Computing the average number of distinct sites visited by a single t-step random walker on a d-dimensional lattice, denoted by $S_t(t)$, is by now a classic problem with a variety of applications ranging from the annealing of defects in crystals to the size of the territory covered by a diffusing animal during the foraging period. First posed and studied by Dvoretzky and Erdős in 1951 [1], this problem was solved exactly in a number of papers in the 1960s [2,3]. It is well established (see [4] for a review) that asymptotically for large $t$, $S_t(t) \sim t^{d/2}$ for $d < 2$, $\sim t^{d/(d-1)}$ for $2 < d < d_c(N)$, $\sim t^\nu$ where the exponent $\nu = N - d(N-1)/2$ varies with $N$ and $d$. For $d > d_c(N)$, $W_{\nu}(t) \to \text{const as } t \to \infty$. Exactly at the critical dimensions there are logarithmic corrections: for $d = 2$, we get $W_{\nu}(t) \sim t/\ln(t)^{\nu}$, while for $d = d_c(N)$, $W_{\nu}(t) \sim t$ for large $t$. Our analytical predictions are verified in numerical simulations.

We compute analytically the mean number of common sites, $W_{\nu}(t)$, visited by $N$ independent random walkers each of length $t$ and all starting at the origin at $t = 0$ in $d$ dimensions. We show that in the $(N-d)$ plane, there are three distinct regimes for the asymptotic large-$t$ growth of $W_{\nu}(t)$. These three regimes are separated by two critical lines $d = 2$ and $d = d_c(N) = 2N/(N-1)$ in the $(N-d)$ plane. For $d < 2$, $W_{\nu}(t) \sim t^{d/2}$ for large $t$ (the N dependence is only in the prefactor). For $2 < d < d_c(N)$, $W_{\nu}(t) \sim t^\nu$ where the exponent $\nu = N - d(N-1)/2$ varies with $N$ and $d$. For $d > d_c(N)$, $W_{\nu}(t) \to \text{const as } t \to \infty$. Exactly at the critical dimensions there are logarithmic corrections: for $d = 2$, we get $W_{\nu}(t) \sim t/\ln(t)^{\nu}$, while for $d = d_c(N)$, $W_{\nu}(t) \sim t$ for large $t$. Our analytical predictions are verified in numerical simulations.

DOIs: 10.1103/PhysRevE.86.021135 PACS number(s): 05.40.Fb, 02.50.Cw, 05.40.Jc, 24.60.–k

The statistics of the number of most popular sites, i.e., the sites visited by all the walkers, arises quite naturally in a number of contexts such as sociology, ecology, artificial networks (e.g., the internet, transport, and engineering networks), and polymer networks just to name a few. For example, in a multiple-user network such as the internet, the most popular “hub” sites visited by all the users are known to play a very important role in the dissemination of information [27]. The knowledge of how many of them are there fundamental for many applications. In the tourism industry, it is important to know the number of the most popular sites in a given area or city that are visited by all the tourists. Motivated by this general question, in this paper we study the statistics of the number of most popular sites in perhaps the simplest model, namely, for $N$ independent random walkers in the $d$-dimensional space, and show that even in this simple model, the asymptotic temporal growth of the mean number of common sites frequented by all $N$ walkers exhibits surprisingly rich behavior. We show that our results also have close connections to the probability of nonintersection of random walks studied in the mathematics literature [28,29]. Given the abundance of random walks used as fundamental models to study numerous natural and artificial systems, and the richness of our exact results, we believe that they will be useful in more specific applications in the future.

We consider $N$ independent t-step walkers on a d-dimensional lattice, each starting at the origin. To compute the number of common sites visited by all the N walkers, it is first useful to introduce a binary random variable $\sigma_{k,N}(\vec{x},t)$ associated with each site $\vec{x}$ such that $\sigma_{k,N}(\vec{x},t) = 1$ if the site...
\[ \text{FIG. 1.} \text{ (Color online) A realization of three random walks, each of six steps [denoted respectively by solid (blue), dashed (purple), and dash-dotted (green) lines] on a square lattice, all starting at the origin O. There are two sites [each marked by a filled (red) circle] that are visited by all three walkers.} \]

\[ x \text{ is visited by exactly } k \text{ of the } N \text{ walkers and } \sigma_{k,N}(x,t) = 0 \text{ otherwise. Then the sum } V_{k,N}(t) = \sum \sigma_{k,N}(x,t) \text{ represents the number of sites visited by exactly } k \text{ of the } N \text{ walkers, each of } t \text{ steps, in a particular realization of the walks. Clearly, } V_{k,N}(t) \text{ is a random variable that fluctuates from one sample to another. Taking the average gives the mean number of sites visited by exactly } k \text{ walkers, } \langle V_{k,N}(t) \rangle = \sum \sigma_{k,N}(x,t), \text{ where } P_{k,N}(x,t) = (\sigma_{k,N}(x,t)) \text{ is the probability that the site } x \text{ is visited by exactly } k \text{ of the } N \text{ walkers. Since the walkers are independent, one can write} \]

\[ P_{k,N}(x,t) = \left( \begin{array}{c} N \\ k \end{array} \right) [p(x,t)]^k [1 - p(x,t)]^{N-k}, \]

where \( p(x,t) \) is the probability that the site \( x \) is visited by a single \( t \)-step walker starting at the origin. Thus,

\[ \langle V_{k,N}(t) \rangle = \left( \begin{array}{c} N \\ k \end{array} \right) \sum x [p(x,t)]^k [1 - p(x,t)]^{N-k}. \]

Finally, the mean number of common sites visited by all the \( N \) walkers is simply

\[ W_N(t) = \langle V_{N,N}(t) \rangle = \sum x [p(x,t)]^N. \]

Hence, once the basic quantity \( p(x,t) \) for a single walker is known, we can determine \( \langle V_{k,N}(t) \rangle \) and, in particular, \( W_N(t) \) just by summing over all sites as in Eq. (4). Note that, by definition, \( p(0,t) = 1 \) for all \( t \geq 0 \) since the walker starts at the origin.

The probability \( p(x,t) \) can be fully determined for a lattice walker with discrete time steps using the standard generating function technique [3]. However, since we are interested here mainly in the asymptotic large-\( t \) regime, it is much easier to work directly in the continuum limit where we treat both space \( x \) and time \( t \) as continuous variables. Consider then a single Brownian motion of length \( r \) and diffusion constant \( D \) in \( d \) dimensions starting at the origin. We are interested in \( p(x,t) \), the probability that the site \( x \) is visited (at least once) by the walker up to time \( t \). Let \( r \) denote the last time before \( t \) that the site \( x \) was visited by the walker. Then, clearly,

\[ p(x,t) = \int_0^t G(x,\tau) q(t - \tau) d\tau, \]

where \( G(x,\tau) = e^{-x^2/(4Dt)}/(4\pi Dt)^{d/2} \) (where \( x = |x| \)) is the standard Green’s function denoting the probability that the particle is at \( x \) at time \( \tau \) and \( q(\tau) \) denotes the persistence, i.e., the probability that, starting at \( x \), the walker does not return to its starting point up to time \( \tau \). Note that \( q(\tau) \) does not depend on the starting point \( x \) and is the same as the probability of no return to the origin up to time \( \tau \). Indeed, \( q(\tau) = \int_0^\infty f(t) d\tau \), where \( f(t) = -d\sigma/d\tau \) is the standard first-passage probability to the origin [30].

The no-return probability \( q(t) \) for a Brownian walker has been studied extensively, and it is well known that for large \( t \), \( q(t) \sim t^{-d/2-1} \) for \( d < 2 \), and \( q(t) \sim 1/\ln t \) for \( d = 2 \), while it approaches a constant for \( d > 2 \) since the walker can escape to infinity with a finite probability for \( d > 2 \) [30]. One can show that to analyze the large-\( t \) behavior of \( p(x,t) \) in Eq. (5) in the scaling regime where \( x \to \infty \), \( t \to \infty \) but keeping \( x/\sqrt{t} \) fixed, it suffices to substitute only the asymptotic behavior of \( q(t) \) in Eq. (5). This gives, for large \( t \)

\[ p(x,t) \sim \int_0^t G(x,\tau)(t - \tau)^{-d/2-1} d\tau \quad \text{for } d < 2, \]

\[ p(x,t) \sim \int_0^t G(x,\tau) d\tau \quad \text{for } d > 2, \]

where we have dropped unimportant constants for convenience. For \( d = 2 \), \( p(x,t) \sim \int_0^t G(x,\tau) d\tau /\ln(t - \tau) \). Substituting the exact Green’s function \( G(x,\tau) = e^{-x^2/(4Dt)}/(4\pi Dt)^{d/2} \) one finds that \( p(x,t) \) has the following asymptotic scaling behavior:

\[ p(x,t) \sim f_{\leq}(\frac{x}{\sqrt{4Dt}}) \quad \text{for } d < 2, \]

\[ p(x,t) \sim t^{1-d/2} f_{>}(\frac{x}{\sqrt{4Dt}}) \quad \text{for } d > 2, \]
where the scaling functions for $d < 2$ and $d > 2$ can be expressed explicitly as

$$ f_<(z) = \int_0^1 e^{-z u} u^{d/2} (1 - u^{d/2-1}) du, $$

$$ f_>(z) = \int_0^1 e^{-z u} u^{d/2} du. $$

Equation (10) can be analytically continued to real the sum by an integral over space. Note that even though Eqs. (8) and (9) do not, in general, hold for very small $x$, the scaling regime can actually be extended all the way to $d < 2$, the scaling regime can actually be extended all the way to $d > 2$, the scaling regime can actually be extended all the way to $d > 2$. The scaling functions for $d > 2$ are

$$ f_<(z) \approx z^{-d} e^{-z} \quad \text{as} \quad z \to 0, $$

$$ f_>(z) \approx z^{-(d-2)} e^{-z} \quad \text{as} \quad z \to \infty. $$

At $d = 2$, one finds $f_2(z) \sim -2 \ln(z)$ as $z \to 0$, and $f_2(z) \sim e^{-z^2}/z^2$ as $z \to \infty$. Note that the scaling forms postulated in Eqs. (8) and (9) do not, in general, hold for very small $x$. For $d < 2$, the scaling regime can actually be extended all the way to $d > 2$. The scaling regime can actually be extended all the way to $d > 2$. The scaling regime can actually be extended all the way to $d > 2$. The scaling regime can actually be extended all the way to $d > 2$. The scaling regime can actually be extended all the way to $d > 2$. The scaling regime can actually be extended all the way to $d > 2$. The scaling regime can actually be extended all the way to $d > 2$. The scaling regime can actually be extended all the way to $d > 2$. The scaling regime can actually be extended all the way to $d > 2$. The scaling regime can actually be extended all the way to $d > 2$. The scaling regime can actually be extended all the way to $d > 2$. The scaling regime can actually be extended all the way to $d > 2$. The scaling regime can actually be extended all the way to $d > 2$. The scaling regime can actually be extended all the way to $d > 2$. The scaling regime can actually be extended all the way to $d > 2$. The scaling regime can actually be extended all the way to $d > 2$. The scaling regime can actually be extended all the way to $d > 2$. The scaling regime can actually be extended all the way to $d > 2$. The scaling regime can actually be extended all the way to $d > 2$. The scaling regime can actually be extended all the way to $d > 2$. The scaling regime can actually be extended all the way to $d > 2$. The scaling regime can actually be extended all the way to $d > 2$. The scaling regime can actually be extended all the way to $d > 2. $
walkers start at the origin, clearly the number of common sites visited must be at least 1, implying $V_{N,N}(t) \geq 1$. When $V_{N,N}(t) = 1$, it corresponds to the event that the walkers do not intersect further up to step $t$ and the origin at $t = 0$ remains the only site visited by all of them up to step $t$. Thus, the probability of no further intersection up to step $t$ is $F_N(t) = \Pr[V_{N,N}(t) = 1]$. Lawler studied the decay of $F_N(t)$ for large $t$ rigorously in special cases [28], and Duplantier showed [29] that $F_N(t)$ approaches a constant as $t \to \infty$ for $d > d_c(N) = 2N/(N-1)$. For $d < d_c(N)$, $F_N(t) \sim t^{-\xi}$ and the exponent $\xi$ was computed using an $\epsilon$ expansion around the critical dimension [29]. In contrast, in this paper we have computed the mean of the random variable $V_{N,N}(t)$, i.e., $W_N(t) = \langle V_{N,N}(t) \rangle$. Note that while $F_N(t)$ is not exactly computable in all $d$, $W_N(t)$ is, as we have shown here.

Another interesting related problem is to compute the mean number of $N$-fold self-intersections of a single ideal polymer chain of length $t$. In Ref. [31], it was stated that in $d = 3$ this grows as $t^{(3-N)/2}$, which looks similar to our result $W_N(t) \sim t^{(3-N)/2}$ in the intermediate phase in $d = 3$ and for $1 < N < 3$. However, the two problems are not exactly identical, and even the single-chain result in Ref. [31] was qualitatively argued for, not rigorously proved, and the logarithmic correction for $N = 3$ was not mentioned.

In summary, we have presented exact asymptotic results for the mean number of common sites $W_N(t)$ visited by $N$ independent random walkers in $d$ dimensions. We have shown that, as a function of $N$ and $d$ in the $(N,d)$ plane, there are three distinct regimes for the growth of $W_N(t)$, including, in particular, an anomalous intermediate regime $2 < d < d_c(N) = 2N/(N-1)$.

We conclude with a few additional remarks. In this paper, we have computed analytically the scaling behavior of $p(\vec{x},t)$, the probability that the site $\vec{x}$ is visited by a single $t$-step walker. This result turns out to be the key ingredient to address other related questions. For instance, it would be easy to compute the mean number of sites visited exactly by $k$ walkers (out of $N$) up to time $t$ using our result in Eq. (2). Here we have restricted our attention only to the $k = N$ case for simplicity. Furthermore, it follows by putting $k = 0$ in Eq. (2) that the probability a site $\vec{x}$ is not visited by any of the $N$ walkers is simply $P_{0,N}(\vec{x},t) = [1 - p(\vec{x},t)]^N$. Hence, the probability that a site $\vec{x}$ is visited by at least one of the walkers is $1 - P_{0,N}(\vec{x},t) = 1 - [1 - p(\vec{x},t)]^N$. Summing over $\vec{x}$, one then gets the mean number of distinct sites visited by the walkers, $S_N(t) = \sum_{\vec{x}} [1 - [1 - p(\vec{x},t)]^N]$. Thus, knowing the behavior of $p(\vec{x},t)$, one can fully analyze $S_N(t)$ and recover rather simply the results of Ref. [14].

There are several directions in which our work can be generalized. It would be interesting to consider cases where the walkers have different step lengths or when they start at different positions [32]. Also, computing the full distribution of the number of common sites visited by all walkers remains a challenging open problem.

This work was partly done during M.V.T.’s several visits at LPTMS, Orsay and he is very grateful for the warm hospitality he received there. S.N.M. acknowledges support by ANR Grant No. 2011-BS04-013-01 WALKMAT and by the Indo-French Centre for the Promotion of Advanced Research under Project No. 4604-3. M.V.T. acknowledges support by PALM LABEX ProNet and FP7-PEOPLE-2010-IRSES 269139 DCP-PhysBio grants.
Applications of the Random Walk (North-Holland, Amsterdam, 1994).