SOME INTERSECTION PROPERTIES OF RANDOM WALK PATHS*

By

P. ERDŐS (Budapest), corresponding member of the Academy,
and S. J. TAYLOR (Birmingham)

1. Introduction. We are concerned with properties of the infinite symmetric random walk over the lattice of points with integer coordinates in Euclidean space of d dimensions. (For a precise definition see [4].) We consider two problems.

PROBLEM A. *Suppose $\Pi^{(d)}(a, b)$ denotes the set of points $S_d(n)$ $(a \leqq n \leqq b)$ of the random walk in d-space; suppose $f(n)$ is an increasing function of n. What are the conditions on the rate of increase of $f(n)$ which are necessary and sufficient to ensure that the sets*

$$\Pi^{(d)}(0, n), \quad \Pi^{(d)}(n+f(n), \infty)$$

have points in common for infinitely many values of n with probability 1?

We complete the solution of this problem in Section 3. Clearly, there is no problem for $d=1$ or 2. The solution takes a different form in the cases $d=3$, $d=4$, and $d \geqq 5$. For example, if $d=4$, an interesting consequence of the result is that, with probability 1, there are infinitely many n for which $\Pi^4(0, n)$ and $\Pi^4(2n, \infty)$ have a point in common. This in turn implies that any two independent random walks in 4-space have infinitely many points in common. This at first surprised us, because two independent Brownian motion paths in 4-space have no points in common with probability 1 (this follows from the result of [2]). The explanation is as follows: with probability 1, two independent Brownian paths in 4-space approach arbitrarily close to each other for arbitrarily large values of t; thus they have infinitely many near misses, but fail to intersect because the fine structure of the paths is not sufficiently dense (in fact, the paths have zero 2-dimensional measure, see [5]). It can be shown similarly that, with probability 1, 3 independent random walks in 3-space have infinitely many points in common, while 3 Brownian motion paths have no common points (this last follows from the result of [3]).

* This paper and the paper [4] were written while P. ERDŐS was visiting the University of Birmingham.

PROBLEM B. *A point $S_d(n)$ of a random walk path is called 'good' if there are no points common to $\Pi^{(d)}(0, n)$ and $\Pi^{(d)}(n+1, \infty)$. For $d=1$ or 2 there are no good points with probability 1. For $d \geqq 3$, the result of Pólya implies that there must be some good points: how many are there?*

We prove, in Section 4, that there are absolute constants τ_d $(d \geqq 5)$ such that, with probability 1, a random walk in d-space has good points at a subsequence of density τ_d. For $d=3$ or 4 the subsequence of good points has density zero with probability 1. We obtain in these cases asymptotic bounds for the number of good points.

We start, in Section 2, by obtaining some preliminary results, and collecting results which are already known but are needed in the sequel.

2. Preliminary results. If E is a condition on the random walk path, we write $\mathbf{P}(E)$ for the probability that the condition is satisfied. If $E_1, E_2, \ldots$ $\ldots, E_k, \ldots$ is a sequence of conditions, we write

$$\mathbf{P}\{E_k \ i.o.\}$$

for the probability that the path satisfies infinitely many of the conditions E_k.

$\mathcal{E}\{Q\}$, $\sigma^2\{Q\}$ denote the mean and variance of a random variable Q.

$[x]$ denotes the largest integer not greater than the real number x.

ε will always denote a positive number.

$c_1, c_2, \ldots, c_{38}$ will denote suitable finite positive real constants.

If X is a vector in d-space, $|X|$ denotes the distance from X to the origin.

For paths in d-space, $\gamma_d(n)$ denotes the probability that in the first $n-1$ steps, the path does not return to the origin. It is proved in [1] that, for $d \geqq 3$, there are positive constants γ_d such that

$$\gamma_d < \gamma_d(n) < \gamma_d + O(n^{1-d/2}). \tag{2.1}$$

$S_d(n)$ denotes the position at the n^{th} step of a random walk in d-space. If L is any lattice point in d-space,

$$u_d(L, n) = \mathbf{P}\{S_d(n) = L\}.$$

Clearly, all points can be reached either only in an even number of steps or only in an odd number of steps, We need the following easy estimates for $u_d(L, n)$. Suppose L is a point which can be reached in an even number of steps, and $|L| = \varrho$. Then

(i) if $\varrho = 0$, we have

$$u_d(0, 2m) = 2\left(\frac{d}{4m\pi}\right)^{d/2} + O\left(\frac{1}{m^{1+d/2}}\right); \tag{2.2}$$

(ii) if $m > \varrho^2$,

$$u_d(L, 2m) = 2\left(\frac{d}{4m\pi}\right)^{d/2}\left[1 + O\left(\frac{\varrho^2}{m}\right)\right]; \tag{2. 3}$$

(iii) if $m > \frac{1}{20}\varrho^2$, there are constants c_1, c_2 with

$$\frac{c_1}{m^{d/2}} > u_d(L, 2m) > \frac{c_2}{m^{d/2}}; \tag{2. 4}$$

(iv) if $m < \varrho^2$, $d \geqq 2$,

$$u_d(L, 2m) \leqq 2\left(\frac{d}{4m\pi}\right)^{d/2} e^{-\varrho^2/2m}\left(1 + O\left(\frac{1}{m}\right)\right). \tag{2. 5}$$

The same asymptotic formulae are valid for $u_d(L, 2m+1)$ in the case where L can be reached in an odd number of steps.

We now need estimates for entering a point L at least once.

LEMMA 1. *Suppose L is a lattice point in d-space ($d \geqq 3$), with $|L| = \varrho > 0$, and $v_d(L)$ is the probability that an infinite random walk starting from O will enter L at least once. Then there are finite positive constants f_d, g_d such that*

$$\frac{f_d}{\varrho^{d-2}} < v_d(L) < \frac{g_d}{\varrho^{d-2}} \qquad (d = 3, 4, \ldots).$$

PROOF. We may clearly assume that $\varrho > 100$. Considering the last time of passage through L, we have

$$v_d(L) \geqq \gamma_d \sum_{n=1}^{\infty} u_d(L, n),$$

since $\gamma_d u_d(L, n)$ is the probability that the path is at L at the n^{th} step and does not return again to L and these events are mutually exclusive. Using (2. 4) we have

$$v_d(L) \geqq \gamma_d \sum_{m > \frac{1}{20}\varrho^2}^{\infty} \frac{c_2}{m^{d/2}} \geqq \frac{f_d}{\varrho^{d-2}} \quad \text{for a suitable } f_d > 0.$$

In the other direction, it is clear that

$$v_d(L) \leqq \sum_{n=1}^{\infty} u_d(L, n) = \sum_{n \leqq \frac{1}{20}\varrho^2} u_d(L, n) + \sum_{n > \frac{1}{20}\varrho^2} u_d(L, n) \leqq \frac{g_d}{\varrho^{d-2}}$$

for a suitable $g_d > 0$,

using (2. 5) and (2. 4).

This completes the proof of the lemma.

A modified form of the same proof suffices to prove

LEMMA 2. *Suppose L is a lattice point in d-space* ($d \geqq 3$) *with* $|L| = \varrho > 0$ *and* $w_d(L, n)$ *is the probability that a random walk of n steps starting from O will enter L at least once. Then there are finite constants* h_d *such that if* $n > \frac{1}{5}\varrho^2$, *then*

$$w_d(L, n) > \frac{h_d}{\varrho^{d-2}} \qquad (d = 3, 4, \ldots).$$

For $d \geqq 3$, we know that random walks wander off to infinity. We need estimates for the probability that they are not too far from O at some time in a given range of t.

LEMMA 3. *For every integer* $N \geqq 0$ *and real number* $r > 0$, *put*

$$Q_d(r, N) = \mathbf{P}\{|S_d(n)| \leqq r \text{ for some } n \geqq N\},$$

then we have for $d \geqq 3$, $N > r^2$

$$\frac{1}{2} e_d \left(\frac{r}{\sqrt{N}}\right)^{d-2} \leqq Q_d(r, N) \leqq e_d \left(\frac{r}{\sqrt{N}}\right)^{d-2}$$

for suitable $e_d > 0$.

This gives the probability of being within a distance r of the origin at some time after the N^{th} step. The corresponding result for Brownian motion is proved in [1]; the random walk result follows because of the relationship between random walk and Brownian motion.

A proof is also given in [1] for

LEMMA 4. *For every integer* $N \geqq 0$ *and real number* $r > 0$, *put*

$$P_d(r, N) = \mathbf{P}\{S_d(n) \leqq r \text{ for some } N \leqq n \leqq 4N\};$$

then for $d \geqq 3$, $P_d(r, N) > \frac{1}{10} e_d \left(\frac{r}{\sqrt{N}}\right)^{d-2}$.

For a random walk in d-space, we put $\varrho_d(n) = |S_d(n)|$. In [4] we briefly studied the average behaviour of $\varrho_d(n)$ and in doing so proved

LEMMA 5. *If* $\varrho_4(n)$ *is the distance from the origin at the* n^{th} *step of a random walk in* 4 *dimensions, then there exists a constant* $c_3 > 0$ *such that*

$$\mathcal{E}\left(\frac{1}{1 + \{\varrho_4(n)\}^2}\right) = \frac{c_3}{n}(1 + o(1)).$$

Similarly one can prove

LEMMA 6. *If $\varrho_3(n)$ is the distance from the origin at the n^{th} step of a random walk in 3-space, then there exist constants c_3', c_3'' such that*

$$\frac{c_3'}{\sqrt{n}} < \mathcal{E}\left\{\frac{1}{1+\varrho_3(n)}\right\} < \frac{c_3''}{\sqrt{n}}.$$

3. Solution of Problem 1. In order to estimate the probability of an intersection between $\Pi^{(d)}(0, n)$ and $\Pi^{(d)}(n+f(n), \infty)$ we need first of all to estimate the probability of at least one intersection between two independent random walks starting at different points. It is critical to suppose that if the starting points are separated by a distance ϱ, then one of the paths takes approximately ϱ^2 steps, while the other either takes ϱ^2 steps or is infinite.

LEMMA 7. *Suppose $\Pi_1^{(d)}(0, n)$, $\Pi_2^{(d)}(0, n)$ are independent random walk paths of n steps in d-space ($d \geqq 3$), the first starting at the origin and the second at P where $|P| = \varrho$ and $\frac{1}{2}\varrho^2 < n < 2\varrho^2$; then there are constants $c_4, c_5, c_6^{(d)}$ such that*

(i) $\mathbf{P}\{\Pi_1^{(3)}(0, n)$ *intersects* $\Pi_2^{(3)}(0, n)\} > c_4$,

(ii) $\mathbf{P}\{\Pi_1^{(4)}(0, n)$ *intersects* $\Pi_2^{(4)}(0, n)\} > \dfrac{c_5}{\log n}$,

(iii) $\mathbf{P}\{\Pi_1^{(d)}(0, n)$ *intersects* $\Pi_2^{(d)}(0, n)\} > \dfrac{c_6^{(d)}}{n^{(d-4)/2}}$ *for* $d = 5, 6, \ldots$.

PROOF OF (i). In Euclidean space of 3 dimensions consider a cube b with centre at the origin and side $\frac{1}{2}\varrho$. Let η be a large positive number and

$$t_3 = [\eta n^{1/6}]. \tag{3.1}$$

Let $\mathfrak{L}$ be the subset of the lattice points with integer coordinates obtained by taking points all of whose coordinates are multiples of t_3 which lie in b but are not within $\frac{1}{4}\varrho$ of the origin. The number $\nu(\mathfrak{L})$ of points of $\mathfrak{L}$ clearly satisfies

$$\nu(\mathfrak{L}) \geqq c_7 \frac{n}{\eta^3} \tag{3.2}$$

and the distance of any point L in $\mathfrak{L}$ from each of O, P lies between $\frac{1}{4}\varrho$ and 2ϱ.

Let $p_1(L)$ denote the probability that $\Pi_1^{(3)}(0, n)$ passes through L, and $p_2(L)$ the probability that $\Pi_2^{(3)}(0, n)$ passes through L. Since the paths are independent, the probability that L lies on both paths is $p_1(L)p_2(L)$. Using Lemmas 1 and 2 we have immediately that

$$\frac{4g_3}{\varrho} \geqq \begin{Bmatrix} p_1(L) \\ p_2(L) \end{Bmatrix} \geqq \frac{h_3}{2\varrho} \tag{3. 3}$$

so that

$$p_1(L)p_2(L) \geqq \frac{c_8}{n}. \tag{3. 4}$$

If interference could be neglected, (3. 2) and (3. 4) would be sufficient to obtain the desired result. We now show that provided η is chosen large enough, the interference is small. Let $p_1(L, M)$, $p_2(L, M)$, respectively, be the probabilities that $\Pi_1^{(3)}(0, n)$, $\Pi_2^{(3)}(0, n)$ pass through both L and M. Clearly, it is sufficient to show that

$$\sum_{L \in \mathfrak{L}} p_1(L)p_2(L) - \sum_{L, M \in \mathfrak{L}}\sum p_1(L, M)p_2(L, M) > c_4. \tag{3. 5}$$

In the notation of Lemma 1, $v_3(L)$ is the probability that an infinite path from the origin will pass through L. It is clear that

$$p_1(L, M) \leqq v_3(L-M)\big(p_1(L) + p_1(M)\big),$$
$$p_2(L, M) \leqq v_3(L-M)\big(p_2(L) + p_2(M)\big).$$

By (3. 3),

$$\sum p_1(L, M)p_2(L, M) \leqq \left(\frac{8g_3}{\varrho}\right)^2 \sum [v_3(L-M)]^2 < \frac{c_9}{n} \sum [v_3(L-M)]^2. \tag{3. 6}$$

For fixed $L \in \mathfrak{L}$, let $\nu_r(L)$ be the number of points of $\mathfrak{L}$ whose distance from L lies between 2^r and 2^{r+1} $(r = 1, 2, \ldots)$. It is clear that

$$\nu_r(L) \leqq c_{10} \frac{2^{3r}}{\eta^3 n^{1/2}}. \tag{3. 7}$$

For a point M whose distance from L is at least 2^r we know that

$$v_3(L-M) \leqq \frac{g_d}{2^r}.$$

Hence

$$\sum_{M \in \mathfrak{L}} v_3(L-M)^2 \leqq \sum_{1 \leqq 2^r \leqq \varrho} \nu_r(L) \frac{g_d^2}{2^{2r}} < \frac{c_{11}}{\eta^3},$$

using (3. 7). By (3. 6), it follows that, for each L in $\mathfrak{L}$,

$$\sum_{M \in \mathfrak{L}} p_1(L, M)p_2(L, M) < \frac{c_9}{n}\frac{c_{11}}{\eta^3} < \frac{1}{2} p_1(L)p_2(L),$$

by (3.3), provided η is chosen large enough. This together with (3.2) and (3.4) is clearly sufficient to establish (3.5).

PROOF OF (ii). A very similar argument will work using for $\mathfrak{L}$ only points whose co-ordinates are all divisible by

$$t_4 = [\eta(\log n)^{1/4}]$$

for large enough η.

PROOF OF (iii). Again the same argument works using this time points whose co-ordinates are all divisible by t_d where t_d is a sufficiently large fixed integer.

The simplest way of obtaining an upper bound corresponding to Lemma 7 (ii) seems to be a calculation of the expectation of the number of points common to the two paths by two different methods. We first obtain a lower bound for this expectation in the case where both paths start at the origin.

LEMMA 8. *Suppose $\Pi_1^{(d)}(0, n)$, $\Pi_2^{(b)}(0, n)$ are independent random walks of n steps in d-space ($d = 3$ or 4), both starting from the origin. Let $D^{(d)}(n)$ denote the number of points common to the two paths. Then there are constants c_{12}, c_{13} such that*

(i) $\mathcal{E}\{D^{(3)}(n)\} > c_{12} n^{1/2}$,

(ii) $\mathcal{E}\{D^{(4)}(n)\} > c_{13} \log n$.

PROOF OF (i). Consider the points P in 3-space with integer co-ordinates whose distance from O is less than $\sqrt{n}$. It is clear that if $v_3(P)$ is the probability that a path of n steps will enter P,

$$\mathcal{E}\{D^{(3)}(n)\} \geqq \sum_{1 \leqq |P| < n^{1/2}} [v_3(P)]^2 \geqq \sum_{1 \leqq 2^r < n^{1/2}} \sum_{2^{r-1} \leqq |P| < 2^r} [v_3(P)]^2.$$

If $2^r > |P|$, we have, by Lemma 2,

$$v_3(P) \geqq \frac{c_{14}}{|P|} > \frac{c_{14}}{2^r}.$$

As the number of points P with $2^{r-1} \leqq |P| < 2^r$ is at least $c_{15} 2^{3r}$, it follows immediately that

$$\mathcal{E}\{D^{(3)}(n)\} \geqq c_{12} n^{1/2}.$$

PROOF OF (ii). The same method works. As above,

$$\mathcal{E}\{D^{(4)}(n)\} \geqq \sum_{1 \leqq 2^r < n^{1/2}} \sum_{2^{r-1} \leqq |P| < 2^r} [V_4(P)]^2 \geqq c_{16} \sum_{r=0}^{\left[\frac{1}{2}\log_2 n\right]} 1 > c_{13} \log n.$$

REMARK 1. It is clear that there is a constant c_{17} such that for any n and $d \geqq 5$

$$1 \leqq \mathcal{E}\{D^{(d)}(n)\} < c_{17}.$$

REMARK 2. By carrying out the computations more carefully it is possible to prove that

$$\mathcal{E}\{D^{(3)}(n)\} = c_{18} n^{1/2}(1+o(1)),$$

$$\mathcal{E}\{D^{(4)}(n)\} = c_{19} \log n(1+o(1)),$$

$$\mathcal{E}\{D^{(d)}(n)\} = c_{20}^{(d)}(1+o(1)) \qquad (d \geqq 5),$$

for suitable constants $c_{18}, c_{19}, c_{20}^{(d)}$. We do not prove these as they are not required in the sequel.

Though we do not require it, we prove the following lemma to complete our results:

LEMMA 9. *Suppose $\Pi_1^{(d)}(0, \infty)$, $\Pi_2^{(d)}(0, \infty)$ are independent random walks in d-space ($d \geqq 4$), the first starting at O, the second at P where $|P| = \varrho$ and $\frac{1}{2}\varrho^2 < n < 2\varrho^2$; then there are constants c_{21} and $c_{22}^{(d)}$ ($d = 5, 6, \ldots$) such that*

(i) $\mathbf{P}\{\Pi_1^{(4)}(0, n)$ *intersects* $\Pi_2^{(4)}(0, \infty)\} < \dfrac{c_{21}}{\log n}$,

(ii) $\mathbf{P}\{\Pi_1^{(d)}(0, \infty)$ *intersects* $\Pi_2^{(d)}(0, \infty)\} < \dfrac{c_{22}^{(d)}}{n^{(d-4)/2}}$ *for* $d = 5, 6, \ldots$.

PROOF OF (i). Calculate the expected number of points common to $\Pi_1^{(4)}(0, n)$ and $\Pi_2^{(4)}(0, \infty)$. This turns out to be finite. However, if there is an intersection between $\Pi_1^{(4)}\left(0, \left[\frac{n}{2}\right]\right)$ and $\Pi_2^{(4)}(0, \infty)$, both paths can continue independently after the intersection, both of them for at least $\left[\frac{n}{2}\right]$ steps. By Lemma 8 (ii), the conditional expectation of the number of points common is at least $c_{13} \log \frac{n}{2}$ given that an intersection occurs. (i) now follows as otherwise the expected number of common points would not remain finite.

PROOF OF (ii). In the notation of Lemma 7, let $p_1(L), p_2(L)$ be the probabilities that $\Pi_1^{(d)}(0, \infty)$, $\Pi_2^{(d)}(0, \infty)$, respectively, pass through the lattice point L. By using Lemma 1, the result (ii) follows immediately on summing $p_1(L)p_2(L)$ over all the lattice points with integer co-ordinates.

The main result of this section is contained in

THEOREM 1. *Suppose $f(n)$ is an integer-valued function of n which increases to infinity as $n \to \infty$ and $E_n^{(d)}$ is the event that the random walk path in d-space is such that $\Pi^{(d)}(0, n)$ and $\Pi^{(d)}(n+f(n), \infty)$ have at least one point in common.*

(i) *For $d=3$, if $f(n)=n\{\varphi(n)\}^2$ and $\varphi(n)$ is monotonic increasing, then*

$$\mathbf{P}\{E_n^{(3)} \text{ i. o.}\}=0 \quad or \quad 1,$$

according as $\sum_{k=1}^{\infty} \frac{1}{\varphi(2^k)}$ *converges or diverges.*

(ii) *For $d=4$, if $f(n)=n\psi(n)$ and $\psi(n)$ is monotonic increasing, then*

$$\mathbf{P}\{E_n^{(4)} \text{ i. o.}\}=0 \quad or \quad 1,$$

according as $\sum_{k=1}^{\infty} \frac{1}{k\psi(2^k)}$ *converges or diverges.*

(iii) *For $d \geqq 5$, if* $\sup_{m \geqq n} \frac{f(m)}{m} \leqq c_{23} \frac{f(n)}{n}$,[1] *then*

$$\mathbf{P}\{E_n^{(d)} \text{ i. o.}\}=0 \quad or \quad 1,$$

according as $\sum_{n=1}^{\infty} \frac{1}{f(n)^{(d-2)/2}}$ *converges or diverges.*

PROOF OF (i). Our first object will be to obtain

$$\mathbf{P}(E_n^{(3)}) < \frac{c_{24}}{\varphi(n)}, \tag{3.8}$$

for a suitable $c_{24}>0$. Let $Q^{(3)}(n)$ be the number of integers r $(0 \leqq r \leqq 2n)$ such that $\Pi^{(3)}(n+f(n), \infty)$ returns to the point $S_3(r)$. By (2. 2) it is clear that

$$\mathcal{E}\{Q^{(3)}(n)\} \leqq c_{24} \sum_{r=0}^{2n} \frac{1}{[n-r+f(n)]^{1/2}}.$$

Thus

$$\mathcal{E}\{Q^{(3)}(n)\} < \frac{c_{25} n^{1/2}}{\varphi(n)}, \tag{3.9}$$

since $f(n)=n\{\varphi(n)\}^2$ and $\varphi(n) \to +\infty$ as $n \to \infty$.

As in the proof of Lemma 9 (i) we can estimate $\mathcal{E}\{Q^{(3)}(n)\}$ by another method. If $E_n^{(3)}$ occurs, this means that there exist integers r_1, r_2 with $0 \leqq r_1 \leqq n$ and $r_2 > n+f(n)$ such that $S_3(r_1)=S_3(r_2)$. Now think of $\Pi^{(3)}(r_1, 2n)$ and

[1] This condition is not really necessary for the truth of the theorem. It is inserted because it simplifies the proof slightly.

$\Pi^{(3)}(r_2, \infty)$ as two independent random walks of length $\geqq n$ starting from the same point. Since $r_2 > n\{\varphi(n)\}^2$ and $\varphi(n) \to \infty$, the knowledge that $S_3(r_2) = S_3(r_1)$ will have no appreciable effect on the behaviour of $\Pi^{(3)}(r_1, 2n)$.[2] Hence by Lemma 8 (i) the conditional expectation of $Q^{(3)}(n)$ given $E_n^{(3)}$ satisfies

$$\mathcal{E}\{Q^{(3)}(n)/E_n^{(3)}\} > c_{26} n^{1/2}.$$

Hence

$$\mathcal{E}\{Q^{(3)}(n)\} \geqq \mathbf{P}(E_n^{(3)})\mathcal{E}\{Q^{(3)}(n)/E_n^{(3)}\} > c_{26} n^{1/2} \mathbf{P}(E_n^{(3)}).$$

Using (3. 9), this immediately establishes (3. 8).

For $k = 2, 3, \ldots$, let $R_k^{(3)}$ be the event that $\Pi^{(3)}(0, 2^k)$ and $\Pi^{(3)}(2^k + 2^{k-1}\{\varphi(2^{k-1})\}^2, \infty)$ have a point in common. By (3. 8)

$$\mathbf{P}(R_k^{(3)}) < \frac{2c_{24}}{\varphi(2^{k-1})} \qquad (k = 2, 3, \ldots). \tag{3.10}$$

Hence, if $\sum \frac{1}{\varphi(2^k)}$ converges, then $\sum \mathbf{P}(R_k^{(3)})$ converges and, by Borel—Cantelli,

$$\mathbf{P}\{R_k^{(3)} \ i.o.\} = 0.$$

But from the definition of $R_k^{(3)}$, the non-occurrence of $R_k^{(3)}$ implies that no $E_n^{(3)}$ occurs for $n_{k-1} \leqq n \leqq n_k$. Hence

$$\mathbf{P}\{E_n^{(3)} \ i.o.\} = 0.$$

Conversely, suppose $\sum \frac{1}{\varphi(2^k)}$ diverges. We have independence difficulties in applying Borel—Cantelli this time: instead we prove that there exists $\eta > 0$ such that

$$\mathbf{P}\{E_n^{(3)} \ i.o.\} > \eta. \tag{3.11}$$

By the law of 0 or 1 this implies $\mathbf{P}\{E_n^{(3)} \ i.o.\} = 1$, and so completes the proof of (i).

Let $F_k^{(3)}$ $(k = 2, 3, \ldots)$ be the event that $\Pi^{(3)}\left(\frac{1}{2}n, n\right)$ and $\Pi^{(3)}(n + f(n), n + 5f(n))$ have a point in common for $n = 2^k$. Let $H_k^{(3)}$ $(k = 2, 3, \ldots)$ be the event that $\Pi^{(3)}(n + f(n), n + 4f(n))$ returns within $\left(\frac{1}{2}n\right)^{1/2}$ of $S_3\left(\frac{1}{2}n\right)$

[2] Strictly speaking the conditional probability distribution for $\{S_3(r) - S_3(r_1)\}$ $(r_1 \leqq r \leqq 2n)$ given $S_3(r_2) = S_3(r_1)$, should be obtained. It can be shown by simple but laborious computation that the conditional probability differs from the 'free' probability by a factor which $\to 1$ as $n \to \infty$; hence it is clearly justifiable to use Lemma 8 (i) with suitable different values for the constants.

for $n=2^k$. By Lemma 4,

$$\mathbf{P}(H_k^{(3)}) > \frac{\frac{1}{10}e_3}{\varphi(2^k)}, \tag{3.12}$$

and after having returned within $\left(\frac{1}{2}n\right)^{1/2}$ of $S_3\left(\frac{1}{2}n\right)$ there are still at least $f(n)>n$ steps to complete $\Pi^{(3)}(n+f(n), n+5f(n))$. Again, since $\varphi(n)\to\infty$, the departure from the situation of two independent random walks each of length $\geqq n$ starting at a distance apart $\sqrt{n}$ becomes negligable for large n. Hence by Lemma 6 (i),

$$\mathbf{P}(F_k^{(3)}/H_k^{(3)}) > c_{27},$$

so that, by (3.12),

$$\mathbf{P}(F_k^{(3)}) > \frac{c_{28}}{\varphi(2^k)}. \tag{3.13}$$

Hence $\sum_{k=2}^{\infty}\mathbf{P}(F_k^{(3)})$ diverges. Since $F_k^{(3)}\subset R_k^{(3)}$, it follows from (3.10) and the fact that $\varphi(n)\to\infty$ that $\mathbf{P}(F_k^{(3)})\to 0$ as $k\to\infty$. Hence for any $\eta>0$ there is a $K=K(\eta)$ such that given $k_1\geqq K$ one can find a $k_2\geqq k_1$ such that

$$2\eta < \sum_{k_1\leqq 4r\leqq k_2}\mathbf{P}(F_{4r}^{(3)}) < 3\eta. \tag{3.14}$$

We shall show that, provided η is chosen small enough, (3.14) will imply

$$\mathbf{P}\{\bigcup_{k_1\leqq k\leqq k_2} F_k^{(3)}\} > \eta \tag{3.15}$$

which clearly implies (3.11).

Suppose now that $F_k^{(3)}$ has happened. This implies that there are integers r_3, r_4 such that $2^{k-1}\leqq r_3\leqq 2^k$, $2^k+2^k\{\varphi(2^k)\}^2\leqq r_4\leqq 2^k+5\cdot 2^k\{\varphi(2^k)\}^2$ and $S_3(r_3)=S_3(r_4)$. In considering whether or not $F_{k+4}^{(3)}$ happens, there are at least

$$(16-5)2^k\{\varphi(2^{k+4})\}^2 > \frac{1}{2}f(2^{k+4})$$

steps after r_4 before we look at $\Pi^{(3)}(n+f(n), n+5f(n))$ for $n=2^{k+4}$. The method used earlier for estimating $R_k^{(3)}$ shows immediately that

$$\mathbf{P}\{F_{k+4}^{(3)}/F_k\} < \frac{c_{29}}{\varphi(2^{k+4})},$$

so that, by (3.13), it follows that

$$\mathbf{P}\{F_{k+4}^{(3)}\cap F_k^{(3)}\} < c_{30}\mathbf{P}(F_k)\mathbf{P}(F_{k+4}).$$

The same argument shows that for an integer $l \geqq 4$,

$$\mathbf{P}(F_{k+l} \cap F_k) < c_{30} \mathbf{P}(F_k) \mathbf{P}(F_{k+l}). \tag{3.16}$$

If (3.14) is satisfied, (3.16) implies that

$$\mathbf{P}\{F_{4r} \cap \bigcup_{4r < 4t \leqq k_2} F_{4t}\} < 3\eta c_{30} \mathbf{P}(F_{4r}),$$

so that if η is chosen small enough to ensure that $3\eta c_{30} < \frac{1}{2}$, it follows that for $k_1 \leqq 4r \leqq k_2$

$$\mathbf{P}\{F_{4r} - F_{4r} \cap \bigcup_{4r < 4t \leqq k_2} F_{4t}\} > \frac{1}{2} \mathbf{P}(F_{4r}).$$

Using (3.14), this clearly establishes (3.15) as required.

PROOF OF (ii). A very similar argument will work. Instead of (3.10), (3.13) one obtains

$$\mathbf{P}(R_k^{(4)}) < \frac{c_{31}}{(k-1)\psi(2^{k-1})}, \quad \mathbf{P}(F_k^{(4)}) > \frac{c_{32}}{k\varphi(2^k)}.$$

PROOF OF (iii). Define a sequence of integers by: n_1 is some fixed integer such that $f(n_1) \geqq 4$,

$$n_{k+1} = n_k + \left[\frac{1}{2} f(n_k)\right] \quad (k = 1, 2, \ldots). \tag{3.17}$$

Under the conditions given it is clear that $\sum_{n=1}^{\infty} \frac{1}{f(n)^{(d-2)/2}}$ converges or diverges with the series

$$\sum_{k=1}^{\infty} \frac{1}{f(n_k)^{(d-4)/2}}. \tag{3.18}$$

Suppose first that (3.18) converges. For $d \geqq 5$, let $Q^{(d)}(k)$ be the number of points common to $\Pi^{(d)}(0, n_k)$ and $\Pi^{(d)}\left(n_k + \left[\frac{1}{2} f(n_{k-1})\right], \infty\right)$. Then the argument used in (i) shows that

$$\mathcal{E}\{Q^{(d)}(k)\} < \frac{c_{33}}{\{f(n_{k-1})\}^{(d-4)/2}}.$$

Hence

$$\mathbf{P}\{Q^{(d)}(k) \geqq 1\} < \frac{c_{33}}{\{f(n_{k-1})\}^{(d-4)/2}}.$$

By Borel—Cantelli, there is probability 1 that $Q^{(d)}(k) = 0$ except for finitely many values of k. Further, if $Q^{(d)}(k)=0$, then $E_n^{(d)}$ does not happen for $n_{k-1} \leqq n \leqq n_k$. Thus

$$\mathbf{P}\{E_n^{(d)} \ i.\,o.\}=0.$$

Conversely suppose that (3.18) diverges. Let $F_k^{(d)}$ be the event that there is a point common to $\Pi^{(d)}(n_k-f(n_k), n_k)$ and $\Pi^{(d)}(n_k+f(n_k), n_k+ +2f(n_k))$. By (3.17), under the conditions satisfied by $f(n)$, the events $F_k^{(d)}$ and $F_{k+l}^{(d)}$ are completely independent when $l \geqq c_{34}$ where c_{34} is a suitable integer. Hence it is sufficient to prove that $\sum_{r=1}^{\infty} \mathbf{P}(F_{rc_{34}}^{(d)})$ diverges and apply Borel—Cantelli.

Now if $H_k^{(d)}$ is the event

$$|S_d(n_k)-S_d(n_k+f(n_k))|<(f(n_k))^{1/2},$$

then an easy computation shows that $\mathbf{P}(H_k^{(d)})>c_{35}^{(d)}$. Further, by Lemma 6 (iii),

$$\mathbf{P}(F_k^{(d)}/H_k^{(d)})>\frac{c_6^{(d)}}{\{f(n_k)\}^{(d-4)/2}},$$

so that

$$\mathbf{P}(F_k^{(d)})>\frac{c_{36}^{(d)}}{\{f(n_k)\}^{(d-4)/2}}.$$

Since $f(n)$ is increasing with n and (3.18) diverges, it follows that $\sum_{r=1}^{\infty} \mathbf{P}(F_{rc_{34}}^{(d)})$ also diverges. This completes the proof of the theorem.

REMARK. The result of Theorem 1 (ii) implies that there are infinitely many values of n such that $\Pi^4(0, n)$ and $\Pi^4(2n, \infty)$ have points in common. A somewhat simplified version of the same proof is sufficient to prove the following interesting result:

THEOREM 2. (i) *Two independent infinite random walks in* 4*-space which start from any two given fixed points have infinitely many common points with probability* 1; *whereas two random walk paths in d-space* ($d \geqq 5$) *meet only finitely often, with probability* 1.

(ii) *Three independent infinite random walks in* 3*-space which start from any* 3 *given fixed points have infinitely many points in common with probability* 1; *whereas three random walk paths in d-space* ($d \geqq 4$) *have only finitely many common points, with probability* 1.

The relationship between these results and those for Brownian motion was discussed in the introduction.

4. Points where the past and future do not intersect. We first need to estimate the probability that two random walk paths starting from the origin have no points (other than their starting point) in common.

For $d \geqq 3$, let $\tau_d(n)$ be the probability that two independent paths of n steps in d-space starting from O have no intersection. Let $\mu_d(n)$ be the corresponding probability for two paths, one of n steps, the other of infinite length. Clearly,

$$\begin{aligned} &\tau_d(1) \geqq \tau_d(2) \geqq \cdots \geqq \tau_d(n) \geqq \ldots, \\ &\mu_d(1) \geqq \mu_d(2) \geqq \cdots \geqq \mu_d(n) \geqq \ldots \end{aligned} \tag{4.1}$$

and

$$\tau_d(n) \geqq \mu_d(n) \qquad (n = 1, 2, \ldots). \tag{4.2}$$

Let $S_d(1), S_d(2), \ldots, S_d(n)$ be the sequence of points in path Π_1, and $S_d^1(1), S_d^1(2), \ldots$ the points of the path Π_2. Put

$$\varrho_d(r) = |S_d(r)|.$$

Then we can enumerate the possibilities that $S_d(r)$ is entered by Π_2, but later points of Π_1 are not entered. This gives

$$\tau_d(n) \sum_{r=0}^{n} p_r \leqq 1 \tag{4.3}$$

where p_r is the probability that Π_2 enters $S_d(r)$ before the n^{th} step, since the probability of later non-intersection is greater than that for the non-intersection of two paths of n-steps.

For $d = 3$, by Lemma 2,

$$p_r > c_{37} \mathscr{E} \left\{ \frac{1}{1 + \varrho_3(r)} \right\} > \frac{c_{38}}{r^{1/2}}, \quad \text{by Lemma 6.}$$

Hence by (4.3) we have

$$\tau_3(n) \leqq \frac{c_{39}}{n^{1/2}}. \tag{4.4}$$

For $d = 4$, by Lemma 2,

$$p_r > c_{40} \mathscr{E} \left\{ \frac{1}{1 + \{\varrho_4(r)\}^2} \right\} > \frac{c_{41}}{r}, \quad \text{by Lemma 5.}$$

Hence by (4.3) we have

$$\tau_4(n) \leqq \frac{c_{42}}{\log n}. \tag{4.5}$$

To obtain estimates in the other direction, it is easier to use $\mu_d(n)$. Clearly, for $0 \leqq k_1 \leqq n$,

$$\mu_d(n-k_1) \sum_{k=1}^{k_1} q_k + \sum_{k=k_1+1}^{n} q_k \geqq 1, \tag{4.6}$$

where q_k is the probability that Π_2 enters $S_d(k)$ at least once.

For $d=3$, the method does not give a good lower bound for $\mu_3(n)$. More complicated computations show that $\mu_3(n) \geqq c_{43} n^{1/2}$, but we do not prove this as it is not required in the sequel.

For $d=4$, take $k_1 = n - \left[\frac{n}{\log n}\right]$. Then, if $S_4(k)=0$, Π_2 certainly enters $S_4(k)$ at least once. If $S_4(k) \neq 0$, then by Lemma 1, the probability that Π_2 enters $S_4(k)$ is at most $c_{44}/\{\varrho_4(k)\}^2$. Since $\mathbf{P}\{S_4(k)=0\} < c_{45}/k^2$, it follows that $q_k < \frac{c_{45}}{k^2} + 2c_{44}\mathcal{E}\left\{\frac{1}{1+\{\varrho_4(k)\}^2}\right\}$. Using Lemma 5 this gives $q_k < \frac{c_{46}}{k}$. Substituting this in (4.6) gives immediately

$$\mu_4\left(\left[\frac{n}{\log n}\right]\right) c_{46} \sum_{k=1}^{n} \frac{1}{k} + O\left(\frac{1}{\log n}\right) \geqq 1$$

which implies immediately

$$\mu_4(n) \geqq \frac{c_{47}}{\log n}. \tag{4.7}$$

For $d \geqq 5$, a similar argument shows that $\sum_{k=1}^{\infty} q_k$ is a convergent series. Hence, if $k_1 = \left[\frac{1}{2} n\right]$,

$$\mu_d\left(\left[\frac{1}{2} n\right]\right) c_{48} + o(1) \geqq 1$$

which implies that

$$\mu_d(n) \geqq c_{49}^{(d)}. \tag{4.8}$$

By (4.1), (4.2) it follows that for $d \geqq 5$, there are constants $\tau_d \geqq \mu_d > 0$ such that

$$\tau_d(n) = \tau_d(1+o(1)). \tag{4.9}$$

In fact, the results of Section 3 imply that $\tau_d = \mu_d$ for $d \geqq 5$, and further that the probability of the non-intersection of two infinite random walks starting from the origin is also τ_d.

THEOREM 3. *For $d \geqq 3$, let $G^{(d)}(n)$ be the number of integers r $(1 \leqq r \leqq n)$ for which $\Pi^{(d)}(0, r)$ and $\Pi^{d}(r+1, \infty)$ have no points in common, then*

(i) *for any $\varepsilon > 0$,*

$$\mathbf{P}\{G^{(3)}(n) > n^{1/2+\varepsilon} \quad i.\,o.\} = 0;$$

(ii) $$\mathbf{P}\left\{0 = \liminf_{n\to\infty} \frac{G^{(4)}(n)\log n}{n} \leqq \limsup_{n\to\infty} \frac{G^{(4)}(n)\log n}{n} \leqq c_{50}\right\} = 1;$$

(iii) *for $d \geqq 5$,*

$$\mathbf{P}\left\{\lim_{n\to\infty} \frac{G^{(d)}(n)}{n} = \tau_d\right\} = 1.$$

PROOF OF (i). By (4.3), we have $\mu_3(n) \leqq c_{39} n^{-1/2}$, so that

$$\mathcal{E}\{G^{(3)}(n)\} \leqq c_{39} \sum_{r=1}^{n} r^{-1/2} < c_{51} n^{1/2}.$$

Hence

$$\mathbf{P}\left\{G^{(3)}(n) > \frac{1}{2} n^{1/2+\varepsilon}\right\} < \frac{c_{52}}{n^{\varepsilon}},$$

so that if $n_k = [k^{2/\varepsilon}]$, the series

$$\sum \mathbf{P}\left\{G^{(3)}(n_k) > \frac{1}{2} n_k^{1/2+\varepsilon}\right\}$$

converges and there is probability 1 that there exists k_0 such that $G^{(3)}(n_k) \leqq \frac{1}{2} n_k^{1/2+\varepsilon}$ for $k \geqq k_0$. Since $G^{(3)}(n)$ increases with n, and $\frac{n_{k+1}}{n_k} \to 1$ as $k \to \infty$, this implies that for $n_k \leqq n \leqq n_{k+1}$, $k \geqq k_0$, $G^{(3)}(n) \leqq n^{1/2+\varepsilon}$.

PROOF OF (ii). Let $T^{(4)}(r)$ be the event that $\Pi^{(4)}\left(r - \left[\frac{r}{(\log r)^4}\right], r\right)$ and $\Pi^{(4)}\left(r+1, r + \left[\frac{r}{(\log r)^4}\right]\right)$ have no point in common and let $H^{(4)}(n)$ be the number of events $T^{(4)}(r)$ which occur for $1 \leqq r \leqq n$. Clearly,

$$H^{(4)}(n) \geqq G^{(4)}(n). \tag{4.10}$$

However, by (4.5) and (4.7),

$$\frac{c_{53}}{\log r} \leqq \mathbf{P}\{T^{(4)}(r)\} \leqq \frac{c_{54}}{\log r},$$

so that

$$c_{53} \frac{n}{\log n} \leqq \mathcal{E}\{H^{(4)}(n)\} \leqq c_{55} \frac{n}{\log n}. \tag{4.11}$$

Now the events $T^{(4)}(k)$ and $T^{(4)}(r)$ are completely independent whenever

$k > r + 4\dfrac{n}{(\log n)^4}$, so that the variance $\sigma^2\{H^{(4)}(r)\}$ of $H^{(4)}(n)$ satisfies

$$\sigma^2\{H^{(4)}(n)\} \leqq c_{56}\frac{n^2}{(\log n)^4}. \tag{4.12}$$

By Chebyshev's inequality, using (4.11) and (4.12), we have

$$\mathbf{P}\left\{|H^{(4)}(n) - \mathcal{E}\{H^{(4)}(n)\}| > \frac{\varepsilon n}{\log n}\right\} < \frac{c_{57}}{\varepsilon^2}\frac{1}{(\log n)^2}. \tag{4.13}$$

For the sequence $m_k = [e^{k/\log k}]$ an application of the Borel—Cantelli lemma shows that for every $\varepsilon > 0$

$$|H^{(4)}(m_k) - \mathcal{E}\{H^{(4)}(m_k)\}| \leqq \frac{\varepsilon m_k}{\log m_k}$$

except for finitely many integers k with probability 1. Since $\dfrac{m_{k+1}}{m_k} \to 1$ as $k \to \infty$, this implies that

$$\lim_{n\to\infty} \frac{\log n}{n}|H^{(4)}(n) - \mathcal{E}\{H^{(4)}(n)\}| = 0$$

with probability 1. By (4.11), (4.10) it follows immediately that

$$\mathbf{P}\left\{\limsup_{n\to\infty} \frac{\log n}{n} H^{(4)}(n) \leqq c_{55}\right\} = 1,$$

which by (4.10) implies the right side inequality of (ii).

By Theorem 1 (ii), there are, with probability 1, infinitely many integers n for which

$$\Pi^{(4)}(0, n) \quad \text{and} \quad \Pi^{(4)}(n + n[\log\log n], \infty)$$

have a point in common. For such values of n there are no points of non-intersection between n and $n + n[\log\log n]$, so that $G^{(4)}(n) = G^{(4)}(n[\log\log n])$. Thus it follows immediately from the right side inequality of (ii) that

$$\mathbf{P}\left\{\liminf_{n\to\infty} \frac{\log n}{n} G^{(4)}(n) = 0\right\} = 1.$$

PROOF OF (iii). For $d \geqq 5$, let $T^{(d)}(r)$ be the event that $\Pi^{(d)}(r - [r^{3/4}], r)$ and $\Pi^{(d)}(r + 1, r + [r^{3/4}])$ have no points in common and let $H^{(d)}(r)$ be the number of events $T^{(d)}(r)$ $(1 \leqq r \leqq n)$ which occur. By (4.9),

$$\mathbf{P}\{T^{(d)}(r)\} = \tau_d(1 + o(1)),$$

so that

$$\mathcal{E}\{H^{(d)}(r)\} = n\tau_d(1 + o(1)). \tag{4.14}$$

If $k > r + 4n^{3/4}$, the events $T^{(d)}(r)$ and $T^{(d)}(k)$ are completely independent. Hence

$$\sigma^2\{H^{(d)}(n)\} < c_{58} n^{7/4}. \tag{4.15}$$

The usual Chebyshev estimate using the sequence $r_k = k^{10}$ $(k = 1, 2, \ldots)$ now establishes using (4. 14), (4. 15) that

$$\mathbf{P}\left\{\lim_{n\to\infty} \frac{H^{(d)}(n)}{n} = \tau_d\right\} = 1. \tag{4. 16}$$

However, from Theorem 1 (iii) there is, with probability 1, an integer n_0 such that for $n \geqq n_0$ (a) $\Pi^{(d)}(0, n-[n^{3/4}])$ and $\Pi^{(d)}(n, \infty)$ do not intersect and (b) $\Pi^{(d)}(0, n)$ and $\Pi^{(d)}(n+[n^{3/4}], \infty)$ do not intersect. For $n \geqq n_0$ the difference $G^{(d)}(n) - H^{(d)}(n)$ remains fixed, so that

$$\mathbf{P}\left\{\lim_{n\to\infty} \frac{G^{(d)}(n) - H^{(d)}(n)}{n} = 0\right\} = 1.$$

This result, together with (4. 16), establishes

$$\mathbf{P}\left\{\lim_{n\to\infty} \frac{G^{(d)}(n)}{n} = \tau_d\right\} = 1,$$

and completes the proof of the theorem.

REMARK 1. The result of Theorem 3 (i) is clearly not best-possible: even the methods we have used can give a better upper bound for the density $\frac{1}{n} G^{(3)}(n)$. We expect that $\frac{1}{n^{1/2}} G^{(3)}(n)$ is bounded but did not succeed in proving this.

REMARK 2. An obvious problem arises out of the result of Theorem 2 (ii). Is $\limsup_{n\to\infty} \frac{\log n}{n} G^{(4)}(n)$ positive or zero?

REMARK 3. It is clear that τ_d, the density of good points, increases with d. In fact, $\tau_d \to 1$ as $d \to \infty$.

(Received 16 November 1959)

References

[1] A. DVORETZKY and P. ERDŐS, Some problems on random walk in space, *Proc. Second Berkeley Symposium*, (1950), pp. 353—368.

[2] A. DVORETZKY, P. ERDŐS and S. KAKUTANI, Double points of Brownian paths in n-space, *Acta Sci. Math. Szeged*, **12** (1950), pp. 75—81.

[3] A. DVORETZKY, P. ERDŐS, S. KAKUTANI and S. J. TAYLOR, Triple points of Brownian paths in 3-space, *Proc. Cambr. Phil. Soc.*, **53** (1957), pp. 856—862.

[4] P. ERDŐS and S. J. TAYLOR, Some problems concerning the structure of random walk paths, *Acta Math. Acad. Sci. Hung.*, **11** (1960), pp. 137—162.

[5] S. J. TAYLOR, The Hausdorff α-dimensional measure of Brownian paths in n-space, *Proc. Cambr. Phil. Soc.*, **49** (1953), pp. 31—39.