—NOTES—

ON SOME PROBABILITIES IN CONTINUOUS RANDOM WALK*

By Z. A. MELZAK (*University of British Columbia, Vancouver*)

1. Introduction. By a (continuous) n-step random walk we shall understand a sequence of n straight unit segments, called steps, in the plane; the first step starts at the origin and each successive step starts at the end of the previous one; every step is in random direction with uniform distribution in angle. This is historically the first random walk ever considered ([1], 1905). In the present article we exploit a certain method, well known in principle, and we compute some probabilities relating to an n-step random walk, such as: (a) the probable number of its self-intersections, (b) the probable number of its crossings by a fixed straight line, and (c) the probable number of intersections of two such independent random walks. The method in question could be called the "method of characteristic event" and goes back at least as far as the work of Markov [2]; its essence is to express the desired probability as an integral $P = \int_R f\, dx$, where R is a complicated subset of the n-dimensional cube C, and then to replace that integral by $\int_C fg\, dx$, where g is the characteristic function of R. A suitable integral representation $g = \int_D h\, du$ is used, giving $P = \int_C \int_D fh\, du\, dx$, the interchange of the order of integrations is justified, and the integrations are then carried out as far as possible.

2. Expected number of self-intersections. Neglecting certain events of probability 0, we define a self-intersection as an event in an n-step random walk when for some i and j, with $1 \leq i < j \leq n$ and $j > i + 1$, the ith and the jth step have in common exactly one point which is interior to each step. We suppose that $n \geq 3$ and we let $f(n)$ be the expected number of self-intersections; it will be proved that

$$f(n) = \frac{n}{4} \sum_{p=2}^{n-1} (1 - p/n)\Big[1 - 4\pi^{-2} \int_0^\infty \int_0^\infty (uv)^{-1}[J_0^{p-1}(u - v) - J_0^{p-1}(u + v)]\, du\, dv$$

$$+ \pi^{-5} \int_0^\infty \int_0^\infty \int_0^\infty \int_0^\infty \int_0^{2\pi} (uvwz)^{-1} \sum_{i=1}^{8} \epsilon_i \cos(c_i \sin\theta) J_0^{p-1}$$

$$((a_i^2 + 2\eta_i a_i b_i \cos\theta + b_i^2)^{1/2})\, d\theta\, du\, dv\, dw\, dz\Big]. \qquad (1)$$

Here J_0 is the Bessel function of the first kind and order 0, and the quantities indexed by i are as follows:

* Received December 30, 1970.

i	ϵ_i	η_i	a_i	b_i	c_i	
1	1	-1	$w - z$	$u - v$	$v - w$	
2	1	1	$w - z$	$u - v$	$v + w$	
3	-1	-1	$w - z$	$u + v$	$v + w$	
4	-1	1	$w - z$	$u + v$	$v - w$	
						(2)
5	-1	-1	$w + z$	$u - v$	$v - w$	
6	-1	1	$w + z$	$u - v$	$v + w$	
7	1	-1	$w + z$	$u + v$	$v + w$	
8	1	1	$w + z$	$u + v$	$v - w$	

Further, for n large we have asymptotically

$$f(n) \sim 2\pi^{-2} n \log n. \tag{3}$$

Let the ith step be from (x_i, y_i) to (x_{i+1}, y_{i+1}) and let it make the angle α_i with the x-axis. The line L_i which carries the ith step has the equation

$$(y - y_i) \cos \alpha_i - (x - x_i) \sin \alpha_i = 0$$

which we write as $L_i(x, y) = 0$. We exclude from consideration the 0-probability event that (x_i, y_i) lies on an L_j with j other than $i - 1$ or i. Therefore the ith and the jth step intersect if and only if

$$L_i^i L_i^{i+1} < 0 \quad \text{and} \quad L_j^i L_j^{i+1} < 0$$

where L_i^j stands for $L_i(x_j, y_j)$. Hence the expression

$$D_{ij} = \tfrac{1}{4}[1 - \operatorname{sgn} L_i^j \operatorname{sgn} L_i^{j+1}][1 - \operatorname{sgn} L_j^i \operatorname{sgn} L_j^{i+1}] \tag{4}$$

is 1 if the ith and jth steps cut, and 0 otherwise. The total number $F = F(\alpha_1, \cdots, \alpha_n)$ of self-intersections is therefore

$$F = \sum_{j=3}^{n} \sum_{i=1}^{j-2} D_{ij} \tag{5}$$

and the expected number of self-intersections is

$$f(n) = (2\pi)^{-n} \int_0^{2\pi} \cdots \int_0^{2\pi} F \, d\alpha_1 \cdots d\alpha_n . \tag{6}$$

To represent the discontinuous factor sgn x we use the integral

$$\operatorname{sgn} x = 2\pi^{-1} \int_0^{\infty} u^{-1} \sin ux \, du. \tag{7}$$

We have

$$x_i = \sum_{k=1}^{i-1} \cos \alpha_k , \qquad y_i = \sum_{k=1}^{i-1} \sin \alpha_k$$

so that

$$L_i^j = S(i + 1, j - 1; i), \qquad L_i^{j+1} = S(i + 1, j; i),$$
$$L_j^i = -S(i, j - 1; j), \qquad L_j^{i+1} = -S(i + 1, j - 1; j)$$

where we use the abbreviation

$$S(p, q; m) = \sum_{k=p}^{q} \sin (\alpha_k - \alpha_m). \tag{8}$$

By (7) we have therefore

$$\operatorname{sgn} L_i^i \operatorname{sgn} L_i^{i+1} = J_2 , \qquad \operatorname{sgn} L_j^i \operatorname{sgn} L_j^{i+1} = J_3$$

where

$$J_2 = 4\pi^{-2} \int_0^\infty \int_0^\infty (uv)^{-1} \sin [uS(i+1, j-1; i)] \sin [vS(i+1, j; i)]\, du\, dv, \tag{9}$$

$$J_3 = 4\pi^{-2} \int_0^\infty \int_0^\infty (uv)^{-1} \sin [uS(i, j-1; j)] \sin [vS(i+1, j-1; j)]\, du\, dv. \tag{10}$$

We also let $J_1 = 1$ and

$$J_4 = 16\pi^{-4} \int_0^\infty \int_0^\infty \int_0^\infty \int_0^\infty (uvwz)^{-1} \sin [uS(i+1, j-1; i)] \sin [vS(i+1, j; i)]$$
$$\cdot \sin [wS(i, j-1; j)] \sin [zS(i+1, j-1; j)]\, du\, dv\, dw\, dz. \tag{11}$$

Now (6) becomes

$$f(n) = \tfrac{1}{4}(2\pi)^{-n} \sum_{j=3}^{n} \sum_{i=1}^{j-2} (I_1 - I_2 - I_3 + I_4) \tag{12}$$

where

$$I_s = \int_0^{2\pi} \cdots \int_0^{2\pi} J_s\, d\alpha_1 \cdots d\alpha_n , \qquad s = 1, 2, 3, 4. \tag{13}$$

We evaluate the integral I_2. To begin with, we show that

$$\int_0^{2\pi} \cdots \int_0^{2\pi} \int_0^\infty \int_0^\infty \cdots du\, dv\, d\alpha_1 \cdots d\alpha_n$$
$$= \int_0^\infty \int_0^\infty \int_0^{2\pi} \cdots \int_0^{2\pi} \cdots d\alpha_1 \cdots d\alpha_n\, du\, dv, \tag{14}$$

the integrand being the same as in (9). This follows by a slight extension, to take care of multiple integrations, of the criterion given in [3]. First, if we replace the infinite upper limits of integrations by any finite ones, then (14) holds. The integral J_2, as a function of $\alpha_1, \cdots, \alpha_n$, is boundedly convergent over the cube $H: 0 \leq \alpha_i \leq 2\pi$, $i = 1, \cdots, n$. If the intersections of H with the hyperplanes $\alpha_k = \alpha_i$ ($k = i + 1, \cdots, j$) are removed from H then J_2 is uniformly convergent on any compact subset of the remainder of H. The removed set has n-dimensional content 0, and thus (14) is justified. Therefore

$$I_2 = 4\pi^{-2} \int_0^\infty \int_0^\infty (uv)^{-1} G(u, v)\, du\, dv \tag{15}$$

where

$$G(u, v) = \int_0^{2\pi} \cdots \int_0^{2\pi} \sin [uS(i+1, j-1; i)] \sin [vS(i+1, j; i)]\, d\alpha_1 \cdots d\alpha_n . \tag{16}$$

We put $p = j - i$ and we introduce new variables

$$\theta_1 = \alpha_{i+1} - \alpha_i , \qquad \theta_2 = \alpha_{i+2} - \alpha_i , \qquad \theta_p = \alpha_j - \alpha_i , \qquad \theta_{p+1} = \alpha_i ,$$

the Jacobian of the transformation is 1; now, using the periodicity of the sine, we can write (16) as

$$G(u, v) = 2\pi^{-2}(2\pi)^{n-p} \operatorname{Re} \left\{ \int_0^{2\pi} \cdots \int_0^{2\pi} \left[\exp \left(i(u - v) \sum_{k=1}^{p-1} \sin \theta_k - iv \sin \theta_p\right) \right.\right.$$
$$\left.\left. - \exp \left(i(u + v) \sum_{k=1}^{p-1} \sin \theta_k + iv \sin \theta_p\right) \right] d\theta_1 \cdots d\theta_p \right\}.$$

Recalling the Poisson integral for $J_0(x)$, [4], we evaluate the above and we use this together with (15) to obtain

$$I_2 = 2\pi^{-2}(2\pi)^n \int_0^\infty \int_0^\infty (uv)^{-1} J_0(v)[J_0^{p-1}(u - v) - J_0^{p-1}(u + v)]\, du\, dv. \tag{17}$$

By the symmetry of the integrands in (13) for $s = 2$ and $s = 3$ we find

$$I_2 = I_3 . \tag{18}$$

To evaluate I_4 we proceed similarly. First the interchange of the order of integrations is justified as before, then we pass over to the same new variables $\theta_1 , \cdots , \theta_{p+1}$, next we use some trigonometry to transform a product of four sines into a sum of eight cosines, and we get

$$I_4 = 16\pi^{-4} \int_0^\infty \int_0^\infty \int_0^\infty \int_0^\infty (uvwz)^{-1} H(u, v, w, z)\, du\, dv\, dw\, dz \tag{19}$$

where

$$H = \tfrac{1}{8}(2\pi)^{n-p} \sum_{i=1}^{8} K_i$$

and each K_i is of the form

$$K_i = \epsilon_i \int_0^{2\pi} \cdots \int_0^{2\pi} \cos \left[a_i \sum_{k=1}^{p-1} \sin \theta_k + b_i \sum_{k=1}^{p-1} \sin (\theta_k - \theta_p) + c_i \sin \theta_p \right] d\theta_1 \cdots d\theta_p ,$$

with a suitable signature $\epsilon_i = \pm 1$ and a_i , b_i , c_i expressed in terms of u, v, w, z. Some elementary though rather lengthy trigonometric manipulations enable us to carry out successively the integrations with respect to $\theta_1 , \cdots , \theta_{p-1}$, though not with respect to θ_p , culminating in

$$H = (16\pi)^{-1} \int_0^{2\pi} \sum_{i=1}^{8} \epsilon_i J_0^{p-1}((a_i^2 + 2\eta_i a_i b_i \cos \theta + b_i^2)^{1/2}) \cos (c_i \sin \theta)\, d\theta \tag{20}$$

where the quantities indexed by i are given in (2). Since I_1 is trivially $(2\pi)^n$, we put together (12), (17), (18), (19), (20) and obtain $f(n)$ as a double series of the form

$$f(n) = \tfrac{1}{4} \sum_{j=3}^{n} \sum_{i=1}^{j-2} \gamma(p)$$

where, as throughout, $p = j - i$; changing now to a single series we get (1).

3. Asymptotic formula. For the proof of (3) we write (1) as

$$f(n) = n \sum_{p=2}^{n-1} (1 - p/n)\, d_p \tag{21}$$

and observe that d_p has an important probabilistic interpretation: letting i/j stand for the event that the ith and the jth steps intersect, we have $d_p = \text{Prob}\ (i/i + p)$. It will be shown that for large p

$$d_p \sim 2\pi^{-2} p^{-1} \tag{22}$$

which together with (21) proves (3). Let $s_j^2 = x_j^2 + y_j^2$; observe that $d_{n-1} = \text{Prob}\ (1/n)$ and the event $1/n$ can occur only if $s_{n-1} \leq 2$. Therefore by an elementary use of conditional probabilities

$$d_{n-1} = \text{Prob}\ (s_{n-1} \leq 2)\ \text{Prob}\ (1/n \mid s_{n-1} \leq 2). \tag{23}$$

By Kluyver's theorem on random walks [4], [5],

$$\text{Prob}\ (s_{n-1} \leq r) = r \int_0^\infty J_1(rx) J_0^{n-1}(x)\, dx$$

and so, by the standard asymptotics of the Bessel functions [4], we have for n large

$$\text{Prob}\ (s_{n-1} \leq r) \sim r^2/n. \tag{24}$$

Hence in particular

$$\text{Prob}\ (s_{n-1} \leq 2) \sim 4/n. \tag{25}$$

To obtain the other term on the right-hand side of (23) we observe the following: the asymptotic formula (24) implies that for large n the hitting probability after $n - 1$ steps is asymptotically uniform over the circle $s_{n-1} \leq 2$. In other words, for large n the probability that (x_n, y_n) lies in a measurable subset A of that circle is asymptotically of the form const. times n^{-1} times the measure of A. Hence the conditional probability $\text{Prob}\ (1/n \mid s_{n-1} \leq 2)$ tends to a limit P as n increases. It is now not hard to show that the limit P is the probability that a unit step, in a random direction and from a random point inside the circle of radius 2 about the origin, intersects the interval [0, 1] on the x-axis. Expressed thus, P can be calculated by elementary means to be $(2\pi^2)^{-1}$, so that for large n

$$\text{Prob}\ (1/n \mid s_{n-1} \leq 2) \sim (2\pi^2)^{-1}.$$

This together with (23) and (25) gives us (22) and so (3) is proved.

4. Two independent random walks. Formulas (1) and (3) allow us to answer the following question: in two independent continuous random walks, of n and m steps respectively, what is the expected number $f(n, m)$ of times the two paths cross? It is easy to show by symmetry considerations that the answer is obtained by taking a single random walk of $n + m$ steps, and subtracting from the number of its self-crossings the numbers of self-crossings of the first n and of the last m steps, giving us $f(n, m) = f(n + m) - f(n) - f(m)$ with f given by (1). By (3) one gets for a fixed t and n large

$$f(n, tn) \sim 2\pi^{-2} n[\log (1 + t) + t \log (1 + t^{-1})]$$

and in particular $f(n, n) \sim \pi^{-2} \log 16\, n$.

5. Line crossings. In much the same way, though rather more simply, it is possible to find a formula for the expected number $f(n, b)$ of crossings of an n-step random walk by a fixed line L whose distance from the origin is b. Without loss of generality we take L to have the equation $y = b$; proceeding as before we find that the expression

$$\tfrac{1}{2}[1 - \operatorname{sgn}(y_{j+1} - b)\operatorname{sgn}(y_j - b)]$$

is 1 if the jth step cuts L, and 0 otherwise. Therefore

$$f(n, b) = (n/2) - \tfrac{1}{2}(2\pi)^{-n} \sum_{j=1}^{n} \int_0^{2\pi} \cdots \int_0^{2\pi} \operatorname{sgn}(y_{j+1} - b)\operatorname{sgn}(y_j - b)\, d\alpha_1 \cdots d\alpha_n .$$

We express y_j and y_{j+1} in terms of the angles $\alpha_1, \cdots, \alpha_n$, use again the integral representation (7), justify the interchange of the order of integrations as before, and perform the angle integrations, to obtain

$$f(n, b) = (n/2) - \pi^{-2} \sum_{j=0}^{n-1} \int_0^{\infty} \int_0^{\infty} (uv)^{-1} J_0(v)$$
$$\cdot \{\cos[b(u - v)]J_0^j(u - v) - \cos[b(u + v)]J_0^j(u + v)\}\, du\, dv\cdot$$

6. Acknowledgements. The author is grateful to Mr. Robert Main for pointing out how to express $f(n, m)$ of Sec. 4 by means of $f(n)$ of Sec. 1, to the National Research Council of Canada for supporting the research reported upon, and to the referees for some suggestions for general improvement.

References

[1] K. Pearson, Nature **72**, 292, 318, 342 (1905)
[2] A. A. Markov, *Wahrscheinlichtkeitsrechnung,* Leipzig, 1912
[3] E. C. Titchmarsh, *The theory of functions,* 2nd ed., Oxford Univ. Press, Oxford, 1939.
[4] G. N. Watson, *Theory of Bessel functions,* 2nd ed., Cambridge Univ. Press, Cambridge, 1958
[5] J. C. Kluyver, Proc. Soc. Sci. Köningl. Akad. Wetensch. Amsterdam 341–350 (1906)