Journal of Statistical Physics, Vol. 44, Nos. 1/2, 1986

Variance of the Range of a Random Walk

David C. Torney[1]

Received October 8, 1985; final February 28, 1986

T_n, the expectation of the square of the number of distinct sites occupied by a random walk in steps 1 through n, is obtained from its relation to the dual first occupancy probability $F_{ij}(x, x')$, and the latter quantity is obtained from a recursion with the first occupancy probability $F_k(x'')$. The variance V_n of the number of distinct sites occupied is calculated directly from T_n; the procedure is illustrated by the calculation of V_n $(4096 \geqslant n)$ and the derivation of asymptotic expansions for V_n for a particular random walk in dimensions 1 through 3.

KEY WORDS: Number of distinct sites; asymptotic expansions; generating functions; one dimension (1-d); two dimensions (2-d); three dimensions (3-d); body-centered cubic (b.c.c) random walk.

INTRODUCTION

Consider an aggregate of M *equivalent* lattice points. Such aggregates have the property that for each point the set $\mathbf{L}$ of M displacements to all points is identical, an example being a d-dimensional Euclidean lattice torus.[1] A *step* of a random walk is a displacement by $\boldsymbol{l} \in \mathbf{L}$; $\boldsymbol{l}$ is chosen with the probability $P_1(\boldsymbol{l})$. The walk then *occupies* the lattice point it is displaced to by the vector $\boldsymbol{l}$. The convolution of j probabilities $P_1(\boldsymbol{l})$ gives the probability $P_j(\mathbf{x})$ of being displaced by $\mathbf{x}$ from the point of origin on step j and defines *random walk*. The number of distinct sites occupied in steps 1 through n is a random variable—the *range* of the random walk.

This paper is concerned with V_n, the variance of the range of the random walk. The expectation of the range E_n is well-characterized;[1]

Work completed under the auspices of the United States Department of Energy.

[1] Theoretical Biology and Biophysics, Theoretical Division, Los Alamos National Laboratory, Group T-10, Mail Stop K710, Los Alamos, New Mexico 87545 U.S.A.

0022-4715/86/0700-0049$05.00/0

822/44/1-2-4

therefore, to find V_n it is sufficient to focus on T_n the expectation of the square of the range because

$$V_n \equiv T_n - E_n^2 \tag{1}$$

T_n is also[(2)] the probability that each of two sites, randomly selected with replacement, has been occupied in steps 1 through n multiplied by the square of the number of lattice points,[2] and therefore can be found by a summation of the dual first-occupancy probability $F_{ij}(\mathbf{x}, \mathbf{x}')$, which is the probability that the site displaced by $\mathbf{x}$ from the origin is occupied for the first time on step i, and that the site displaced by $\mathbf{x}'$ from the origin is occupied for the first time on step j.

The theory of the range of a random walk was initiated primarily by Dvoretzky and Erdös[(2)] and there was subsequent significant progress in characterizing the random variable made principally by Jain and Pruitt.[(3,4,5,6)] In particular, it is directly relevant that formulas for V_n were obtained for arbitrary walks on infinite lattices in dimensions 1 through 6 in the limit that n is large. However, these were not asymptotic series and there is no indication of how large n has to be before the given formulas become valid. This paper develops methods for the calculation of V_n, useful for small n, and for the generation of asymptotic series for V_n, useful for large n. The methods are illustrated for a particular random walk in dimensions 1 through 3, and the leading terms of the asymptotic series corroborate the large n formulas discussed above.

THE VARIANCE

The fundamental probabilities for a random walk are the $P_j(\mathbf{x})$, the probabilities of being displaced from the point of origin to the site $\mathbf{x}$ on the jth step. One can obtain the $F_j(\mathbf{x})$, the probabilities of being displaced from the point of origin to the site $\mathbf{x}$ for the first time on the jth step,[(7)] recursively from the $P_j(\mathbf{x})$

$$F_j(\mathbf{x}) = P_j(\mathbf{x}) - \sum_{i=1}^{j-1} P_{j-i}(\mathbf{o})\, F_i(\mathbf{x}) \qquad j \geqslant 1 \tag{2}$$

[2] The correlation is demonstrated in the following context. Each of the W possible walks is represented by a box containing M balls, one corresponding to each lattice point. In a given box, one distinguishes a subset of balls, for example, by making them red—the number of balls in the subset divided by M being the fraction of distinct sites occupied in the walk represented by the box. The probability t of selecting two red balls, if one randomly chooses two balls from the box with replacement, is the square of this fraction. The arithmetic average $\bar{t}$ of the value t from all the boxes is the average of the square of the fraction of balls that are red.

The expectation of the range in steps one through n E_n can now be found using the theorem of Dvoretzky and Erdös[2] relating the probability of a walker reaching a new site to the probability of nonreturn to its origin[3]

$$\sum_{\mathbf{x}\in\mathbf{L}} F_j(\mathbf{x}) = 1 - \sum_{i=1}^{j-1} F_i(\mathbf{o})$$
$$E_n = \sum_{j=1}^{n} \left\{1 - \sum_{i=1}^{j-1} F_i(\mathbf{o})\right\} \tag{3}$$

One also calculates S_n, the expectation of the range in steps zero through n;[8] the simple relationship is

$$E_n = S_{n-1}$$

To obtain V_n we introduce the dual first-occupancy probability $F_{ij}(\mathbf{x}, \mathbf{x}')$ that a walker first occupy the site with a displacement $\mathbf{x}$ from its origin on the ith step and first occupy the site with the displacement $\mathbf{x}'$ from its origin on the jth step. These probabilities have the following recursion with the probabilities $F_k(\boldsymbol{l})$

$$\mathbf{x} = \mathbf{x}'$$
$$F_{ij}(\mathbf{x}, \mathbf{x}) = \delta_{ij} F_i(\mathbf{x}) \tag{4a}$$
$$\mathbf{x} \neq \mathbf{x}'$$
$$F_{ij}(\mathbf{x}, \mathbf{x}') = F_{j-i}(\mathbf{x}' - \mathbf{x}) \left\{F_i(\mathbf{x}) - \sum_{k=1}^{i-1} F_{ik}(\mathbf{x}, \mathbf{x}')\right\} \qquad i < j$$
$$F_{ij}(\mathbf{x}, \mathbf{x}') = 0 \qquad i = j$$
$$F_{ij}(\mathbf{x}, \mathbf{x}') = F_{i-j}(\mathbf{x} - \mathbf{x}') \left\{F_j(\mathbf{x}') - \sum_{k=1}^{j-1} F_{jk}(\mathbf{x}', \mathbf{x})\right\} \qquad i > j \tag{4b}$$

Some examples of elementary formulas for the probabilities occurring in (4a) and (4b) are

$$F_1(\mathbf{x}' - \mathbf{x})$$
$$= P_1(\mathbf{x}' - \mathbf{x})$$

[3] Although the theorem of Dvoretzky and Erdös was proven for an infinite "cubic" lattice, one can easily prove the theorem for any finite lattice with equivalent sites using generating functions.

$$F_j(\mathbf{x}'-\mathbf{x}) = \sum_{\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3,\ldots,\mathbf{x}_{j-1}\in\mathbf{L}_1} P_1(\mathbf{x}'-\mathbf{x}_1)\,P_1(\mathbf{x}_1-\mathbf{x}_2)\,P_1(\mathbf{x}_2-\mathbf{x}_3)\cdots P_1(\mathbf{x}_{j-1}-\mathbf{x})$$
$$j>1$$

$$F_{1j}(\mathbf{x},\mathbf{x}') = F_{j-1}(\mathbf{x}'-\mathbf{x})\,F_1(\mathbf{x})$$
$$j>1$$

$$F_{ij}(\mathbf{x},\mathbf{x}') = F_{j-i}(\mathbf{x}'-\mathbf{x}) \sum_{\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3,\ldots,\mathbf{x}_{i-1}\in\mathbf{L}_2} P_1(\mathbf{x}-\mathbf{x}_1)\,P_1(\mathbf{x}_1-\mathbf{x}_2)\cdots P_1(\mathbf{x}_{i-1})$$
$$j>i>1$$

$P_1(\boldsymbol{l})$ is the probability of being displaced by $\boldsymbol{l}$ in one step, $\mathbf{L}_1$ is all lattice points in $\mathbf{L}$ but $\mathbf{x}'$, and $\mathbf{L}_2$ is all lattice points in $\mathbf{L}$ but $\mathbf{x}$ and $\mathbf{x}'$. These elementary formulas and their analogues for the other cases can be rearranged to give (4a) and (4b), which is a proof that the quantities one obtains from the recursion are identical with those defined by the elementary formulas. This method of proof is adopted from the proof of the recursion between $F_j(\mathbf{x})$ and $P_j(\mathbf{x})$.[9]

From the definition of the dual first occupancy probability, it follows that

$$\begin{aligned}
T_n &= \sum_{i,j=1}^{n} \sum_{\mathbf{x}',\mathbf{x}''\in\mathbf{L}} F_{ij}(\mathbf{x}',\mathbf{x}'') \\
&= \sum_{i,j=1}^{n} \sum_{\mathbf{x}_0,\mathbf{x}\in\mathbf{L}} F_{ij}(\mathbf{x}_0-\mathbf{x},\mathbf{x}_0) \\
&\equiv \sum_{i,j=1}^{n} \sum_{\mathbf{x}\in\mathbf{L}} F_{ij}(\mathbf{x}) \\
&= \sum_{j=1}^{n} \sum_{\mathbf{x}\in\mathbf{L}} [\tilde{F}_j(\mathbf{x})+F_j(\mathbf{x})] \\
&\equiv \sum_{j=1}^{n} \tilde{F}_j + E_n
\end{aligned} \tag{5}$$

with $F_{ij}(\mathbf{x})$ and $\tilde{F}_j$ defined in passing and with the following additional definition

$$\tilde{F}_j(\mathbf{x}) \equiv \sum_{i=1}^{j-1} [F_{ij}(\mathbf{x})+F_{ji}(\mathbf{x})]$$

The recursion for the $\tilde{F}_j(\mathbf{x})$ follows from (4a) and (4b)

$$\tilde{F}_j(\mathbf{x}) = \sum_{i=1}^{j-1} F_{j-i}(\mathbf{x}) \left\{ \sum_{\mathbf{x}' \in \mathbf{L}} [2 \cdot F_i(\mathbf{x}')] - \tilde{F}_i(\mathbf{x}) \right\} \qquad \mathbf{x} \neq 0$$
$$= 0 \qquad \mathbf{x} = 0 \qquad (6)$$

under the assumption that the random walk has inversion symmetry, that is, that $F_j(l)$ equals $F_j(-l)$. In (5), one uses the equivalence of the lattice points to reduce a double sum over the lattice to a single sum. Nevertheless, to obtain T_n a separate function must be determined for each lattice point that can be occupied by the random walk. With T_n determined from (5) and (6), one can calculate V_n from (1), using (3) to find E_n.

A PARTICULAR RANDOM WALK

To illustrate the use of (3)–(6), we calculate V_n for analogous random walks on an infinite lattice in dimensions one through three. The walk is separable; the probability of being displaced to each of the 2^d vertexes of a d-dimensional cube is 2^{-d} and illustrated in Fig. 1.

We have numerically solved for the $\tilde{F}_j(\mathbf{x})$, performed the sum over lattice sites indicated in (5), and calculated the E_n to get the V_n for $4096 \geqslant n$. The results are shown in Fig. 2. A fast-fourier transform was used to deconvolve and thereby obtain the $\tilde{F}_j(\mathbf{x})$ directly from the $P_j(\mathbf{x})$; a computer program included at the end of the Appendix sets out the method. A complete listing of the data graphed in Fig. 2 and the other programs used herein is available.[4] Selected values of E_n, T_n and V_n are given in Table I. The accuracy of the numerical method was determined by increasing the number of lattice points and the number of transform points by a factor of 2 until no further changes in the reported values would be expected on iteration, as no rigorous error estimate is available. In three dimensions all sites within a cube of 680 sites centered on the origin were included and the number of transform points is 2^{16}.

The asymptotic behavior of V_n is found using generating functions; the methods are given in the Appendix. The results follow. One dimension

$$V_n \sim 0.22610963,..., n + O(n^{1/2}) \qquad (\text{A}.6)$$

$$E_n^2 \sim 2.5464791,..., n + O(n^{1/2})$$

[4] The probabilities of being at the origin and of first return to the origin on even step numbers and the expectation, expectation of the square, and variance of the range of the separable random walk in dimensions one through three on step numbers one through 4096 are given to seven places. See document no. 04401 of the ASIS National Auxiliary Publication Service, c/o Microfiche Publications, P. O. Box 3513, Grand Central Station, New York, N.Y., 10163.

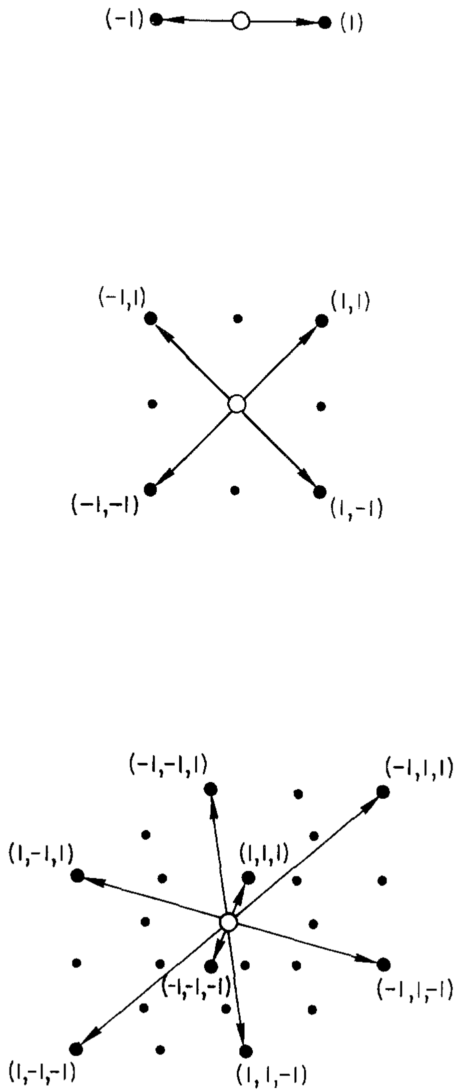

Fig. 1. The particular random walk that is used in the text, the separable walk, is illustrated. In one dimension (top) the displacements $(+1)$ and (-1) each occur with probability $\frac{1}{2}$. In two dimensions (middle) the displacements $(+1, +1)$, $(+1, -1)$, $(-1, +1)$, and $(-1, -1)$ each occur with probability $\frac{1}{4}$. In three dimensions (bottom) the displacements $(+1, +1, +1)$, $(+1, +1, -1)$, $(+1, -1, +1)$, $(-1, +1, +1)$, $(+1, -1, -1)$, $(-1, +1, -1)$, $(-1, -1, +1)$, and $(-1, -1, -1)$ each occur with probability $\frac{1}{8}$.

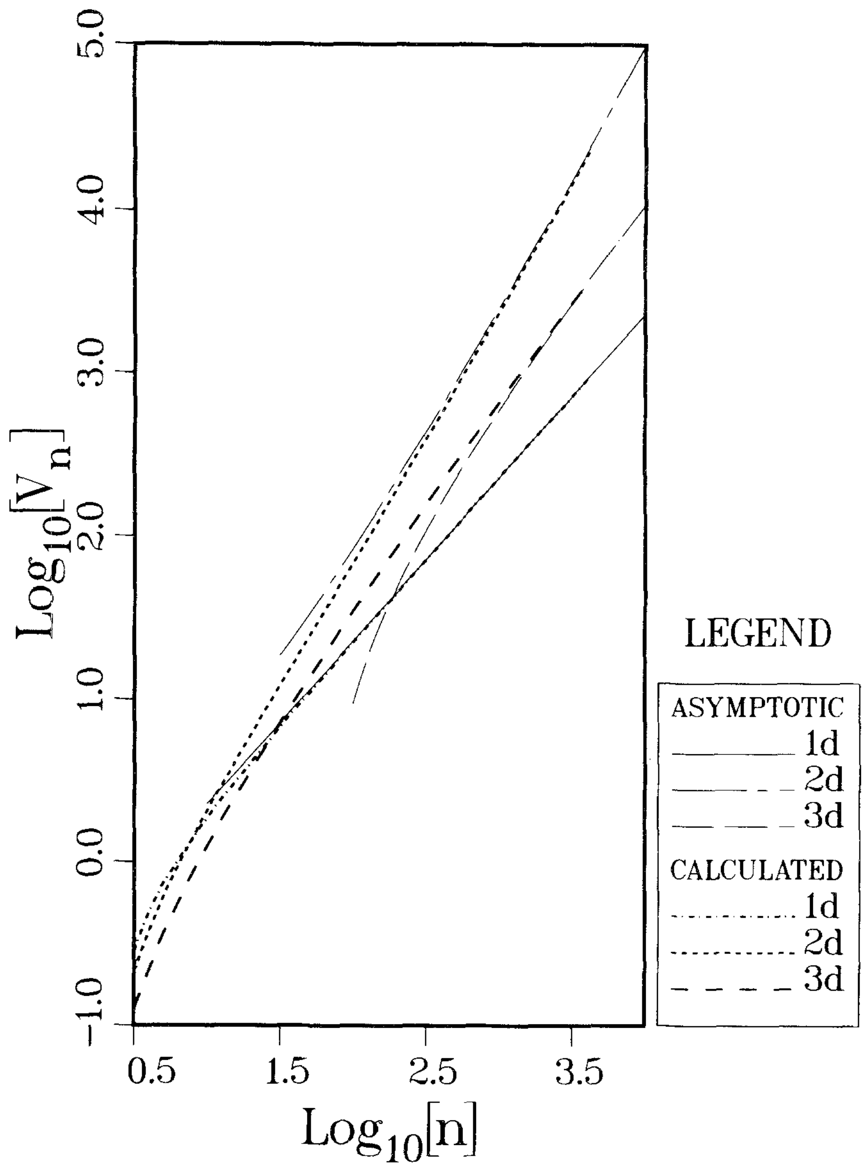

Fig. 2. Variance of range V_n for the separable random walk in dimensions 1 through 3. The axes have a logarithmic scale with the x axis being $\log_{10} n$ ($0.5 \leqslant \log_{10} n \leqslant 4.0$) and the y axis being $\log_{10} V_n$ ($-1.0 \leqslant \log_{10} V_n \leqslant 5.0$). The asymptotic formulas plotted are the sums of the terms whose coefficients have been determined in (A.6), (A.10), and (A.14) for one, two, and three dimensions. The asymptotic formula for one dimension is plotted for $1.0 \leqslant \log_{10} n \leqslant 4.0$, that for two dimensions is plotted for $1.5 \leqslant \log_{10} n \leqslant 4.0$, and that for three dimensions is plotted for $2.0 \leqslant \log_{10} n \leqslant 4.0$. The computed curves extend from $0.5 \leqslant \log_{10} n \leqslant \log_{10} 4096$ and are linear interpolations from the computed values of V_n, $3 \leqslant n \leqslant 4096$. The legend gives the pattern correspondence between the six curves and the three asymptotic formulas and three sets of calculated values; the thicker curves are computed values, and the thinner curves are asymptotic formulas.

Table I. Calculated Expectation, Expectation of Square, and Variance[a]

	$d=1$			$d=2$			$d=3$		
n	E_n	T_n	V_n	E_n	T_n	V_n	E_n	T_n	V_n
2	$2.00000e+00$	$4.00000e+00$	0.00	$2.00000e+00$	$4.00000e+00$	0.00	$2.00000e+00$	$4.00000e+00$	0.00
4	$3.00000e+00$	$9.50000e+00$	$5.00e-01$	$3.50000e+00$	$1.26250e+01$	$3.75e-01$	$3.75000e+00$	$1.42813e+01$	$2.19e-01$
8	$4.37500e+00$	$2.05313e+01$	$1.39e+00$	$6.10156e+00$	$3.86431e+01$	$1.41e+00$	$7.06299e+00$	$5.07439e+01$	$8.58e-01$
16	$6.28418e+00$	$4.26960e+01$	$3.21e+00$	$1.06852e+01$	$1.18492e+02$	$4.32e+00$	$1.34252e+01$	$1.82854e+02$	$2.62e+00$
32	$8.95680e+00$	$8.70491e+01$	$6.82e+00$	$1.88714e+01$	$3.68504e+02$	$1.24e+01$	$2.57820e+01$	$6.71939e+02$	$7.23e+00$
64	$1.27164e+01$	$1.75768e+02$	$1.41e+01$	$3.36531e+01$	$1.16740e+03$	$3.49e+01$	$4.99808e+01$	$2.51701e+03$	$1.89e+01$
128	$1.80188e+01$	$3.53212e+02$	$2.85e+01$	$6.05840e+01$	$3.76901e+03$	$9.86e+01$	$9.76552e+01$	$9.58443e+03$	$4.79e+01$
256	$2.55074e+01$	$7.08102e+02$	$5.75e+01$	$1.10018e+02$	$1.23861e+04$	$2.82e+02$	$1.91984e+02$	$3.69761e+04$	$1.18e+02$
512	$3.60905e+01$	$1.41788e+03$	$1.15e+02$	$2.01339e+02$	$4.13584e+04$	$8.21e+02$	$3.79204e+02$	$1.44081e+05$	$2.85e+02$
1024	$5.10521e+01$	$2.83745e+03$	$2.31e+02$	$3.70976e+02$	$1.40053e+05$	$2.43e+03$	$7.51608e+02$	$5.65592e+05$	$6.77e+02$
2048	$7.22075e+01$	$5.67658e+03$	$4.63e+02$	$6.87619e+02$	$4.80136e+05$	$7.32e+03$	$1.49354e+03$	$2.23225e+06$	$1.58e+03$
4096	$1.02123e+02$	$1.13548e+04$	$9.26e+02$	$1.28119e+03$	$1.66384e+06$	$2.24e+04$	$2.97335e+03$	$8.84444e+06$	$3.65e+03$

[a] The notation $e+N$ means the number is to be multiplied by 10^N

Two dimensions

$$V_n \sim 16.768193,\ldots, n^2(\log 8n)^{-4} - 16.399478,\ldots, n^2(\log 8n)^{-5} + 106.67852,\ldots, n^2(\log 8n)^{-6} + O[n^2(\log 8n)^{-7}] \qquad \text{(A.10)}$$
$$E_n^2 \sim 9.8696044,\ldots, n^2(\log 8n)^{-2} + O[n^2(\log 8n)^{-3}]$$

Three dimensions

$$V_n \sim 0.21514511,\ldots, n\log n - 0.8970,\ldots, n + O(n^{1/2})$$
$$E_n^2 \sim 0.51519379,\ldots, n^2 + O(n^{3/2}) \qquad \text{(A.14)}$$

All coefficients given to eight places are expressible in terms of constants and are known exactly (see Appendix). On the other hand, the coefficient given to four places in (A.14) was determined numerically. All coefficients have been "rounded up" if the digit one beyond the last given exceeded 4.

The asymptotic formulas for the variance, (A.6), (A.10), and (A.14), are plotted alongside the computed values in Fig. 2. In Table II, selected values of the asymptotic formulas for T_n (from the Appendix) and V_n are given. Having more than one term in the asymptotic series greatly improves the agreement of these series with the computed values in two and three dimensions for $n \simeq 4096$. The agreement is improving more rapidly in three dimensions as is expected from the orders of the first neglected terms of the series.

The leading term in one dimension was given in Ref. 3 with recourse to the relation between continuous diffusion and a discrete process. The leading term in two dimensions was given in Ref. 4 with the determinant of the covariance matrix of an equivalent walk that reaches all lattice points being one-fourth. In three dimensions the separable walk can be mapped into one that reaches all lattice points for which the determinant of the covariance matrix is one-sixteenth; the coefficient of $n \log n$ given in Ref. 5 is then in exact agreement with that in (A.14).

Table II. Asymptotic Formulas

	$d=1$		$d=2$		$d=3$	
n	T_n	V_n	T_n	V_n	T_n	V_n
512	1.41957e+03	1.16e+02	4.14954e+04	8.95e+02	1.44029e+05	2.28e+02
4096	1.13565e+04	9.26e+02	1.66495e+06	2.32e+04	8.84447e+06	3.66e+03

For all n greater than 32, one observes the asymptotic ordering: that V_n is greatest in two dimensions, intermediate in three dimensions, and smallest in one dimension. Using the Chebyshev inequality, the probability that the absolute value of the difference between the range and E_n exceeds E_n/C is less than $V_n(C/E_n)^2$. Thus, in dimensions one through three the width of the probability distribution of the range divided by its expectation decreases as d increases, being negligible asymptotically for $d \geqslant 2$, and therefore the effect of the variance on a physical process dependent on the range is expected to decrease with increasing dimensionality.

CONCLUSION

The formulas given above allow a refined calculation of the variance of the range of an arbitrary walk on a lattice with equivalent sites. Both the expectation and the variance may enter into a theory of a process which in fact depends on the distribution of the range. An example is the theory of the survival of a random walker on a lattice containing randomly placed traping sites.[10] If one includes the effect of the variance in a theory based on the expectation of the range and the effect is small within the time period of interest, one thereby substantiates and improves the theory. Thus, the methods and results of this paper have direct applications in the formulation of diffusion limited processes.

ACKNOWLEDGMENTS

An algorithm provided by Vance Faber and Milton Wing (Los Alamos National Laboratory, Computation Division) greatly facilitated the computation of V_n. George Weiss gave valuable assistance. The manuscript was prepared by Patricia Reitemeier and Sherry Richmond. The support of the United States Department of Energy is gratefully acknowledged.

APPENDIX

It is convenient to use generating functions defined as follows

$$\tilde{F}(\mathbf{x}) = \sum_{j=1}^{\infty} \tilde{F}_j(\mathbf{x})\, z^j, \qquad P(\mathbf{x}) = \sum_{j=1}^{\infty} P_j(\mathbf{x})\, z^j, \dots \tag{A.1}$$

Equations (2) and (6) give

$$\begin{aligned} \tilde{F}(\mathbf{x}) &= [2z/(1-z)]\{[1+P(\mathbf{o})]^{-1} - [1+P(\mathbf{o})+P(\mathbf{x})]^{-1}\}, && \mathbf{x} \neq 0 \\ &= 0 && \mathbf{x} = 0 \end{aligned} \tag{A.2}$$

From (2)

$$F(\mathbf{x}) = P(\mathbf{x})/[1 + P(\mathbf{o})] \leqslant 1, \qquad 0 \leqslant z \leqslant 1$$

thus with $1 > F(\mathbf{x})$, (A.2) can be written as a geometric series

$$\tilde{F}(\mathbf{x}) = \{2z/([1-z][1+P(\mathbf{o})])\} \sum_{m=0}^{\infty} (-)^m \{P(\mathbf{x})/[1+P(\mathbf{o})]\}^{m+1}$$

$$\tilde{F} \equiv \sum_{m=1}^{\infty} \tilde{F}^{(m)} \equiv [2z/(1-z)] \sum_{m=1}^{\infty} [1+P(\mathbf{o})]^{-m-1} \sum_{\mathbf{x}}{}' (-)^{m+1} [P(\mathbf{x})]^m; \tag{A.3}$$

the primed summation is over all lattice points except the origin.

One Dimension

In one dimension[1]

$$P(0) = (1-z^2)^{-1/2} - 1$$

and for $k \neq 0$

$$P(k) = x^{|k|}(1-z^2)^{-1/2}$$

with $x = [1-(1-z^2)^{1/2}]/z$.

Substituting into (A.2) gives

$$\tilde{F} = 4z[(1+z)/(1-z)]^{1/2} \sum_{j=1}^{\infty} x^j/(1+x^j) \tag{A.4}$$

The sum at the right in (A.4) is related, by changing the plus sign to a minus sign, to a power series of number theory, with the coefficient d_j of x^j being the number of divisors of j (including unity). Thus the asymptotic expansion[11] of the sum of first n d_j gives the first term in the asymptotic expansion of the sum at the right of (A.4) in half-integral powers of $(1-z)$, and using the Tauberian theorem given in Ref. 1.

$$\sum_{j=1}^{n} \tilde{F}_j \sim 4n \log 2 + O(n^{1/2})$$

In one dimension E_n is asymptotically[1]

$$E_n \sim (8n/\pi)^{1/2} + O(n^{-1/2}) \tag{A.5}$$

Thus, to leading order, we can neglect the contribution of E_n to T_n [given in (5)] and

$$V_n \sim (4 \log 2 - 8/\pi)\, n + O(n^{1/2})$$
$$= 0.22610963,...,\, n + O(n^{1/2}) \tag{A.6}$$

Two Dimensions

For the separable random walk, the generating function for $P(\mathbf{x})$ at the point $(0, 0)$ is simply related[7] to the complete elliptic integral $K(z)$, and a double integral equals $P(\mathbf{x})$ at all other points.[1] In the limit that $z \to 1^-$, $K(z)$ goes to infinity as $\frac{1}{2} \log[8/(1-z)] + O[(1-z)\log(1-z)]$.

It will be shown that to get the three leading terms in the asymptotic expansion of V_n, one needs the first five terms in the asymptotic expansion of T_n that are found from the leading order term in the expansion in powers and logarithms of each of the first five $\tilde{F}^{(m)}$. For $\tilde{F}^{(1)}$ through $\tilde{F}^{(5)}$ one obtains the first term of the asymptotic expansions using the leading order behavior of $P(\mathbf{x})$ for $|\mathbf{x}| \gg 1$.[8,12]

$$\begin{aligned}
\tilde{F}^{(1)} &\sim 2\pi^2(1-z)^{-2} \{\log[8/(1-z)]\}^{-2} \\
\tilde{F}^{(2)} &\sim -2\pi^2(1-z)^{-2} \{\log[8/(1-z)]\}^{-3} \\
\tilde{F}^{(3)} &\sim 8a_3\pi^2(1-z)^{-2} \{\log[8/(1-z)]\}^{-4} \\
\tilde{F}^{(4)} &\sim -16a_4\pi^2(1-z)^{-2} \{\log[8/(1-z)]\}^{-5} \\
\tilde{F}^{(5)} &\sim 32a_5\pi^2(1-z)^{-2} \{\log[8/(1-z)]\}^{-6}
\end{aligned} \tag{A.7}$$

with

$$a_m = \int_0^\infty dr\, r[K_0(r)]^m$$

$$a_3 = 0.58597681,... \qquad a_4 = 1.0517998,... \qquad a_5 = 2.4965992,...$$

where $K_0(r)$ is the modified Bessel function of order zero. The integrals a_1 and a_2 are 1 and $\frac{1}{2}$, and a_3 can be transformed to a simpler integral

$$a_3 = -\frac{1}{2} \int_0^1 dx \log x/(1 - x + x^2)$$

Thus the exact equivalence of coefficient of the leading term of this solution with the previously given coefficient,[4] which includes the latter integral, is easily proved. The numerical value of the integral is known.[13] The integrals a_m, $m \geqslant 3$, were performed numerically using the GAUS8 and

BESK0 routines from the Common Los Alamos National Laboratory Mathematical Software.

The asymptotic behavior of the sum of the first n coefficients of the generating functions $\tilde{F}^{(m)}$ is found by dividing (A.7) by $1-z$ and saddle point integration using the method of Ref. 14. If a generating function has the form

$$A(1-z)^{-3}\{\log[8/(1-z)]\}^{-l}$$

with A a constant, the asymptotic expansion for the nth coefficient is

$$An^2(\log 8n + d_y)^{-l}(y!)^{-1}|_{y=2}$$

where d_y is differentiation with respect to y and the expression is to be expanded as a power series in $d_y/\log 8n$. Thus, (A.7) contains the contributions to the asymptotic expansion of T_n of orders $n^2[\log(8n)]^{-2}$ through $n^2[\log(8n)]^{-6}$, all neglected contributions to T_n being $O\{n^2[\log(8n)]^{-6}\}$. Collecting terms gives

$$\begin{aligned} T_n \sim (\pi n)^2 [\log(8n)]^{-2} \{1 &+ 0.84556867,..., [\log(8n)]^{-1} \\ &+ 0.94534482,..., [\log(8n)]^{-2} - 4.3489012,..., [\log(8n)]^{-3} \\ &+ 8.8781213,..., [\log(8n)]^{-4} + O[\log(8n)]^{-5}\} \end{aligned} \tag{A.8}$$

The leading terms in the expansion of E_n are those given[14] for S_n

$$\begin{aligned} E_n \sim \pi n[\log(8n)]^{-1} \{1 &+ 0.42278434,..., [\log(8n)]^{-1} \\ &- 0.46618747,..., [\log(8n)]^{-2} - 1.1465466,..., [\log(8n)]^{-3} \\ &- 0.58925976,..., [\log(8n)]^{-4} + O[\log(8n)]^{-5}\} \end{aligned} \tag{A.9}$$

The first two terms in the expansions of T_n and $(E_n)^2$ are identical; therefore

$$\begin{aligned} V_n \sim n^2\{&16.768193,..., [\log(8n)]^{-4} - 16.399478,..., [\log(8n)]^{-5} \\ &+ 106.67852,..., [\log(8n)]^{-6} + O[\log(8n)]^{-7}\} \end{aligned} \tag{A.10}$$

These methods will yield subsequent coefficients of the terms in this expansion.

Three Dimensions

$P(0, 0, 0)$ is related to the square of the complete elliptic integral $K(z')$, and[12]

$$\begin{aligned} P(0,0,0)+1 &\sim 1.3932039,\ldots,\ -0.90031632,\ldots,\ (1-z)^{1/2} \\ &\quad +0.84205258,\ldots,\ (1-z)+O(1-z)^{3/2} \\ &= c_0 - c_1(1-z)^{1/2} + c_2(1-z) + O(1-z)^{3/2} \qquad \text{(A.11)} \end{aligned}$$

The leading terms in the asymptotic expansion of $T_n - E_n$ in functions of $1-z$ as $z \to 1^-$ can be found by considering the $\tilde{F}^{(m)}$. $\tilde{F}^{(1)}$ and $\tilde{F}^{(2)}$ are found from the triple integral[1] equal to $P(\mathbf{x})$, and their expansions follow from (A.11). Including the origin in the summation over lattice points in (A.3) cancels when one considers $\tilde{F}^{(1)} + \tilde{F}^{(2)}$

$$\begin{aligned} \tilde{F}^{(1)} + \tilde{F}^{(2)} &\sim 2c_0^{-2}(1-z)^{-2} + 3c_1 c_0^{-3}(1-z)^{-3/2} \\ &\quad + c_0^{-2}\{3c_1^2 c_0^{-2} - 2(c_2/c_0) - 4\}(1-z)^{-1} + O(1-z)^{-1/2} \end{aligned}$$

With $m \geqslant 3$, it is most direct to numerically perform the lattice sum indicated in (A.3).

$$\begin{aligned} \tilde{F}^{(3)} &\sim [2/(1-z)\,c_0] \left\{ \sum_{\mathbf{x};R \geqslant |\mathbf{x}|}' [P(\mathbf{x})/c_0]^3 \Big|_{z=1} \right. \\ &\quad \left. + \lim_{z\to 1} \pi \int_R^\infty dr\, r^2 [(2/\pi r c_0)\exp(-r[2(1-z)/z]^{1/2})]^3 \right\} \\ &\sim [8/c_0^4 \pi^2 (1-z)] \log[1/(1-z)] - (0.8976,\ldots)(1-z)^{-1} \end{aligned}$$

$$\begin{aligned} \sum_{m=4}^{\infty} \tilde{F}^{(m)} &\sim [-2/(1-z)\,c_0] \left\{ \sum_{\mathbf{x};R \geqslant |\mathbf{x}|}' [P(\mathbf{x})/c_0]^4/[1+P(\mathbf{x})/c_0] \Big|_{z=1} \right. \\ &\quad \left. + (8/\pi^2 c_0^3)\log(1+2/\pi c_0 R) \right\} \\ &\sim -(0.14325,\ldots)(1-z)^{-1} \end{aligned}$$

In both cases the summation with $R \geqslant |\mathbf{x}|$ was performed in two parts; for $32 \geqslant |\mathbf{x}| \geqslant 1$, an integral for $P(\mathbf{x})$ with $z=1$ was evaluated numerically. An asymptotic formula[8,12,15] for $P(\mathbf{x})$ for large $|\mathbf{x}| = r$

$$P(\mathbf{x})|_{z=1} \sim (2/\pi r)[1 + (3/4r^2) - (5/4r^6)(x^4 + y^4 + z^4) + O(r^{-4})]$$

was then used for $2048 \geqslant |\mathbf{x}| \geqslant 32$. The second term in the brace in each case results from using only the leading term in the asymptotic formula for $P(\mathbf{x})$ given above and its analogue[8] for $1 > z$ and replacing the summation by integration for $R = |\mathbf{x}| \geqslant 2048$. The term containing $\log[1/(1-z)]$ is the resulting leading term in the asymptotic expansion of the exponential integral. There was no change in the first five places of the numerically derived coefficients whether $R = 1024$ or $R = 2048$.

Therefore, the generating function for $T_n - E_n$ is

$$\sum_{n=1}^{\infty} (T_n - E_n) z^n \sim 1.0303876,..., (1-z)^{-3} + 0.99878567,..., (1-z)^{-5/2}$$
$$+ 0.21514511,..., (1-z)^{-2} \log[1/(1-z)]$$
$$- (0.8976,..., + 0.1432,..., + 2.0381,...)(1-z)^{-2}$$
$$+ O(1-z)^{-3/2}$$

with the coefficients and terms not given above arising from $\tilde{F}^{(1)}$ and $\tilde{F}^{(2)}$. The asymptotic expansion for T_n is found by considering the asymptotic coefficients in the power series for the factors of $(1-z)^{-j}$ in this equation except for the term containing $\log[1/(1-z)]$ to which the Tauberian theorem of Ref. 1 is applied. One must also add the asymptotic expansion[12] for E_n

$$E_n \sim 0.71777001,..., n + 0.52338443,..., n^{1/2}$$
$$- 0.13407925,..., + O(n^{-1/2}) \tag{A.12}$$

$$T_n \sim 0.51519379,..., n^2 + 0.75133930,..., n^{3/2}$$
$$+ 0.21514511,..., n \log n - 0.8156,..., n + O(n^{1/2}) \tag{A.13}$$

Since the first two terms in the expansions of $(E_n)^2$ and T_n are identical

$$V_n \sim 0.21514511,..., n \log n - 0.8970,..., n + O(n^{1/2}) \tag{A.14}$$

COMPUTATION OF THE VARIANCE OF THE RANGE (3 – *d*)

```
program var3 (tape3)
parameter (nfft=2**15,nmax=2*2048,nmxt=2048,nsum=2*165,nsmt=165)
dimension p(nmax,nsmt)
dimension c(0:4),e(nmax),f(nmax),g(nmax),q1(nmxt),q2(nmxt),v(nmax)
dimension wsave(4*nfft+15)
complex po(nfft), px(nfft), pr(nfft), pt(nfft)
call ebm
nmpo=nmax+1
nstp=nsmt+1
nmfo=nmax/4
```

```
      c(0)=96.0/(1.0*nfft)
      c(1)=48.0/(1.0*nfft)
      c(2)=16.0/(1.0*nfft)
      c(3)=24.0/(1.0*nfft)
      c(4)=12.0/(1.0*nfft)
      do 10 n=1,nmxt
      q1(n)=0.0
      q2(n)=0.0
 10   continue
      do 20 n=1,nfft
      po(n)=0.0
      pt(n)=0.0
 20   continue
      q1(1)=0.5
      p(1,1)=0.5
      do 50 n=2,nmxt
      q2(1)=q1(1)
      do 30 i=2,n
      q2(i)=0.5*(q1(i-1)+q1(i))
 30   continue
      do 40 i=1,n
      q1(i)=0.5*(q2(i)+q2(i+1))
 40   continue
      do 50 i=1,nsmt
      p(2*n-2,i)=q2(i+1)
      p(2*n-1,i)=q1(i)
 50   continue
      do 60 i=2,nstp
      p(nmax,i-1)=0.5*(q1(i-1)+q1(i))
 60   continue
      po(1)=1.0
      do 70 n=2,nmax,2
      po(n+1)=p(n-1,1)**3
 70   continue
      call cffti (nfft,wsave)
      call cfftf (nfft,po,wsave)
      do 80 n=1,nfft
      pr(n)=(1.0/po(n))
 80   continue
      coeff=c(4)
      do 110 i=1,nsmt
      px(1)=0.0
      do 90 n=3,nmax,2
      px(n-1)=0.0
      px(n)=p(n-1,i)*p(n-2,1)**2
 90   continue
      do 100 n=nmax,nfft
      px(n)=0.0
100   continue
      call cfftf (nfft,px,wsave)
      do 110 n=1,nfft
      pt(n)=pt(n)+coeff*(pr(n)-(1.0/(po(n)+px(n))))
110   continue
      do 140 i=1,nsmt
      do 140 j=1,i
      ij=shiftr((i.eq.j),63)
      coeff=c(1+2*ij)
      px(1)=0.0
      do 120 n=3,nmax,2
      px(n-1)=0.0
      px(n)=p(n-1,i)*p(n-1,j)*p(n-2,1)
120   continue
      do 130 n=nmax,nfft
      px(n)=0.0
130   continue
      call cfftf (nfft,px,wsave)
      do 140 n=1,nfft
      pt(n)=pt(n)+coeff*(pr(n)-(1.0/(po(n)+px(n))))
140   continue
      do 210 i=1,nsmt
```

```
      do 210 j=1,i
      do 210 k=1,j
      ij=shiftr((i.eq.j),63)
      jk=shiftr((j.eq.k),63)
      coeff=c(ij+jk)
      do 150 n=2,nmax,2
      px(n-1)=0.0
      px(n)=p(n-1,i)*p(n-1,j)*p(n-1,k)
  150 continue
      do 160 n=nmpo,nfft
      px(n)=0.0
  160 continue
      call cfftf (nfft,px,wsave)
      do 170 n=1,nfft
      pt(n)=pt(n)+coeff*(pr(n)-(1.0/(po(n)+px(n))))
  170 continue
      px(1)=0.0
      do 180 n=3,nmax,2
      px(n-1)=0.0
      px(n)=p(n-1,i)*p(n-1,j)*p(n-1,k)
  180 continue
      do 190 n=nmax,nfft
      px(n)=0.0
  190 continue
      call cfftf (nfft,px,wsave)
      do 200 n=1,nfft
      pt(n)=pt(n)+coeff*(pr(n)-(1.0/(po(n)+px(n))))
  200 continue
  210 continue
      call cfftb (nfft,pt,wsave)
      f(2)=0.125
      do 220 n=4,nmax,2
      s=0.0
      do 220 m=4,n,2
      s=s+f(m-2)*p((n+1)-m,1)**3
      f(n)=p(n-1,1)**3-s
  220 continue
      e(1)=1.0
      do 230 n=2,nmax
      f(n)=f(n)+f(n-1)
      e(n)=1.0-f(n-1)
  230 continue
      g(1)=0.0
      do 240 n=2,nmax
      g(n)=g(n-1)+real(pt(n))
      if (mod(n,4).eq.0) write (3,290) g(n-3),g(n-2),g(n-1),g(n),n
  240 continue
      g(1)=1.0
      do 250 n=2,nmax
      g(n)=g(n)+e(n)+g(n-1)
      e(n)=e(n)+e(n-1)
      v(n)=g(n)-e(n)**2
  250 continue
      do 260 n=1,nmfo
      write (3,290) g(4*n-3),g(4*n-2),g(4*n-1),g(4*n),4*n
  260 continue
      do 270 n=1,nmxt
      write (3,290) e(2*n-1),e(2*n-1)**2,e(2*n),e(2*n)**2,2*n
  270 continue
      do 280 n=1,nmfo
      write (3,290) v(4*n-3),v(4*n-2),v(4*n-1),v(4*n),4*n
  280 continue
      stop
  290 format (4(1pe20.12,2x),i5)
      end

c     The function cfftf(n,c,wsave) calculates C(j) for j=1,2,...,n
c     with C(j)=the sum from k=0,1,...,n-1 of
c               C(k)*exp(-i*j*k*2*PI/n),
c     where i=sqrt(-1).
c     The function cfftb is identical except that -i is replaced with i.
c     The result is returned in the input vector.  The function cffti
c     initializes arrays used by cfftf and cfftb.  These functions are
c     part of the Common Los Alamos Mathematical Software.
```

REFERENCES

1. E. W. Montroll, *Proc. Symp. Appl. Math. Am.* **16**:193–220 (1964).
2. A. Dvoretzky and P. Erdös, *Proceedings of the 2nd Berkeley Symposium on Mathematical Statistics and Probability* (University of California Press, Berkeley, 1951), pp. 353–367.
3. W. Feller, *Ann. Math. Stat.* **22**:427–432 (1951).
4. N. C. Jain and W. E. Pruitt, *Proceedings of the 6th Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 3 (University of California Press, Berkeley, 1972), pp. 31–50.
5. N. C. Jain and W. E. Pruitt, *J. Analyse Math.* **24**:369–393 (1971).
6. N. C. Jain and W. E. Pruitt, *Z. Wahrsch. Verw. Gebiete* **16**:279–292 (1970).
7. G. Pólya, *Math. Ann.* **84**:149–160 (1921).
8. E. W. Montroll and G. H. Weiss, *J. Math. Phys.* **6**:167–181 (1965).
9. F. Spitzer, *Principles of Random Walk* (Springer-Verlag, New York, 1976), p. 4.
10. G. H. Weiss, *Proc. Nat. Acad. Sci. USA*, **77**:4391–4392 (1980).
11. G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers* (Oxford University Press, London, 1984), p. 265.
12. G. S. Joyce, *J. Math. Phys.* **12**:1390–1414 (1971).
13. I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series and Products*, 2nd ed. (Academic Press, New York, 1980), p. 533.
14. F. S. Henyey and V. Seshadri, *J. Chem. Phys.* **76**:5530–5534 (1982).
15. R. J. Duffin, *Duke Math. J.* **20**:233–251 (1953).