=1$ and makes 3 revisits to $x=1$ (red circles) before being absorbed at $x=0$ (square). This path contributes to $A(6,3)$ (see Eq. (11)).

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 figure 5), and make k revisits to $x=1$, before dying at the $(2n+1)$st step. The number of such paths was found in [22] 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 [23, 24]. 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 [25], which counts the total number of random walks of $2n$ steps that start at $x=0$, remain in the region $x\geqslant 0$, 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 $2(n+1)$. Note the shift $n\to n+1$ to ensure that the walk always remains above $x=0$.

From equations (11) and (12), we have

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

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

J. Stat. Mech. (2018) 103205

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

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

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

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

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

Which was derived in [22] (see also [26]). In terms of the generating function, we immediately obtain the remarkably simple result

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

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 equation (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:

$$\begin{aligned} \langle k^2\rangle_n &= \frac{n(13n-1)}{(2+n)(3+n)}, \\ \langle k^3\rangle_n &= \frac{15n^2(5n-1)}{(2+n)(3+n)(4+n)}, \\ \langle k^4\rangle_n &= \frac{541n^2-196n^2+11n^2+4n}{(2+n)(3+n)(4+n)(5+n)}, \end{aligned} \tag{21a}$$

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:

J. Stat. Mech. (2018) 103205

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

etc. Parenthetically, these numbers are also sequence A000629 in the On-Line Encyclopedia of Integer Sequences [27]

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

$$\begin{aligned} G(x,y) = \sum_{n,k\geqslant 0} \mathcal{P}(n,k)\, x^n\, y^k &= \frac{1}{2}\big[1+g(x,y)\big], \\ &= \frac{1}{2}\left(1+\frac{xy}{2-xy+2\sqrt{1-x}}\right), \end{aligned} \tag{22}$$

where the term 1 in the parenthesis 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 visits 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 visits to $x=2$. To facilitate this calculation, it is useful to define the three-variable generating function

$$\mathcal{G}(x,y,z) = \sum_{n\geqslant 1}\ \sum_{1\leqslant \ell\leqslant n}\ \sum_{\ell\leqslant k\leqslant n} \mathcal{P}\left(n,k,\ell\right)\, x^n\, y^k\, z^\ell, \tag{23}$$

which encodes all paths according to (n,k,ℓ). We also label each successive excursion of the path above $x=1$ by the index $1\leqslant i\leqslant \ell$, 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 (figure 6). As shown in this figure, $2m_i$ counts the number of steps that lie above $x=2$. 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 figure 6):

J. Stat. Mech. (2018) 103205

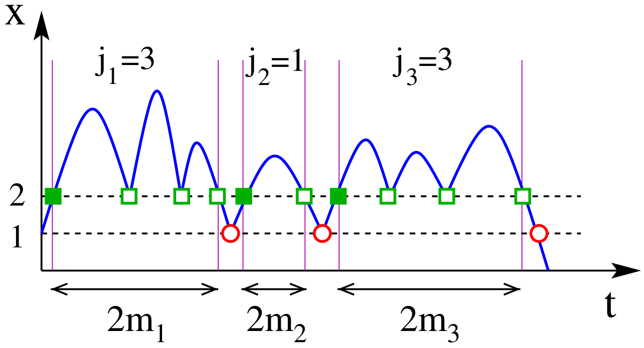

Figure 6. Schematic space-time trajectory of a random walk that starts at $x=1$ and has three revisits to $x=1$ (red circles) and $k=10$ visits to $x=2$ (green squares). There are $\ell=3$ excursions above $x=1$. Immediately after a revisit to $x=1$, the next revisit to $x=2$ is shown as a solid green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

$$\begin{aligned}
&j_1+j_2+\cdots+j_\ell+\ell=k,\\
&m_1+m_2+\cdots+m_l+\ell=n,\\
&j_i\leqslant m_i.
\end{aligned}\tag{24}$$

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 equation (22) as (see the appendix for details of this derivation)

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

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 $x=1$.

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

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

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. Substituting Equation (26) in (25) and comparing the resulting formula with the first line of (19), we obtain

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

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

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

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 $x>2$ 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 $x=1$. 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 section 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 figure 2(a). Clearly, our argument also applies for any starting point of the walk x_0, as long as we restrict to coordinates with $x>x_0+1$.

J. Stat. Mech. (2018) 103205

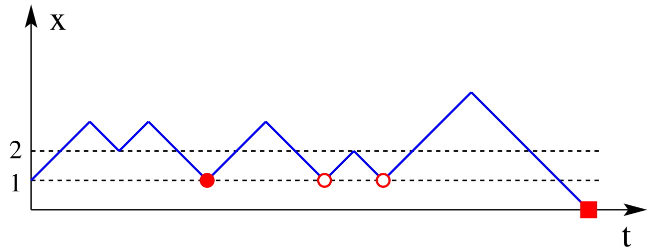

Figure 7. Space-time trajectory of a one-dimensional random walk of $2n+1=19$ that starts at $x=1$ and first revisits $x=1$ at step $2m=6$ (solid circle). Subsequent revisits to $x=1$ are indicated by open circles and the walk is absorbed when it first reaches $x=0$ (square).

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

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

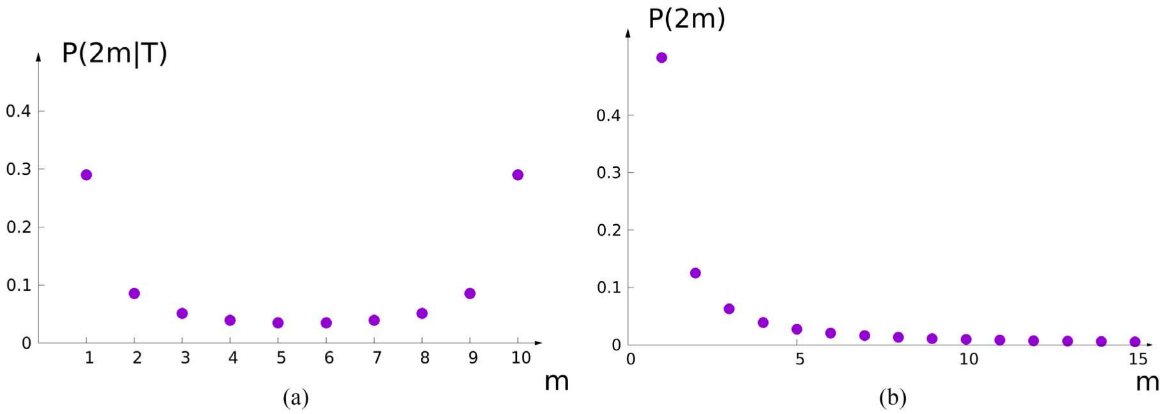

Figure 8. (a) Conditional distribution $P\,(\tau_1 = 2m\,|\,T = 2n+1)$ for $n = 10$. Note that typical (i.e. most likely) values of τ_1 are 2 and $2n$, while the average value is $\langle\tau_1\rangle = n+1$. (b) Distribution $P\,(\tau_1 = 2m)$ of the first revisit time to 1.

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

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

Because $\mathbf{P}\,(\tau_1 = 2m\,|\,T = 2n+1)$ is symmetric under $m \to n+1-m$, the average value of τ_1, the time of the first revisit, conditioned on $T = 2n+1$, can be immediately seen to be

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

The above result $\mathbf{P}\,(\tau_1 = 2m\,|\,T = 2n+1)$ 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 (figure 8(a)), just as in the arcsine law for the time of the last zero of a Brownian motion.

4.2. Time of first return to 1 for any *n*

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

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

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 $x = 2$):

J. Stat. Mech. (2018) 103205

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

This distribution is normalized, because $\sum_{m\geqslant 1} C_{m-1}/2^{2m-1} = 1$. 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, 16]. 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 $m = 1$ (figure 8(b)).

5. Summary

We showed how standard first-passage methods can be used to determine the average 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 average residence time in a given spatial range is simply the integral of the probability distribution over all time in this same range, with the given initial condition and the absorbing boundary condition. This perspective allowed us to also treat, in a relatively simple manner: (a) biased diffusion, (b) diffusion in a finite interval (conditioned on absorption at a given side of the interval), and (c) 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) in the distance between this point and the absorber. That is, a diffusing particle does not linger when it is close to an absorbing point. Another interesting feature, almost intuitive from the analogy with electrostatics, is the fact that, in low dimensions ($d \leqslant 3$), the average residence time at any point beyond the starting point is constant and simply equal to the average residence time at the starting point. This is no longer the case for $d \geqslant 4$, with an abrupt transition between $d = 3$ and $d = 4$. It would be of interest to understand why this transition occurs at a spatial dimension that differs from that of the well known transition between recurrence and transience, which happens at $d = 2$.

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 $x \geqslant 2$, 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

SR wishes to acknowledge the hospitality of Université Paris-1 Panthéon-Sorbonne for their support of a visit during which this project was initiated, as well as financial support from grant DMR-1608211 from the National Science Foundation. We also thank O Bénichou and S N Majumdar for useful discussions and advice.

Appendix. Relation between generating functions

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

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

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

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

$$\begin{aligned}\mathcal{G}(x,y,z) &= \sum_{n=1}^{\infty}\sum_{\ell=1}^{n}\sum_{k=\ell}^{n}\sum_{\substack{j_1+\cdots+j_\ell=k-\ell\\ m_1+\cdots+m_\ell=n-\ell\\ j_i\leqslant m_i}}P(\ell)\,\mathbf{P}\left(m_1,\ldots,m_\ell,j_1,\ldots,j_\ell|\ell\right)\,x^{m_1+\ldots m_\ell+\ell}\,y^{j_1+\ldots j_\ell+\ell}\,z^\ell \\ &= \sum_{n=1}^{\infty}\sum_{\ell=1}^{n}\sum_{\substack{m_1+\cdots+m_\ell=n-\ell\\ 0\leqslant j_i\leqslant m_i}}P(\ell)\,\mathbf{P}\left(m_1,\ldots,m_\ell,j_1,\ldots,j_\ell|\ell\right)\,x^{m_1+\ldots m_\ell+\ell}\,y^{j_1+\ldots j_\ell+\ell}\,z^\ell \\ &= \sum_{\ell\geqslant1}P(\ell)\sum_{n\geqslant\ell}\sum_{\substack{m_1+\cdots+m_\ell=n-\ell\\ 0\leqslant j_i\leqslant m_i}}\mathbf{P}\left(m_1,\ldots,m_\ell,j_1,\ldots,j_\ell|\ell\right)\,x^{m_1+\ldots m_\ell+\ell}\,y^{j_1+\ldots j_\ell+\ell}\,z^\ell \\ &= \sum_{\ell\geqslant1}P(\ell)\sum_{m_1,\ldots,m_\ell\geqslant0}\sum_{j_1,\ldots,j_\ell=0}^{m_1,\ldots,m_\ell}\left[\prod_{i=1}^{\ell}\mathbf{P}\left(m_i,j_i|\ell\right)\,x^{m_i}\,y^{k_i}\right](xyz)^\ell \\ &= \sum_{\ell\geqslant1}P(\ell)\,(xyz)^\ell\sum_{m_1,\ldots,m_\ell\geqslant0}\sum_{j_1,\ldots,j_\ell=0}^{m_1,\ldots,m_\ell}\left[\prod_{i=1}^{\ell}\mathbf{P}\left(m_i,j_i|\ell\right)\,x^{m_i}\,y^{j_i}\right] \\ &= \sum_{\ell\geqslant1}P(\ell)\,(xyz)^\ell\prod_{i=1}^{\ell}\left[\sum_{m_i\geqslant0}\sum_{0\leqslant j_i\leqslant m_i}\mathbf{P}\left(m_i,j_i|\ell\right)\,x^{m_i}\,y^{j_i}\right].\end{aligned} \tag{A.2}$$

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

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

Note that $G(x, y)$, as defined in equation (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.

Equation (A.2) becomes

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

This is equation (25) in the main text.

References

[1] Feller W 1964 *Introduction to Probability Theory and its Applications* 3rd edn (New York: Wiley)
[2] Spitzer F 1964 *Principles of Random Walk* (Princeton, NJ: Van Nostrand)
[3] Montroll E W and Weiss G H 1965 *J. Math. Phys.* **6** 167
[4] Mörters P and Peres Y 2010 *Brownian Motion* (Cambridge: Cambridge University Press)
[5] Chung K L 1976 *Ark. Mat.* **14** 155
[6] Baldassarri A, Colaiori F and Castellano C 2003 *Phys. Rev. Lett.* **90** 060601
[7] Colaiori F, Baldassarri A and Castellano C 2004 *Phys. Rev.* E **69** 041105
[8] Lévy P 1948 *Processus Stochastique et Mouvement Brownien* (Paris: Gauthier-Villars)
[9] Ray D 1963 *Illinois J. Math.* **7** 615 (https://projecteuclid.org/euclid.ijm/1255645099)
[10] Knight F B 1963 *Trans. Am. Math. Soc.* **109** 56
[11] Condamin S, Tejedor V and Bénichou O 2007 *Phys. Rev.* E **76** 050102
[12] Bénichou O and Desbois J 2009 *J. Phys. A: Math. Theor.* **42** 015004
[13] Bénichou O and Voituriez R 2014 *Phys. Rep.* **539** 225
[14] Agmon N 1984 *J. Chem. Phys.* **81** 3644
[15] Daniels H E 1982 *Ann. Stat.* **10** 394
[16] Redner S 2001 *A Guide to First-Passage Processes* (Cambridge: Cambridge University Press)
[17] Ferraro M and Zaninetti L 2001 *Phys. Rev.* E **64** 056107
[18] Ferraro M and Zaninetti L 2004 *Physica* A **338** 307
[19] Ferraro M and Zaninetti L 2006 *Phys. Rev.* E **73** 057102
[20] Galambos J 1987 *The Asymptotic Theory of Extreme Order Statistics* (Malabar, FL: Krieger)
[21] Krapivsky P L, Redner S and Ben-Naim E 2010 *A Kinetic View of Statistical Physics* (Cambridge: Cambridge University Press)
[22] Bhat U, Redner S and Bénichou O 2017 *J. Stat. Mech.* 073213
[23] Bailey D F 1996 *Math. Mag.* **69** 128
[24] Lee K-H and Oh S-J 2016 arXiv:1601.06685
[25] Stanley R P 2015 *Catalan Numbers* (Cambridge: Cambridge University Press)
[26] Stanley R P 1999 *Enumerative Combinatorics* vol 2 (Cambridge: Cambridge University Press)
[27] Sloane N J A 2007 The on-line encyclopedia of integer sequences *Towards mechanized mathematical assistants* 130 (Berlin: Springer) (http://www.research.att.com/~njas/sequences/)