Journal of Statistical Physics, Vol. 89, Nos. 1/2, 1997

Universal Fluctuations in the Support of the Random Walk

F. van Wijland[1] and H. J. Hilhorst[1]

Received November 26, 1996

A random walk starts from the origin of a d-dimensional lattice. The occupation number $n(\mathbf{x}, t)$ equals unity if after t steps site $\mathbf{x}$ has been visited by the walk, and zero otherwise. We study translationally invariant sums $M(t)$ of observables defined locally on the field of occupation numbers. Examples are the number $S(t)$ of visited sites, the area $E(t)$ of the (appropriately defined) surface of the set of visited sites, and, in dimension $d=3$, the Euler index of this surface. In $d \geqslant 3$, the *averages* $\bar{M}(t)$ all increase linearly with t as $t \to \infty$. We show that in $d=3$, to leading order in an asymptotic expansion in t, the *deviations* from average $\Delta M(t) \equiv M(t) - \bar{M}(t)$ are, up to a normalization, all *identical to a single "universal" random variable*. This result resembles an earlier one in dimension $d=2$; we show that this universality breaks down for $d>3$.

KEY WORDS: Simple random walk; asymptotic behavior; topology of the support.

1. INTRODUCTION

We consider the simple random walk on a d-dimensional hypercubic lattice: The walk starts at time $t=0$ at the origin $\mathbf{x}=\mathbf{0}$ and steps at $t=1, 2, 3,...$ with equal probability to one of the $2d$ neighboring lattice sites. The set of the sites that have been visited after t steps is called the support at time t. It is actually the support of the field of occupation numbers $n(\mathbf{x}, t)$ defined by

$$n(\mathbf{x}, t) = \begin{cases} 0 & \text{if } \mathbf{x} \text{ has not yet been visited at time } t \\ 1 & \text{otherwise} \end{cases} \tag{1.1}$$

This paper is dedicated to Bernard Jancovici on the occasion of his 65th birthday.

[1] Laboratoire de Physique Théorique et Hautes Energies, (Laboratoire associé au Centre National de la Recherche Scientifique, URA D0063), Université de Paris-Sud, 91405 Orsay cedex, France.

0022-4715/97/1000-0119$12.50/0

The main variable that characterizes the support is its total number of sites $S(t)$. The probability distribution of this random variable is known explicitly only in dimension $d=1$, although it is well characterized[1] also in higher dimensions. Our interest here will be in the connection between the statistical properties of $S(t)$ and those of other variables characterizing the support. An example of the variables of interest is the $(d-1)$-dimensional surface area $E(t)$ of the support, defined as the total number of nearest-neighbor bonds connecting a site of the support with a site not in it. We may imagine to obtain $E(t)$ by inspecting all pairs of neighboring sites on the lattice and retaining a contribution unity whenever a pattern is found that consists of a visited and an unvisited site. Expressions for $S(t)$ and $E(t)$ in terms of the occupation numbers are easily constructed. For later convenience we shall express them instead in terms of the complementary occupation numbers

$$m(\mathbf{x}, t) \equiv 1 - n(\mathbf{x}, t) \tag{1.2}$$

This leads to

$$\begin{aligned} S(t) &= \sum_{\mathbf{x}} [1 - m(\mathbf{x}, t)] \\ E(t) &= \sum_{\mathbf{x}} \sum_{\mathbf{e}} [m(\mathbf{x}, t) + m(\mathbf{x}+\mathbf{e}, t) - 2m(\mathbf{x}, t)\, m(\mathbf{x}+\mathbf{e}, t)] \end{aligned} \tag{1.3}$$

where $\mathbf{e}$ runs through the d basis vectors of the lattice.

Both $S(t)$ and $E(t)$ count the total number of occurrences in space of specific local patterns of visited and unvisited sites. The quantities inside the summation signs in Eq. (1.3) are the indicator functions of these patterns. Clearly observables can be constructed that count the total number of local patterns of a great many other types. A pattern may consist, for example, of a "dimer" of two visited nearest-neighbor sites oriented along a specified axis; or of an unvisited site having $2d$ visited neighbors; etc. etc. We shall generically denote such observables and their linear combinations by the symbol $M(t)$. The class of all $M(t)$ is defined precisely in Sec. 2.

There certainly exist many nontrivial correlations among the random variables $M(t)$. Nevertheless, it should be possible, in principle, to express them in terms of independently fluctuating degrees of freedom of the support. Examination[2] of dimension $d=2$ has yielded a most surprising result. It appears that to leading order as $t \to \infty$ the deviations from average $\Delta M(t) \equiv M(t) - \overline{M(t)}$ satisfy

$$\Delta M(t) \simeq b_M t (\log t)^{-k_M - 1}\, \eta(t) \qquad (d=2) \tag{1.4}$$

where $\eta(t)$ is a random variable of zero average and unit variance, b_M a proportionality constant, and k_M a nonnegative integer called the order of M (one has for example $k_S=0$ and $k_E=1$). Equation (1.4) says that in the large t limit the whole class of random variables $M(t)$ reduces to just a single fluctuating degree of freedom $\eta(t)$.

In the present work we ask how the above conclusion is affected by the spatial dimensionality d. We are not aware of any simple heuristic argument that leads to (1.4) and that would provide an indication whether or not a similar reduction of degrees of freedom still holds in higher dimension. It should be expected that as d goes up, independently fluctuating degrees of freedom will appear.

We show by explicit calculation that in dimension $d=3$, to leading order as $t \to \infty$, the fluctuating degrees of freedom still collapse to a single one in a manner very similar to $d=2$. Explicitly,

$$\Delta M(t) \simeq a_M (t \log t)^{1/2} \xi(t) \qquad (d=3) \tag{1.5}$$

where $\xi(t)$ is a random variable of zero average and unit variance, and a_M a proportionality constant. Equation (1.5) holds in particular for the choice $M=S$. The fluctuations $\Delta S(t)$ have been thoroughly studied and we know[1] therefore that the probability distribution of $\xi(t)$ is Gaussian. The variables $\xi(t)$ in $d=3$ and $\eta(t)$ in $d=2$ are universal in the sense that they do not depend upon M. Differences between $d=3$ and $d=2$ have to do, among other things, with the appearance of an "order" k_M in $d=2$ but not in $d=3$. Also, $\eta(t)$ is known to be non-Gaussian.

The way to derive the three-dimensional result is closely analogous to what was done in two dimensions. We calculate the correlation of two observables $M(t)$ and $M'(t)$ and find

$$\overline{\Delta M(t)\, \Delta M'(t)} \simeq a_M a_{M'} t \log t + \mathcal{O}(t) \qquad (d=3) \tag{1.6}$$

where a_M and $a_{M'}$ are known proportionality constants. We then exploit the fact that the prefactor in Eq. (1.6) factorizes into an M and an M' dependent part. It follows that the normalized random variables

$$\xi_M(t) \equiv a_M^{-1} (t \log t)^{-1/2} \Delta M(t) \tag{1.7}$$

have $\overline{\xi_M(t)}=0$ and $\overline{\xi_M(t)\, \xi_{M'}(t)}=1$ for all M and M'. Therefore the averages and variances of all differences $\xi_M(t)-\xi_{M'}(t)$ vanish and the $\xi_M(t)$ must all be identical to one and the same random variable that we shall call $\xi(t)$; whence (1.5).

We show, again by explicit calculation, that in all dimensions $d > 3$ the correlation between two arbitrary observables M and M' has the form

$$\overline{\Delta M(t)\, \Delta M'(t)} \simeq c_{MM'} t \qquad (d > 3) \tag{1.8}$$

with a prefactor $c_{MM'}$ that no longer factorizes. Hence for $d > 3$ the fluctuations of the $M(t)$ cannot be described any more by a single random variable.

Our principal result, described above, is of a general nature. In Sec. 5 we use our calculational framework to study three examples of fluctuating observables $M(t)$. One of these is the Euler index $\chi(t)$ of the surface of the support of the three-dimensional random walk. We recall that the Euler index of a closed two-dimensional surface embedded in three-dimensional space (the surface of a "Swiss cheese") is

$$\chi = 2(1 - H + C) \tag{1.9}$$

where H is the number of handles of the surface and C the number of cavities enclosed by it. In Section 5, for appropriately defined "handles" and "cavities," we construct an explicit expression for $\chi(t)$ and show that it is a member of the class of $M(t)$. We then determine its asymptotic time dependence and its correlation with other observables.

2. GENERAL OBSERVABLES $M(t)$

The general property or "observable" $M(t)$ that we shall deal with can now be defined as follows. Let A be a finite (typically small) subset of lattice vectors. Then we write

$$M(t) = \sum_A \mu_A M_A(t) \tag{2.1}$$

where the μ_A are numerical coefficients, where the sum is through all subsets A not equivalent by a translation, and where

$$M_A(t) = \sum_{\mathbf{x}} m_{\mathbf{x}+A}(t) \tag{2.2}$$

with

$$m_{\mathbf{x}+A}(t) = \prod_{\mathbf{a} \in A} m(\mathbf{x} + \mathbf{a}, t) \tag{2.3}$$

Here $\mathbf{x} + A$ is a convenient notation for the set of lattice sites $\{\mathbf{x} + \mathbf{a} \mid \mathbf{a} \in A\}$. When $A = \varnothing$ the RHS of Eq. (2.3) should be assigned the value unity and Eq. (4.11) shows that $M_{\varnothing}(t)$ is the total number of lattice sites. As an example one may check that $S(t)$ has $\mu_A = \pm 1$ for $A = \varnothing$ and $A = \{\mathbf{0}\}$, respectively, and $\mu_A = 0$ otherwise; and that $E(t)$ has $\mu_A = 2d, -2$ for $A = \{\mathbf{0}\}, \{\mathbf{0}, \mathbf{e}\}$, respectively, and $\mu_A = 0$ otherwise. We note that the expansion coefficients μ_A should satisfy the relation $\sum_A \mu_A = 0$ if the sum on $\mathbf{x}$ in Eq. (2.2) is to have a finite value in the infinite lattice limit.

3. AVERAGES OF OBSERVABLES

3.1. Averages $\bar{M}(t)$

The expression (2.3) averaged on all random walk trajectories represents the probability that the set $\mathbf{x} + A$ has not yet been visited by the walk. It is not surprising, therefore, that the average $\bar{M}(t)$ of interest to us is connected to the distribution of first passage times for that set. More precisely, we can express it in terms of the probability $f_{\mathbf{x}+A}(\tau)$ that the walker's first visit to the set $\mathbf{x} + A$ takes place at time τ. Elementary manipulations lead to the explicit relation

$$\hat{\bar{M}}(z) = -\frac{1}{1-z} \sum_{A \neq \varnothing} \mu_A \hat{F}_A(z) \tag{3.1}$$

in which

$$\hat{F}_A(z) = \sum_{\mathbf{x}} \hat{f}_{\mathbf{x}+A}(z) \tag{3.2}$$

and where, here and below, any time dependent quantity $X(t)$ that occurs will be associated with a generating function $\hat{X}(z) = \sum_{t=0}^{\infty} z^t X(t)$. The first passage time probabilities can be expressed in the random walk Green function by standard methods. After some algebra[2] one finds

$$\hat{F}_A(z) = \frac{1}{(1-z)} G_A^{-1}(z) \tag{3.3}$$

in which

$$G_A^{-1}(z) = \sum_{\mathbf{a}, \mathbf{a}' \in A} (\mathbf{G}_A^{-1})_{\mathbf{a}\mathbf{a}'} \tag{3.4}$$

where $\mathbf{G}_A$ is the matrix of elements $\hat{G}(\mathbf{a} - \mathbf{a}', z)$ with $\mathbf{a}$ and $\mathbf{a}'$ restricted to the set A, and $G(\mathbf{x}, t)$ is the probability that a walker starting at the origin occupies site $\mathbf{x}$ after t steps. For the later purpose of discussing the small-$(1-z)$ expansion it will be convenient to split $\hat{G}(\mathbf{x}, z)$ up according to

$$\hat{G}(\mathbf{x}, z) = \hat{G}(\mathbf{0}, z) - g(\mathbf{x}, z) \tag{3.5}$$

With the aid of a little algebra we find the analogous splitting

$$G_A(z) = \hat{G}(\mathbf{0}, z) - g_A(z) \tag{3.6}$$

in which

$$g_A^{-1}(z) = \sum_{\mathbf{a}, \mathbf{a}' \in A} (\mathbf{g}_A^{-1})_{\mathbf{a}\mathbf{a}'} \tag{3.7}$$

where $\mathbf{g}_A$ is the matrix of elements $g(\mathbf{a}-\mathbf{a}', z)$ with $\mathbf{a}$ and $\mathbf{a}'$ restricted to the set A. Using Eqs. (3.6) and (3.3) in Eq. (3.1) and inverting that equation we obtain the final result for the average of a general observable $M(t)$,

$$\bar{M}(t) = -\frac{1}{2\pi \mathbf{i}} \oint \frac{dz}{z^{t+1}} \frac{1}{(1-z)^2} \sum_{A \neq \varnothing} \mu_A \frac{1}{\hat{G}(\mathbf{0}, z) - g_A(z)} \tag{3.8}$$

where the integral runs counterclockwise around the origin. The coefficient $G_A(1)^{-1}$ appears in the literature[3] as the electrostatic capacity of the set A.

3.2. Averages $\overline{M(t)\, M'(t)}$

The calculation of the correlation $\overline{\Delta M(t)\, \Delta M'(t)}$ requires that we now consider the average of the product $M(t)\, M'(t)$. The latter can be rewritten as

$$M(t)\, M'(t) = \sum_{\mathbf{r}} \sum_{A, B} \mu_A \mu'_B M_{A \cup (\mathbf{r}+B)}(t) \tag{3.9}$$

The observable inside the sum on $\mathbf{r}$ is of the same type as $M(t)$ and $M'(t)$, even if it is parametrically dependent on $\mathbf{r}$. The average of (3.9) can therefore be evaluated with the aid of Eq. (3.8). After separating the terms with $A = \varnothing$ or $B = \varnothing$ from the others we find

$$\overline{M(t)\, M'(t)} = -\frac{1}{2\pi \mathbf{i}} \oint \frac{dz}{z^{t+1}} \frac{1}{(1-z)^2} \sum_{A, B \neq \varnothing} \mu_A \mu'_B \sum_{\mathbf{r}} I_{AB}(\mathbf{r}, z) \tag{3.10}$$

in which

$$I_{AB}(\mathbf{r}, z) = \frac{1}{G_{A\cup(\mathbf{r}+B)}(z)} - \frac{1}{G_A(z)} - \frac{1}{G_B(z)} \tag{3.11}$$

This expression involves $G_{A\cup(\mathbf{r}+B)}(z)$, a quantity whose $\mathbf{r}$ dependence needs to be rendered more explicit if the sum on $\mathbf{r}$ in Eq. (3.10) is to be performed. This we are able to achieve only in the scaling limit $r \to \infty$ and $z \to 1$ with $\xi^2 \equiv 2\,dr^2(1-z)$ fixed. In that limit a calculation analogous to that of ref. 2 for the two-dimensional case leads to

$$\frac{1}{G_{A\cup(\mathbf{r}+B)}(z)} \simeq \frac{G_A(z) + G_B(z) - 2\hat{G}(\mathbf{r}, z)}{G_A(z)\,G_B(z) - \hat{G}^2(\mathbf{r}, z)} \tag{3.12}$$

where

$$\hat{G}(\mathbf{r}, z) \simeq d^{d/2}\pi^{-d/2}(1-z)^{d/2-1}\,\xi^{1-d/2}K_{d/2-1}(\xi) \tag{3.13}$$

The expressions of this section up through Eq. (3.11) are exact for all finite times t. Only in Eqs. (3.12) and (3.13) has the scaling limit been taken, is an infinite lattice implied, and does the dimensionality of space appear explicitly.

4. LONG TIME BEHAVIOR

4.1. Averages $\bar{M}(t)$

The long time behavior of $\bar{M}(t)$ is determined by the behavior of $\hat{\bar{M}}(z)$ in the vicinity of $z=1$. We need the small-$(1-z)$ expansion of $G_A(z)$, hence of $\hat{G}(\mathbf{0}, z)$ and $g_A(z)$. The expansion of $\hat{G}(\mathbf{0}, z)$ is well-known[4,5] and reads

$$\hat{G}(\mathbf{0}, z) = \begin{cases} G_0 - \dfrac{3^{3/2}}{2^{1/2}\pi}(1-z)^{1/2} + \mathcal{O}(1-z) & (d=3) \\ G_0 - \dfrac{4}{\pi^2}(1-z)\ln\dfrac{1}{1-z} + c_4(1-z) + o((1-z)^{3/2}) & (d=4) \\ G_0 - c_d(1-z) + \mathcal{O}((1-z)^{3/2}) & (d\geqslant 5) \end{cases} \tag{4.1}$$

where we have abbreviated $G_0 \equiv \hat{G}(\mathbf{0}, 1)$ and where the c_d with $d = 4, 5,...$ are positive constants. Since $g(\mathbf{x}, z)$ is a solution of

$$\frac{1}{2d} \sum_{\mu=1}^{d} [g(\mathbf{x}+\mathbf{e}_\mu, z) - 2g(\mathbf{x}, z) + g(\mathbf{x}-\mathbf{e}_\mu, z)] = \frac{1}{z}\delta_{\mathbf{x},\mathbf{0}} - \frac{1}{z}(1-z)\,\hat{G}(\mathbf{0}, z) \tag{4.2}$$

one may conclude that $g_A(z)$ defined in Eq. (3.7) has an expansion of the form

$$g_A(z) = g_A(1) + (1-z)\,g_A^{(1)} + o(1-z) \tag{4.3}$$

as $z \to 1$, where $g_A^{(1)}$ is an A dependent numerical constant. The $(1-z)$-expansion of $\hat{M}(z)$ therefore reads

$$\hat{M}(z) = \begin{cases} m_1(1-z)^{-2} + \dfrac{3^{3/2}}{2^{1/2}\pi} m_2(1-z)^{-3/2} + \mathcal{O}((1-z)^{-1}) \\ m_1(1-z)^{-2} + \dfrac{4}{\pi^2} m_2(1-z)^{-1} \log\dfrac{1}{1-z} \\ \qquad + (m_2 - c_4 m_2^{(1)})(1-z)^{-1} + o((1-z)^{-1/2}) \\ m_1(1-z)^{-2} + (c_d m_2 + m_2^{(1)})(1-z)^{-1} + \mathcal{O}((1-z)^{-1/2}) \end{cases} \tag{4.4}$$

in dimensions $d = 3$, $d = 4$, and $d \geqslant 5$, respectively; here the coefficients $m_n, m_2^{(1)}$ are determined by the observable M through the relations

$$m_n[M] = -\sum_{A \neq \varnothing} \frac{\mu_A}{G_A^n(1)}, \qquad m_2^{(1)}[M] = -\sum_{A \neq \varnothing} \frac{\mu_A g_A^{(1)}}{G_A^2(1)} \tag{4.5}$$

When no confusion can arise we shall not indicate their M dependence explicitly. For transforming back to the time domain we need the basic expansions, for $s > 0$,

$$\frac{1}{2\pi\mathbf{i}} \oint \frac{dz}{z^{t+1}} \frac{1}{(1-z)^s} = \frac{t^{s-1}}{\Gamma(s)} \left[1 + \frac{s^2 - s}{2} \frac{1}{t} + \mathcal{O}\left(\frac{1}{t^2}\right) \right] \tag{4.6}$$

$$\frac{1}{2\pi\mathbf{i}} \oint \frac{dz}{z^{t+1}} \frac{1}{(1-z)^s} \ln\frac{1}{1-z} = \frac{t^{s-1}}{\Gamma(s)} \left[\log t - \psi(s) + \mathcal{O}\left(\frac{\log t}{t}\right) \right] \tag{4.7}$$

where ψ is the psi function. The first integral is easily found by writing the integrand as a power series in z and the second one follows from it by differentiation with respect to s. The result is

$$\bar{M}(t) = \begin{cases} m_1 t + \dfrac{2^{1/2}3^{3/2}}{\pi^{3/2}} m_2 t^{1/2} + \mathcal{O}(1) & (d=3) \\ m_1 t + \dfrac{4}{\pi^2} m_2 \ln t + m_1 + \dfrac{4\gamma}{\pi^2} m_2 - c_4 m_2 + m_2^{(1)} + o(t^{-1/2}) & (d=4) \\ m_1 t + m_1 + c_d m_2 + m_2^{(1)} + \mathcal{O}(t^{-1/2}) & (d \geq 5) \end{cases} \tag{4.8}$$

where γ denotes Euler's constant. In Sec. 5 some of the coefficients appearing in Eq. (4.8) will be calculated for specific examples of observables $M(t)$.

4.2. Fluctuations and Correlations

In order to study the fluctuations and correlations of two observables $M(t)$ and $M'(t)$ in the large time limit, we now have to evaluate the average (3.10) in the limit $t \to \infty$. The principal problem is to find the expansion of $\sum_{\mathbf{r}} I_{AB}(\mathbf{r}, z)$ as $z \to 1$. Whereas $G_A(z)$ and $G_B(z)$ remain finite in that limit, Eq. (3.13) shows that in all $d > 2$ the quantity $\hat{G}(\mathbf{r}, z)$ vanishes when $z \to 1$ (with ξ fixed). This suggests that different leading powers of $1 - z$ are associated with different powers of $\hat{G}(\mathbf{r}, z)$. We therefore express I_{AB} in the form

$$I_{AB}(\mathbf{r}, z) = I_1(z)\, \hat{G}(\mathbf{r}, z) + \cdots + I_{n-1}(z)\, \hat{G}^{n-1}(\mathbf{r}, z) + f_{AB}^{(n)}(\mathbf{r}, z) \tag{4.9}$$

in which we take for $I_k(z)$ the coefficient of $\hat{G}^k(\mathbf{r}, z)$ in a power series expansion of the RHS of Eq. (3.11) in which for $1/G_{A \cup (\mathbf{r}+B)}(z)$ the RHS of Eq. (3.12) has been substituted. Explicitly,

$$\begin{aligned} I_1(z) &= -2/G_A(z)\, G_B(z) \\ I_2(z) &= (G_A(z) + G_B(z))/G_A^2(z)\, G_B^2(z) \\ I_3(z) &= -2/G_A^2(z)\, G_B^2(z) \end{aligned} \tag{4.10}$$

Even though Eq. (3.12) holds only in the scaling limit, Eq. (4.9) is an exact identity if it is taken as the definition of the remainder $f_{AB}^{(n)}$. It is now possible to find the small $1 - z$ expansion of $\sum_{\mathbf{r}} I_{AB}(\mathbf{r}, z)$ in the following way. After substituting Eq. (4.9) in Eq. (3.9) we evaluate the sum on all space of the powers $\hat{G}^k(\mathbf{r}, z)$ for $k=1, 2, \ldots$. For $k=1$ conservation of probability trivially leads to $\sum_{\mathbf{r}} \hat{G}(\mathbf{r}, z) = (1-z)^{-1}$. For $k=2$ one easily derives the general relation

$$\sum_{\mathbf{r}} \hat{G}^2(\mathbf{r}, z) = \left(1 + z\frac{d}{dz}\right) \hat{G}(\mathbf{0}, z) \tag{4.11}$$

after which use can be made of Eq. (4.1). It follows that

$$\sum_{\mathbf{r}} \hat{G}^2(\mathbf{r}, z) = \begin{cases} \dfrac{3^{3/2}}{2^{3/2}\pi}(1-z)^{-1/2} + \mathcal{O}(1) \\ \dfrac{4}{\pi^2}\ln\dfrac{1}{1-z} + G_0 - c_4 - \dfrac{4}{\pi^2} + \mathcal{O}((1-z)^{1/2}) \end{cases} \tag{4.12}$$

in dimensions $d=3$ and $d=4$, respectively. In dimensions $d>4$ this quantity does not diverge as $z \to 1$. Finally, for $k=3$, we employ the scaling expression (3.13) for $\hat{G}(\mathbf{r}, z)$. In dimension $d=3$ the resulting integral on ξ diverges logarithmically at the origin, and using the lower end cutoff $\xi \sim r\sqrt{1-z}$ we find

$$\sum_{\mathbf{r}} \hat{G}^3(\mathbf{r}, z) = \frac{3^3}{2\pi^2}\ln\frac{1}{1-z} + \mathcal{O}(1) \qquad (d=3) \tag{4.13}$$

In dimensions $d>3$ the sum with $k=3$ remains finite as $z \to 1$. The remainder of the work will be carried out separately for the three qualitatively different cases $d=3$, $d=4$, and $d \geqslant 5$.

4.2.1. Dimension $d=3$. We employ Eq. (4.9) with $n=4$. In the limit $z \to 1$ the sum $\sum_{\mathbf{r}} f_{AB}^{(4)}(\mathbf{r}, 1)$ remains finite. Using the above expansions and also expanding $G_A(z)$ and $G_B(z)$ for small $1-z$ we find for the average of Eq. (3.10) the expression

$$\overline{M(t)\,M'(t)} = \frac{1}{2\pi \mathrm{i}}\oint \frac{dz}{z^{t+1}}\left[2m_1 m_1' \frac{1}{(1-z)^3} + \frac{3^{5/2}}{2^{3/2}\pi}(m_1 m_2' + m_1' m_2)\frac{1}{(1-z)^{5/2}} + \frac{3^3}{2\pi^2} m_2 m_2' \frac{1}{(1-z)^2}\log\frac{1}{1-z} + \cdots\right] \tag{4.14}$$

in which the m_n' are determined by M' via Eq. (4.5). After carrying out the z integral with the aid of the basic relations of Eqs. (4.6) and (4.7) and subtracting the product $\bar{M}(t)\,\bar{M}'(t)$ obtainable from Eq. (4.8) we are led to the final result

$$\overline{\Delta M(t)\,\Delta M'(t)} \simeq \frac{3^3}{2\pi^2} m_2 m_2' t \ln t + c_{MM'} t \tag{4.15}$$

where the constant $c_{MM'}$ is due to contributions from various subleading terms that have not been displayed explicitly above. Equation (4.15) leads to the result announced in Eq. (1.6) of the Introduction, whose consequence is the universality embodied by Eq. (1.5).

4.2.2. Dimensions $d=4$ and $d \geqslant 5$. We employ Eq. (4.9) with $n=3$. In the limit $z \to 1$ the sum $\sum_{\mathbf{r}} f_{AB}^{(3)}(\mathbf{r}, 1)$ remains finite. Proceeding as in the case of $d=3$ we obtain in dimension $d=4$ the final result

$$\begin{aligned}\overline{\Delta M(t)\,\Delta M'(t)} = [&m_1 m_1' - G_0(m_1 m_2' + m_2 m_1') \\ &+ m_1 m_2^{(1)'} + m_2^{(1)} m_1' - \mathscr{F}_{MM'}]\, t + \cdots\end{aligned} \tag{4.16}$$

where

$$\mathscr{F}_{MM'} = \sum_{A,\,B \neq \varnothing} \mu_A \mu_B' \sum_{\mathbf{r}} f_{AB}^{(3)}(\mathbf{r}, 1) \tag{4.17}$$

The only difference that arises when $d \geqslant 5$ is that not only $\sum_{\mathbf{r}} f_{AB}^{(3)}(\mathbf{r}, 1)$ but also $\sum_{\mathbf{r}} \hat{G}^2(\mathbf{r}, 1)$ remains finite as $z \to 1$. Straightforward calculation leads in a slightly different way to exactly the same result (4.16). If we denote the coefficient of the term linear in t again by $c_{MM'}$, then Eq. (4.16) constitutes the result announced in Eq. (1.8).

5. EXAMPLES

We apply the preceding theory to three examples. As a preliminary we render more precise the concept of "surface of the support." We imagine $\mathbb{R}^d$ partitioned into unit cubes that surround the lattice sites. Visited (unvisited) unit cubes correspond to visited (unvisited) lattice sites. The support of the walk then defines a visited volume whose $(d-1)$-dimensional surface will be called the surface of the support. In the second example we shall study the Euler index of the surface of the support of the three-dimensional walk. In order for this topological invariant—which involves the handles and the cavities of the support—to be well defined, we adopt the convention that two visited unit cubes are connected as soon as they share a vertex (and hence a fortiori when they share an edge or a face). Figure 1 shows an example of two connected unit cubes; their connectedness may

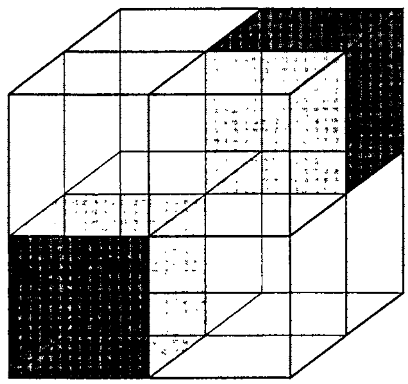


Fig. 1. Example of a local configuration of $2\times2\times2$ sites that may occur in the support of the three-dimensional walk. Shaded (unshaded) unit cubes are centered on visited (unvisited) lattice sites. The volumes of shaded cubes are considered to be connected via the common vertex.

be made visible by an infinitesimal deformation of the surface near the common vertex.

5.1. Surface Area $E(t)$ of the Support

From Eq. (4.8) we deduce that the average surface area of the support in dimension $d\geqslant 3$ behaves as

$$\bar{E}(t)=m_1[E]\,t+o(t) \tag{5.1}$$

In the particular case $d=3$ we have, more precisely,

$$\bar{E}(t)=m_1[E]\,t+\frac{2^{1/2}3^{3/2}}{\pi^{3/2}}m_2[E]\sqrt{t}+\mathcal{O}(1) \qquad (d=3) \tag{5.2}$$

From Eq. (4.5) and the remarks below Eq. (2.3) it follows that for all d

$$m_n[E]=2d[(G_0-\tfrac{1}{2})^{-n}-G_0^{-n}] \tag{5.3}$$

where we used that for $A=\{\mathbf{0}\}, \{\mathbf{0},\mathbf{e}\}$ one has $G_A=G_0, G_0-\frac{1}{2}$, respectively. In $d=3$ we have $G_0=1.5164...$ ([4]) so that

$$m_1[E]=1.95, \quad m_2[E]=3.20 \qquad (d=3) \tag{5.4}$$

In dimension $d\gg 1$, using the expansion $G_0=1+(2d)^{-1}+\mathcal{O}(d^{-2})$ one sees that $\bar{E}(t)\simeq[2d-3+\mathcal{O}(d^{-1})]\,t$, which may also be found by more heuristic arguments.

5.2. Euler Index $\chi(t)$

The Euler index of an arbitrary, not necessarily closed, two-dimensional surface embedded in $\mathbb{R}^d$ (with $d \geqslant 2$) is

$$\chi = 2(1 - H + C) - B \tag{5.5}$$

where H, C, and B are the number of handles, cavities (connected components of the complement of the surface), and boundaries, respectively.

The support of the two-dimensional walk is a flat two-dimensional surface with $H = C = 0$ and therefore has Euler index $\chi = 2 - B$; the number of boundaries B is equal to one plus the number of unvisited islands enclosed by the support. This quantity was studied in ref. 2.

Here we wish to study the Euler index of the surface of the support of the three-dimensional walk. For this closed ($B = 0$) two-dimensional surface the Euler index reduces to $\chi = 2(1 - H - C)$. The handles and cavities are nonlocal objects of varying sizes and shapes, and it is interesting that one should be able to calculate the Euler index. The reason that this is possible is that the Euler index can be expressed, by means of the Gauss–Bonnet theorem, as an integral over the surface of a function of the local curvature.

We first need a discrete analog of the Gauss–Bonnet theorem, that is, an expression for χ as the sum over space of a function defined locally on a two-dimensional surface built out of plaquettes. In order to find the contribution from the surface around a specific vertex, we examine all 2^8 patterns of occupation numbers in the surrounding cubic volume of $2 \times 2 \times 2$ unit cubes. The local contribution to the Euler index will be a weighted sum of the indicator functions of these patterns. There are only 22 patterns not equivalent under symmetry operations, and therefore 22 weights to be determined. One can set up equations for these weights by requiring that the final expression correctly yield the known result for various surfaces having the topology of a sphere ($\chi = 2$) or of a torus ($\chi = 0$). One finds, for example, that the pattern shown in Fig. 1 has weight -3.

It is then a tedious but straightforward matter to cast the expression for χ in the form of Eq. (2.1) by working out the products with factors $1 - m$, summing on $\mathbf{x}$, and rearranging terms. Since the resulting expression for χ may also be of use in other contexts, we have listed the coefficients in Fig. 2. The nonzero coefficients μ_A are the 18 entries of that table, and those obtainable from them by rotations and reflections. Also listed are the values of the inverse capacities $G_A(1)$, found numerically by determining the required values of $\hat{G}(\mathbf{x}, 1)$ and inverting the 18 matrices $\mathbf{G}_A$. These data

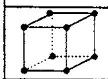
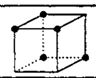
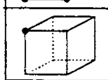
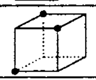
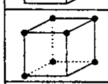
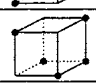
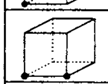
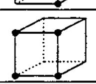
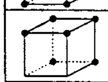
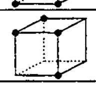
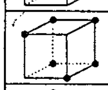
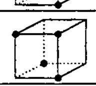
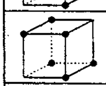
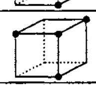
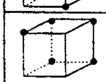
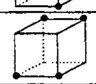
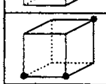
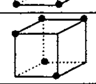
A	μ_A	σ_A	$G_A(1)$	A	μ_A	σ_A	$G_A(1)$
	-8	1	0.540		3/4	16	0.614
	2	1	1.516		3/8	16	0.726
	3	8	0.559		9/8	16	0.602
	-2	3	1.016		2	3	0.720
	-5/6	12	0.591		1/6	24	0.684
	-11/6	12	0.579		-7/12	24	0.661
	-2	4	0.579		-1/2	8	0.690
	1/4	24	0.633		1/3	6	0.656
	-1/8	16	0.749		-3	2	0.627

Fig. 2. The Euler index can be written as $\chi = \sum_A \mu_A M_A$, with the M_A defined in the text. The first and fifth column list the subsets A of lattice sites (heavy dots) that have nonzero μ_A. The other nonzero μ_A correspond to the sets A obtainable from those shown by rotations and/or reflections. The total number of distinct sets related to A by such symmetry operations is σ_A. The coefficients $G_A(1)$ are the inverse capacities needed in the calculation.

then allow the large time behavior of the average $\bar{\chi}(t)$ to be calculated. It has the general form of Eq. (4.8),

$$\bar{\chi}(t) = m_1[\chi]\, t + \frac{2^{1/2}3^{3/2}}{\pi^{3/2}} m_2[\chi]\sqrt{t} + \mathcal{O}(1) \qquad (d=3) \tag{5.6}$$

with the numerical values

$$m_1[\chi] = -3.44 \times 10^{-2}, \quad m_2[\chi] = 2.12 \times 10^{-1} \qquad (d=3) \tag{5.7}$$

The negativity of $m_1[\chi]$ implies that the Euler index is dominated by the number of handles.

The Euler index is defined also for $(d-1)$-dimensional surfaces and can always be expressed as the integral of a local function on this surface.

We have found no general way to obtain discrete analogs of these expressions when the surface dimension is $d-1$, with $d \geqslant 4$, even though for a given dimension this is in principle possible.

5.3. Cavities

A quantity of potential interest is the number of cavities of a given size, shape and orientation. The indicator function of such a cavity is a single product of m's and $(1-m)$'s and their total number is obtained by summing that product on space. This number therefore belongs to the class of obervables $M(t)$ and behaves for large times as shown in Eq. (4.8).

Let $C_1(t)$ be the number of cavities consisting of a single site. We determined the numerical values

$$m_1[C_1] = 7.92 \times 10^{-4}, \quad m_2[C_1] = -7.10 \times 10^{-3} \qquad (d=3) \tag{5.8}$$

These show the rarity of such cavities: Their number increases at the approximate rate of one every one thousand steps. Cavities larger than a single site will certainly be still rarer, so that $C_1(t) \approx C(t)$. It follows that $\chi(t)$ is approximately equal to minus twice the number $H(t)$ of handles of the surface of the support.

5.4. Correlations

We have not included in the above examples the number of sites $S(t)$ of the support, as this quantity has been thoroughly studied in the literature.[6,7] It is now useful to observe that the average $\bar{S}(t)$ behaves, in $d=3$, according to the general law (4.8) with coefficients $m_1[S]$ and $m_2[S]$ that are easily determined via our general rules. One finds

$$m_1[S] = 0.659, \quad m_2[S] = 0.434 \qquad (d=3) \tag{5.9}$$

in agreement with the literature.[7] Upon inserting the various coefficients m_2 determined above in Eq. (4.15) we directly obtain the leading order behavior of all cross-correlations between the quantities studied. It appears that the total surface area $E(t)$ of the support is positively correlated with $S(t)$, and that (since $C_1(t) \approx C(t)$) the numbers of handles and of cavities, $H(t)$ and $C(t)$, are negatively correlated with $S(t)$.

6. DISCUSSION

The asymptotic behavior as $t \to \infty$ of the average and the variance of the total number S of sites in the support is well known in all dimensions.

These known results are reproduced by the relations of this paper if one sets $M = M' = S$. In the dimensions $d = 2$ and $d = 3$ the variable S is known to have anomalous fluctuations in the sense that the rms deviation $\overline{\Delta S^2}^{1/2}$ increases faster with t than proportional to $\bar{S}^{1/2}(t)$. The fluctuations in S do not arise, therefore, as the sum of effectively independent contributions along the random walk trajectory, but must be due to a collective effect. It is easy to accept that similar anomalous fluctuations appear in an entire class of observables $M(t)$ related to the support, and in their correlations. What is surprising, however, is that the $M(t)$ all fluctuate in strict proportionality to one another. This proportionality holds both in dimension $d = 3$ (a result of this work) and in $d = 2$ (derived in ref. 2). In $d \geqslant 4$ the anomalous fluctuations and the rigorous proportionality of the fluctuating observables disappear.

The number of sites $S(t)$ visited by a d-dimensional lattice random walk has its analog, in continuum space, in the d-dimensional volume ("Wiener sausage") swept out during a time interval t by a hypersphere of radius b that performs Brownian motion. The volume of the Wiener sausage has been a recurrent object of investigation (see e.g. ref. 8), and the islands enclosed by it in $d = 2$ have been considered by mathematicians (e.g., recently by Werner[9]). However, we are not aware of any results, in dimensions $d \geqslant 3$, on the total surface area of the Wiener sausage, let alone on such properties as the number of handles or cavities. We see here problems for future research.

REFERENCES

1. J.-F. Le Gall, *Commun. Math. Phys.* **104**:471 (1986).
2. F. van Wijland, S. Caser, and H. J. Hilhorst, *J. Phys. A* **30**:507 (1997).
3. G. Lawler, *Intersections of random walks*, Probability and Its Applications (Birkhäuser, Boston 1991).
4. E. W. Montroll and G. H. Weiss, *J. Math. Phys.* **6**:167 (1965).
5. G. Zumofen and A. Blumen, *J. Chem. Phys.* **76**:3713 (1982).
6. G. H. Weiss, *Aspects and Applications of the Random Walk* (North-Holland, Amsterdam 1994).
7. B. D. Hughes, *Random Walks and Random Environments, Vol.* 1: *Random Walks* (Clarendon Press, Oxford 1996).
8. A. M. Berezhkovskii, Yu. A. Makhnovskii, and R. A. Suris, *J. Stat. Phys.* **57**:333 (1989).
9. W. Werner, *Probab. Theory Relat. Fields* **98**:307 (1994).