ELSEVIER

Physica A 209 (1994) 361–368

PHYSICA A

The expected number of distinct sites visited by N biased random walks in one dimension

Hernan Larralde[a,b,1], George H. Weiss[b], H. Eugene Stanley[a]

[a] Department of Physics, Boston University, Boston, MA 02215, USA
[b] Physical Sciences Laboratory, Division of Computer Research and Technology, National Institutes of Health, Bethesda, MD 20892, USA

Received 20 April 1994

Abstract

We calculate the asymptotic form of the expected number of distinct sites visited by N random walkers moving independently in one dimension. It is shown that to lowest order and at long times, the leading term in the asymptotic result is that found for the random walk of a single biased particle, which implies that the bias is strong enough a factor to dominate the many-body effects in that regime. The lowest order correction term contains the many-body contribution. This is essentially the result for the unbiased random walk.

1. Introduction

Although the problem of calculating the expected number of distinct sites visited by an n-step random walk on a lattice first appeared in the mathematical literature, applications of different properties of this random variable abound in the literature of the physical sciences (cf., for example, the references cited in [1]). It allows, for example, a simple characterization of the amount of territory covered by a diffusing particle, and can be used to extend the Smoluchowski model for chemical reaction rates [2,3]. Similar problems, for example, are suggested by applications in the theory of the diffusion of photons in a turbid medium with applications to the use of optical methods to determine physical properties of human tissue [4,5]. In such applications it is useful to have a measure of how much tissue is explored by the laser-injected photons. A natural characterization of this quantity is the expected number of distinct sites visited by

[1] Present address: Cavendish Laboratory, Cambridge University, Maddingly Road, Cambridge CB3 0NE, UK.

Elsevier Science B.V.
SSDI 0378-4371(94)00103-Z

a random walker in a specified time. The analogous quantity in a diffusion model can be expressed in terms of the volume of the Wiener sausage, which is somewhat more difficult to work with than the related lattice quantity.

It is well known that quite difficult mathematical problems are posed in a determination of the actual distribution of the distinct number of sites visited by an n-step random walk. However, if attention is restricted to the *expected number* of distinct sites visited by a single random walker, a quantity to be denoted by $\langle S_1(n)\rangle$, then a considerable amount is known about this function because a generating function of this quantity can be calculated [6,7]. The application of Tauberian methods can then be used to infer the behavior of $\langle S_1(n)\rangle$. This is true only for the first moment; a determination of similar properties of higher moments requires the use of much more sophisticated mathematical methods [8], and, indeed, to date, information is available only on the asymptotic from of the first two moments (in three or more dimension the asymptotic form of the distribution is also known [8]). We have recently extended the analysis for finding $\langle S_1(n)\rangle$ to that of finding the analogous quantity for N non-interacting random walkers, a function which we denote by $\langle S_N(n)\rangle$ [9]. The behavior of this quantity is surprisingly rich when considered as a function of the two variables n and N. The theory in [9] was developed for unbiased random walks whose second moments of individual steps are finite. In the present paper we consider the extension of that analysis for biased random walks in one dimension. Even in this seemingly simple case we will see that the mathematical formalism required to calculate the asymptotic behavior of $\langle S_N(n)\rangle$ can be quite complicated, although our final result is not too surprising when considered on purely intuitive grounds.

2. General formalism

The general formalism developed in [9] remains the starting point for the present analysis, with differences appearing only in the detailed form of the generating functions to be introduced in the following paragraphs. In order to make the exposition self-contained we summarize the principal features of the formalism.

Assume that all of the random walks start initially from the origin and let $f_n(\boldsymbol{r})$ denote the probability that a single random walker reaches site $\boldsymbol{r}$ for the first time at step n, i.e., it is the first passage time probability to $\boldsymbol{r}$. The generating function of the $f_n(\boldsymbol{r})$ with respect to n can be calculated in terms of the characteristic function, a relation of which considerable use will be made in the analysis to follow. The probability that $\boldsymbol{r}$ has not been visited by step n will be denoted by $\Gamma_n(\boldsymbol{r})$, which is related to the $f_n(\boldsymbol{r})$ by

$$\Gamma_n(\boldsymbol{r}) = 1 - \sum_{j=0}^{n} f_j(\boldsymbol{r})\,. \tag{1}$$

Finally, denote the generating function of $\langle S_N(n)\rangle$ with respect to N by $S(u;n)$, u being the transform parameter. It was shown in [9] that $S(u;n)$ can be written in terms of $\Gamma_n(r)$ as

$$S(u;n)=\frac{u}{1-u}\sum_r \frac{1-\Gamma_n(r)}{1-u\Gamma_n(r)}\,. \tag{2}$$

This result is quite general, and applies to all independent sets of random walks regardless of whether the random walk is biased. The strategy for calculating $\langle S_N(n)\rangle$ will be to first determine the analytic behavior of the generating function $S(u;n)$ in the neighborhood of $u=1$. By appealing to a Tauberian theorem we may then infer the behavior of $\langle S_N(n)\rangle$ in the limit of large N.

If a walker can only reach a finite number of sites in a single step then define $\Omega_N(n)$ to be the number of sites reachable by N random walkers in n steps. It is clear that when n is fixed $\lim_{N\to\infty}\langle S_N(n)\rangle=\Omega_N(n)$ since in this limit all possible trajectories will be traversed. Next we consider the limit $n\to\infty$ when $N\gg 1$. Let $p(j)$ be the probability that a single random walker is displaced by j lattice sites in a single step, let $p_n(j)$ be the probability that the displacement in n steps is equal to j, and let $\hat{p}(\theta)$ be the characteristic function defined by

$$\hat{p}(\theta)=\sum_{j=-\infty}^{\infty} p(j)\exp(\mathrm{i}j\theta)\,. \tag{3}$$

Let $p(j;z)$ be the generating function of the $p_n(j)$ with respect to the variable n. This function can be represented in terms of the characteristic function as

$$p(j;z)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{\mathrm{e}^{-\mathrm{i}j\theta}}{1-z\hat{p}(\theta)}\,\mathrm{d}\theta\,. \tag{4}$$

The importance of $p(j;z)$ lies in the fact that the corresponding generating function for the $\{f_n(j)\}$ is related to it by

$$f(j;z)=p(j;z)/p(0;z)\,,\qquad j\neq 0\,. \tag{5}$$

Only the large-n form of $f_n(j)$ will be required in our analysis of properties of $S(u;n)$.

Since our interest lies in the large-n regime, we need only consider properties of $f(j;z)$ in the singularity that occurs when the limit $z\to 1$ is taken. The characteristic function has the property $\hat{p}(0)=1$ which suggests that the singularity can only be due to the behavior of the integrand of Eq. (4) at $\theta=0$. In order to find this behavior it suffices to expand $\hat{p}(\theta)$ around $\theta=0$ and calculate the effect of the singularity.

We will consider only the case in which the first moment and the variance of the displacement are finite. These parameters will be denoted by $\langle j\rangle=\mu$ and $\langle j^2\rangle-\langle j\rangle^2=\sigma^2$ respectively. The assumption that these quantities are finite allows us to expand $\hat{p}(\theta)$ to second order in the neighborhood of $\theta=0$ as

$$\hat{p}(\theta) \sim 1 + \mathrm{i}\mu\theta - \tfrac{1}{2}\sigma^2\theta^2 \,, \tag{6}$$

which, in turn, leads to the approximation

$$p(j;z) \sim \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} \frac{\mathrm{e}^{-\mathrm{i}j\theta}}{1 - z - \mathrm{i}\mu\theta + \frac{1}{2}\sigma^2\theta^2}\, \mathrm{d}\theta \,, \tag{7}$$

where the limits of integration have been extended to $\pm\infty$ because the behavior of $p(j;z)$ near $z = 1$ is mainly determined by the behavior of denominator in the neighborhood of $\theta = 0$. The integral in this last equation is most simply evaluated by exponentiating the denominator using the identity $u^{-1} = \int_0^\infty \exp(-ut)\,\mathrm{d}t$ and interchanging the orders of integration. The integral with respect to θ can then be evaluated in closed form. In order to most succinctly represent the results we will use units normalized to the value of σ by using the notation

$$\nu = \mu/\sigma \,, \qquad \rho = j/\sigma \,, \tag{8}$$

in which case we find that

$$p(j;z) \sim \frac{1}{\sqrt{2(1-z) + \nu^2}} \exp\{\nu\rho - |\rho|\sqrt{2(1-z) + \nu^2}\} \,. \tag{9}$$

While it is not strictly possible to invert the generating function in this equation we can find its inverse in the limit $n \to \infty$ by regarding it as a Laplace transform. That is to say, we replace $1 - z$ by a Laplace transform parameter s (remembering that we are only interested in $z \sim 1$ or the equivalent $s \sim 0$) which allows an exact inversion of the transform. In such a case the parameter ρ takes on a continuum of values and the function $\Gamma_n(\rho)$ is to be replaced by $\Gamma_t(\rho)$ where t is now to be interpreted as a continuous variable. Let $\hat{\Gamma}_s(\rho)$ denote the transform of $\Gamma_t(\rho)$. This function is explicitly equal to

$$\hat{\Gamma}_s(\rho) \sim \frac{1}{s}[1 - \exp\{\nu\rho - |\rho|\sqrt{2s + \nu^2}\}] \,, \tag{10}$$

whose inverse transform is

$$\Gamma_t(\rho) = 1 - \frac{|\rho|}{\sqrt{2\pi}} \int\limits_0^t \exp\left[-\frac{1}{2}\left(\frac{\rho}{\sqrt{\tau}} - \nu\sqrt{\tau}\right)^2\right] \frac{\mathrm{d}\tau}{\tau^{3/2}} \,. \tag{11}$$

Notice that, due to the presence of ν in the exponential that appears in the integrand, $\Gamma_t(\rho)$ is not symmetric with respect to an interchange in the sign of ρ. In consequence, the behavior of the integral depends on whether ρ is positive or negative. In particular, one expects that when $\nu > 0$ the value of the integral in Eq. (11) will be much smaller when $\rho < 0$ than when $\rho > 0$. This reflects the fact that the bias is towards the direction of increasing x.

To evaluate the integral in Eq. (11) we first replace the integrand by a more convenient representation of the exponential functions appearing in it. For this purpose consider the functions $\varphi_+(\tau)$ and $\varphi_-(\tau)$ defined by

$$\varphi_+(\tau) = \operatorname{erf}\left(\nu\sqrt{\frac{\tau}{2}} + \frac{\rho}{\sqrt{2\tau}}\right), \qquad \varphi_-(\tau) = \operatorname{erf}\left(\nu\sqrt{\frac{\tau}{2}} - \frac{\rho}{\sqrt{2\tau}}\right), \tag{12}$$

whose derivatives are

$$\begin{aligned} \frac{\mathrm{d}\varphi_+}{\mathrm{d}\tau} &= \frac{1}{\sqrt{2\pi}}\left(\frac{\nu}{\tau^{1/2}} - \frac{\rho}{\tau^{3/2}}\right)\exp\left\{-\frac{1}{2}\left(\nu\sqrt{\tau} + \frac{\rho}{\sqrt{\tau}}\right)^2\right\}, \\ \frac{\mathrm{d}\varphi_-}{\mathrm{d}\tau} &+ \frac{1}{\sqrt{2\pi}}\left(\frac{\nu}{\tau^{1/2}} + \frac{\rho}{\tau^{3/2}}\right)\exp\left\{-\frac{1}{2}\left(\nu\sqrt{\tau} + \frac{\rho}{\sqrt{\tau}}\right)^2\right\}. \end{aligned} \tag{13}$$

Another identity required for our analysis is the algebraic relation

$$\frac{1}{2}\left(\nu\sqrt{\tau} + \frac{\rho}{\sqrt{\tau}}\right)^2 = \frac{1}{2}\left(\nu\sqrt{\tau} - \frac{\rho}{\sqrt{\tau}}\right)^2 + 2\nu\rho\,. \tag{14}$$

Therefore we may express the integrand in Eq. (11) in terms of a linear combination of the derivatives of φ_+ and φ_- as

$$\frac{\mathrm{d}\varphi}{\mathrm{d}\tau} - \mathrm{e}^{2\mu\rho}\frac{\mathrm{d}\varphi}{\mathrm{d}\tau} = \rho\sqrt{\frac{2}{\pi\tau^3}}\exp\left\{-\frac{1}{2}\left(\nu\sqrt{\tau} - \frac{\rho}{\sqrt{\tau}}\right)^2\right\}. \tag{15}$$

Hence, on solving for the integrand in Eq. (11) from the formula in this last equation and substituting the result back into the integral, which can now be evaluated in closed form we find that when $\rho > 0$, $\Gamma_t(\rho)$ is equal to

$$\Gamma_t^{(+)}(\rho) = 1 - \frac{1}{2}\left[\operatorname{erfc}\left(\frac{\rho}{\sqrt{2t}} - \nu\sqrt{\frac{t}{2}}\right) + \mathrm{e}^{2\mu\rho}\operatorname{erfc}\left(\frac{\rho}{\sqrt{2t}} + \nu\sqrt{\frac{t}{2}}\right)\right]. \tag{16}$$

A similar argument can be used to show that when $\rho < 0$

$$\Gamma_t^{(-)}(\rho) = 1 - \frac{1}{2}\left[\mathrm{e}^{-2\mu|\rho|}\operatorname{erfc}\left(\frac{|\rho|}{\sqrt{2t}} - \nu\sqrt{\frac{t}{2}}\right) + \operatorname{erfc}\left(\frac{|\rho|}{\sqrt{2t}} + \nu\sqrt{\frac{t}{2}}\right)\right]. \tag{17}$$

In the large-t regime only large values of $|\rho|$ will contribute significantly to the value of the integral. This allows us, as in [9], to replace the sum in Eq. (2) by an integral with respect to ρ, which will be decomposed into two contributions, one for $\rho > 0$ and the second for $\rho < 0$:

$$\begin{aligned} S(u;t) &\sim \frac{1}{1-u}\left\{\int_0^\infty \frac{1-\Gamma_t^{(+)}(\rho)}{1-u\Gamma_t^{(+)}(\rho)}\,\mathrm{d}\rho + \int_{-\infty}^0 \frac{1-\Gamma_t^{(-)}(\rho)}{1-u\Gamma_t^{(-)}(\rho)}\,\mathrm{d}\rho\right\} \\ &\equiv S_+(u;t) + S_-(u;t)\,. \end{aligned} \tag{18}$$

The region of interest in u-space that corresponds to large N is $u \sim 1$. Consider first the term $S_+(u;t)$, making a change of variable from ρ to

$$\omega = \frac{\rho}{\nu t}\,. \tag{19}$$

The integral defining $S_+(u;t)$ is then expressed as

$$S_+(u;t) \sim \frac{\nu t}{1-u} \int_0^\infty \frac{G_t(\omega)}{1-u+G_t(\omega)} \, d\omega \,, \tag{20}$$

where now

$$G_t(\omega) = \frac{1}{2}\left\{\operatorname{erfc}\left[\nu\sqrt{\frac{t}{2}}(\omega - 1)\right] + e^{2\mu^2 t\omega} \operatorname{erfc}\left[\nu\sqrt{\frac{t}{2}}(\omega + 1)\right]\right\}. \tag{21}$$

It is readily shown that $S_-(u;t)$ has the form

$$S_-(u;t) = \frac{\nu t}{1-u} \int_0^\infty \frac{1-\Gamma_t^{(-)}(-\omega)}{1-u\Gamma_t^{(-)}(-\omega)} \, d\omega \tag{22}$$

in which $\Gamma_t^{(-)}(-\omega)$ can be written in terms of $\Gamma_t^{(+)}(\omega)$ as

$$\Gamma_t^{(-)}(-\omega) = 1 - e^{-2\mu^2 t}[1 - \Gamma_t^{(+)}(\omega)] \,. \tag{23}$$

This identity allows us to write the sum of the two contributions to $S(u;t)$ as

$$S(u,t) \sim \frac{\nu t}{1-u} \int_0^\infty \left\{ \frac{G_t(\omega)}{1-u+G_t(\omega)} + \frac{G_t(\omega)}{1+e^{2\mu^2 t\omega}[1-u+G_t(\omega)]} \right\} d\omega \,. \tag{24}$$

It is clear that setting $u = 1$ in this representation causes the first integral on the right side to diverge but not to second, since $\lim_{t\to\infty} G_t(\omega) = 0$. This limiting value is reached in such a way that

$$\int_0^\infty \frac{G_t(\omega)}{1+e^{2\mu^2 t\omega}\, G_t(\omega)} \, d\omega = U(t) < \infty \,. \tag{25}$$

Since $G_t(\omega)$ decreases monotonically with t it follows that $\lim_{t\to\infty} U(t) = 0$, which implies that any singular behavior arising from the integral is necessarily due to the first term on the right-hand side of Eq. (24). In what follows we therefore ignore the contribution from $U(t)$ in comparison with that term and focus on the behavior of $S_+(u;t)$ for large values of t.

In order to find this limiting behavior we observe that $\operatorname{erfc}(-\infty) = 2$ and $\operatorname{erfc}(\infty) = 0$. Hence the asymptotic behavior of $S_+(u;t)$ depends on whether ω is less than or greater than 1, i.e., it approaches 2 in the former case and 0 in the latter. This is not true of the second term because the argument of the complementary error function in that term remains positive for all values of ω. This suggests a decomposition of the range of integration $(0,\infty)$ into two segments as $(0,1) + (1,\infty)$.

Consider the behavior of the integral in the first interval, that is, in $(0,1)$. On fixing ω, making use of the asymptotic result

$$\operatorname{erfc}(\Omega) \sim \frac{e^{-\Omega^2}}{\sqrt{\pi}\,\Omega}, \qquad \Omega \gg 1 \,, \tag{26}$$

and letting t increase, we find that $G_t(\omega)$ can be approximated by

$$G_t(\omega) \sim 1 - \frac{1}{\nu}\sqrt{\frac{2}{t}}\,\frac{e^{-\mu^2 t/2}}{\omega^2 - 1}\,. \tag{27}$$

Since we may set $u = 1$ in the second term appearing in the integrand of Eq. (24) without causing the integral to diverge it follows that the contribution to $S(u; t)$ from the range $(0, 1)$ is equal to $\nu t/(1-u)$. This corresponds to a contribution to $\langle S_N(n)\rangle$ which is equal to νn in the limit $n \to \infty$. This is not a surprising result since the bias effectively reduces some of the fluctuations inherent in the random walk with the result that many-body effects do not show up, at least in the lowest order approximation.

Any many-body effects in lowest order must therefore be due to the contribution from the range $(1, \infty)$. Let us fix ω and let $t \to \infty$. In this limit the function $G_t(\omega)$ has the limiting form

$$G_t(\omega) \sim \frac{1}{\nu}\sqrt{\frac{2}{\pi t}}\,\frac{\omega}{\omega^2 - 1}\exp\{-\tfrac{1}{2}\omega^2 t(\omega - 1)^2\}\,. \tag{28}$$

This can be translated into an approximation for $S(u; t)$ which, after changing the variable of integration to $v = \omega - 1$ and defining a parameter ϵ as

$$\epsilon = \nu(\tfrac{1}{2}\pi)^{1/2}(1 - u) \tag{29}$$

we can express $\mathcal{S}(u; t) \equiv S(u; t) - \mu t$ as

$$\begin{aligned} \mathcal{S}(u; t) &\sim \frac{t}{1-u}\int_0^\infty \frac{dv}{\epsilon t^{1/2} v\left(\dfrac{v+2}{v+1}\right) e^{(\mu^2 t/2)v^2} + 1} \\ &= \frac{\sqrt{2t}}{1-u}\int_0^\infty \frac{d\rho}{\epsilon \nu^{-1}\sqrt{2}\rho\left(\dfrac{\rho + \nu\sqrt{2t}}{\rho + \nu\sqrt{t/2}}\right) e^{\rho^2} + 1}\,. \end{aligned} \tag{30}$$

Let us keep in mind that in the large N limit with t fixed but large the value of the integral will be dominated by behavior of the integrand at large ρ. This being the case we can simplify Eq. (24) still further to

$$\mathcal{S}(u; t) \sim \frac{\sqrt{2t}}{1-u}\int_0^\infty \frac{d\rho}{\epsilon'\rho\, e^{\rho^2} + 1}\,, \tag{31}$$

in which we have set $\epsilon' = \sqrt{2}(\epsilon/\nu)$. The analytic behavior of this integral has been determined in [9]. On inserting the results of that analysis into this last expression we find that

$$\mathcal{S}(u; t) \sim \sqrt{\frac{t}{2}}\,\frac{1}{1-u}\sqrt{\ln\left(\frac{1}{1-u}\right)}\,, \tag{32}$$

which implies the asymptotic expression

$$\langle S_N(t)\rangle - \mu t \sim \sqrt{\tfrac{1}{2}t\ln(N)}\,. \tag{33}$$

The expression on the right-hand side of this equation exhibits essentially the same dependence on N that was found in the case of the unbiassed random walk and is independent of ν. The final result therefore consists of the sum of a dominant term which represents the contribution due to bias alone, which has shown to be independent of the number of random walkers, and a second term which reflects the many-body contribution.

The results obtained so far indicate that when the displacement probabilities include a bias there is a crossover in one dimension at $t = \mathrm{O}(\ln(N))$. At very long times the bias will be the predominant factor in determining the expected number of distinct sites visited while at shorter (but nevertheless long!) times the many-body effects can be more significant than the bias. It is also possible that the constant μ may depend weakly on N in $d > 1$ but have not so far carried out the much more complicated calculations required to verify this point in detail.

This work was supported in part by NSF and ONR.

References

[1] F. den Hollander and G.H. Weiss, in: An Introduction to Some Contemporary Problems in Statistical Physics, ed. G.H. Weiss (SIAM, Philadelphia, 1994);
G.H. Weiss, Aspects and Applications of the Random Walk (Elsevier, Amsterdam, 1994); in: Fractals in Science, A. Bunde and S. Havlin, eds. (Springer, Berlin, 1994), pp. 119–162.
[2] M. von Smoluchowski, Z. Phys. Chem. 29 (1917) 129.
[3] S.A. Rice, Diffusion-Limited Reactions (Elsevier, Amsterdam, 1985).
[4] R.F. Bonner, S. Havlin, R.J. Nossal and G.H. Weiss, J. Opt. Soc. Am. 4 (1987) 423.
[5] A.H. Gandjbakhche and G.H. Weiss, in: Progress in Optics, ed. E. Wolf, in press.
[6] E.W. Montroll, in: Stochastic Processes in Mathematical Physics and Engineering, Proc. Symp. Appl. Math. XVI (Am. Math. Soc., Philadelphia, 1964).
[7] E.W. Montroll and G.H. Weiss, J. Math. Phys. 6 (1965) 167.
[8] N.C. Jain and W.E. Pruitt, in: Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, vol. III (Univ. of Calif., Berkeley, 1970) p. 31.
[9] H. Larralde, P. Trunfio, S. Havlin, H.E. Stanley and G.H. Weiss, Phys. Rev. A 45 (1992) 7128; Nature 355 (1992) 423.