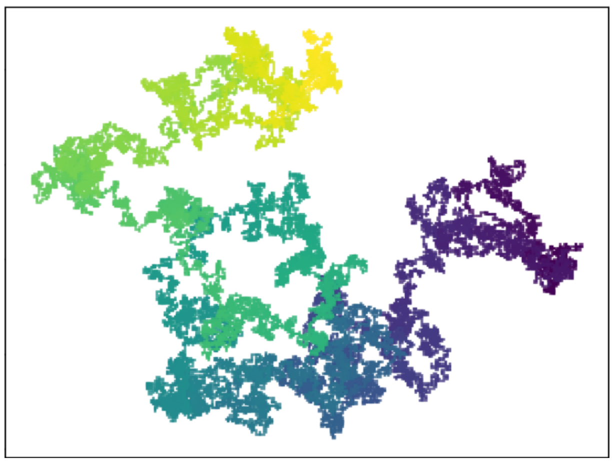
Research Project: Stochastic Crossroads, Navigating Random Walk Intersections

Romea Kelvin



Abstract

This thesis explores some fundamental results on random walks on $\mathbb{Z}^d$, focusing on the intersection of random walks. We present classical results on the recurrence and transience of random walks, as well as the famous Erdős-Taylor theorem, which we will illustrate and then prove. Finally, we will address the computation of intersection coefficients of random walks to conclude this thesis.

1 Introduction

Random walks on $\mathbb{Z}^d$ are fundamental stochastic processes in probability and stochastic processes theory. They model the random movement of a particle in a lattice of dimension d. A random walk is defined by a sequence of i.i.d. random variables X_n taking values in $\mathbb{Z}^d$. At each step, the particle moves in one of the possible directions with equal probability.

2 Theoretical Foundations

In this section, we introduce the basic concepts concerning random walks on $\mathbb{Z}^d$. Let $(X_n)_{n\geq 0}$ be a sequence of independent identically distributed random variables such that $E(X_n) = 0$ and the sequence $(X_n)_{n\geq 0}$ is square integrable with values in $\mathbb{Z}^d$, thus defining a random walk. At each time step n, the particle moves according to the random vector X_n. The position of the particle after n steps is given by the sum of the first n random vectors:

$$S_n = X_1 + X_2 + \cdots + X_n.$$

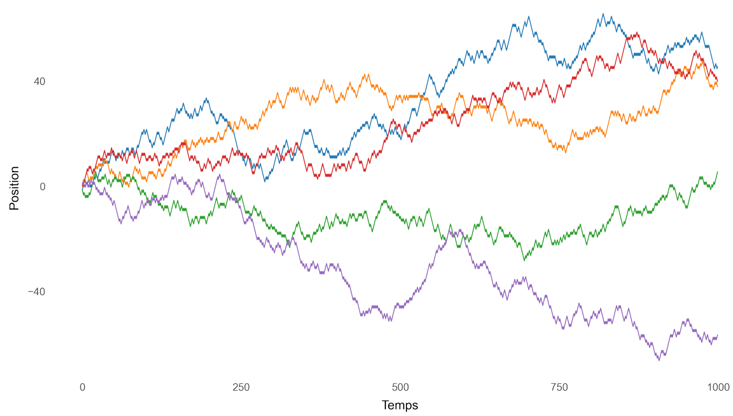


Figure 1: Random Walk $d = 1$

Some Properties of Random Walks:

Property 1: Aperiodicity

A random walk is said to be aperiodic if the Markov chain (S) is irreducible and aperiodic on $\mathbb{Z}^d$.

Anti-concentration Theorem

If $(S_n)_{n>0}$ is a random walk of dimension d and aperiodic, then there exists a constant $C > 0$ such that for all $n > 1$, we have:

$$\sup_{x\in\mathbb{Z}} P(S_n = x) \leq Cn^{-\frac{d}{2}}.$$

Proof:

Noting $\hat{\mu}(\xi) = E[e^{i\xi X_1}]$ and applying a Fourier transformation to $P(S_n = x)$, we obtain:

$$\begin{aligned}
P(S_n = x) &= \sum_{k\in\mathbb{Z}^d} \frac{1}{(2\pi)^d} \int_{(-\pi,\pi)^d} e^{-i\xi\cdot x} e^{i\xi\cdot k} P(S_n = k) d\xi \\
&= \frac{1}{(2\pi)^d} \int_{(-\pi,\pi)^d} e^{-i\xi\cdot x} E[e^{i\xi\cdot S_n}] d\xi \quad \text{by transferring the expectation} \\
&= \frac{1}{(2\pi)^d} \int_{(-\pi,\pi)^d} e^{-i\xi\cdot x} (\hat{\mu}(\xi))^n d\xi,
\end{aligned}$$

since $E[e^{it\cdot S_n}] = E[e^{it\cdot X_1} \dots e^{it\cdot X_n}] = E[e^{it\cdot X_1}] \dots E[e^{it\cdot X_n}] = (\hat{\mu}(t))^n$. Indeed, the X_i are independent and the function $x \mapsto \exp(-ixt)$ is continuous, so by coalition the $e^{it\cdot X_i}$ are also independent.
Noticing that $|e^{-i\xi\cdot x}| \leq 1$, we obtain:

$$P(S_n = x) = \tfrac{1}{(2\pi)^d} \int_{(-\pi,\pi)^d} e^{-i\xi\cdot x} (\hat{\mu}(\xi))^n d\xi \leq \tfrac{1}{(2\pi)^d} \int_{(-\pi,\pi)^d} |\hat{\mu}(\xi)|^n d\xi.$$

For the rest of the proof, we assume that in the case where μ is aperiodic, we have:

$$|\hat{\mu}(\xi)| \leq 1 - \lambda|\xi|^2, \quad \forall \xi \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]^d.$$

We obtain,

$$\int_{(-\frac{\pi}{2};\frac{\pi}{2})^d} |\hat{\mu}(\xi)|^n d\xi \leq \int_{(-\frac{\pi}{2};\frac{\pi}{2})^d} \left(1 - \lambda|\xi|^2\right)^n d\xi.$$

With the change of variable $\sqrt{\lambda}\xi = z/\sqrt{n}$, which implies $\xi = \frac{z}{\sqrt{\lambda n}}$. Changing the differential volume element $d\xi$ in terms of dz. In d-dimensions, we get $d\xi = \left(\frac{1}{\sqrt{\lambda n}}\right)^d dz$.

$$\int_{(-\frac{\pi}{2};\frac{\pi}{2})^d} \left(1 - \lambda|\xi|^2\right)^n d\xi. = \frac{1}{(\sqrt{\lambda n})^d} \int_{\mathbb{R}^d} \left(1 - \frac{|z|^2}{n}\right)^n dz$$

Switching to polar coordinates, in d-dimensions, the differential volume element in polar coordinates is $r^{d-1} dr d\Omega$, where $d\Omega$ is the solid angle volume element in d-dimensional angles.

$$\begin{aligned}
\frac{1}{(\lambda n)^{\frac{d}{2}}} \int_{\mathbb{R}^d} \left(1 - \frac{|z|^2}{n}\right)^n dz &= \frac{1}{(\lambda n)^{\frac{d}{2}}} \int_0^\infty \int_{S^{d-1}} \left(1 - \frac{r^2}{n}\right)^n r^{d-1} dr d\Omega \\
&\leq \frac{C'}{n^{\frac{d}{2}}} \int_0^\infty r^{d-1} \left(1 - \frac{r^2}{n}\right)^n dr
\end{aligned}$$

Setting $C' = \frac{1}{\lambda^{\frac{d}{2}}}$, Moreover, by Taylor series, we have $\left(1 - \frac{r^2}{n}\right)^n \leq e^{-r^2}$ Thus, finally,

$$\frac{C'}{n^{\frac{d}{2}}} \int_0^\infty r^{d-1}(1 - \frac{r^2}{n})^n dr \leq \frac{C'}{n^{\frac{d}{2}}} \int_0^\infty r^{d-1} e^{-r^2} dr = \frac{C}{n^{\frac{d}{2}}}.$$

We have successfully proven that $\sup_{x \in \mathbb{Z}} P(S_n = x) \leq C n^{-\frac{d}{2}}$.

Property 2: Recurrence

If the random walk (S) is recurrent, this means that the probability of returning to the origin infinitely often is equal to 1, i.e:

$$P(|S_n| \to \infty) = 0$$

Recurrence Criterion: Chung-Fuchs Theorem

A random walk on $\mathbb{Z}^d$ is recurrent if and only if we have:

$$\lim_{r \to 1} \int_{-\pi}^{\pi} \operatorname{Re}\left(\frac{1}{1 - r\hat{\mu}(\xi)}\right) d\xi = \infty.$$

where $\hat{\mu}(\xi) = E[e^{i\xi \cdot X_1}]$

Proof for $d = 1$:

The random walk (S) is recurrent if and only if $\sum_{n>0} P(S_n = 0)$ diverges. Starting from the expansion of $P(S_n = x)$ from the proof of the Anti-concentration Theorem with $d = 1$, we have:

$$P(S_n = x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-it \cdot x} (\hat{\mu}(t))^n dt,$$

So for $P(S_n = 0)$ we have:

$$P(S_n = 0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} (\hat{\mu}(t))^n dt,$$

We will sum the $P(S_n = 0)$ and multiply by r^n with $r \in [0, 1]$ to ensure that we can interchange sum and integral. We obtain:

$$\sum_{n \geq 0} r^n P(S_n = 0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{n \geq 0} r^n (\hat{\mu}(t))^n dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{1 - r\hat{\mu}(t)} dt.$$

By geometric series formula, which applies well because $(r\hat{\mu}(t))^n \in [0,1]$. Since the interior of the integral is real, we can take its real part, and moreover by letting r tend to 1 we indeed obtain that the series $\sum_{n>0} P\left(S_n = 0\right) = \infty$ if and only if $\lim_{r\to 1} \int_{-\pi}^{\pi} \mathrm{Re}\left(\frac{1}{1-r\hat{\mu}(\xi)}\right) d\xi = \infty$. The proof in the general case is very similar.

Property 3: Transience

If the random walk is transient, this means that it diverges to infinity with probability 1. In other words:

$$P(|S_n| \to \infty) = 1$$

Transience of Aperiodic Random Walks

Let (S) be a random walk on $\mathbb{Z}^d$ aperiodic, then if $d \geq 3$, (S) is transient.

Proof:

For (S) to be transient, the average number of returns to 0 must be finite. That is, the series $\sum_{n\geq 1} P(S_n = 0)$ must converge. According to the Anti-concentration Theorem, we have:

$$\sum_{n\geq 1} P(S_n = 0) \leq \sum_{n\geq 1} Cn^{-\frac{d}{2}} = C \sum_{n\geq 1} n^{-\frac{d}{2}}.$$

So if $d \geq 3$, the series converges by the Riemann criterion. Thus, (S) is transient.

Property 4: The Trace:

The trace of a random walk up to time n is defined as the sequence of visited positions:

$$R_n = \{S_0, S_1, \ldots, S_n\}.$$

Its cardinality, R_n, represents the number of different positions visited up to time n.

Kesten–Spitzer–Whitman Theorem:

A fundamental result in the study of random walks is the Kesten–Spitzer–Whitman theorem which states that:

$$\frac{R_n}{n} \xrightarrow{p.s.} c$$

where $c \geq 0$ is the probability that the walk never returns to the origin after time 1.

Proof:

Let's start by estimating the first two moments of R_n.

$$E[R_n] = E\left[\sum_{i=0}^{n} 1_{S_j \neq S_i : \forall i < j \leq n}\right]$$

By linearity of expectation and knowing that $E[1_A] = P(A)$, we have:

$$\begin{aligned} E[R_n] &= \sum_{i=0}^{n} P(S_j \neq S_0 : \forall 0 < j \leq n - i) \\ &\sim n \cdot P(S_i \neq S_0 : \forall i > 1) \end{aligned}$$

By Cesàro summation. For the second moment, we start by expanding the square of R_n:

$$\begin{aligned} (R_n)^2 &= \sum_{1 \leq i,j \leq n} 1(S_i \in \{S_1, \ldots, S_{i-1}\}) \cdot 1(S_j \in \{S_1, \ldots, S_{j-1}\}) \\ &= 2 \sum_{1 \leq i < j \leq n} 1(S_i \in \{S_1, \ldots, S_{i-1}\}, X_j \in \{S_1, \ldots, S_{j-1}\}) + R_n. \end{aligned}$$

The expectation of each variable in the above sum can be bounded as follows:

$$\begin{aligned} &P[S_i \in \{S_1, \ldots, S_{i-1}\}, S_j \in \{S_1, \ldots, S_{j-1}\}] \\ &\leq P[S_i \in \{S_1, \ldots, S_{i-1}\}, S_j \in \{S_i, \ldots, S_{j-1}\}] \\ &\leq P[X_i \neq 0, X_i + X_{i-1} \neq 0, \ldots, X_i + \ldots + X_2 \neq 0, X_j \neq 0, X_j + X_{j-1} \neq 0, X_j + \ldots + X_{i+1} \neq 0] \\ &\leq P[X_i \neq 0, X_i + X_{i-1} \neq 0, \ldots, X_i + \ldots + X_2 \neq 0] \cdot P[X_j \neq 0, X_j + X_{j-1} \neq 0, X_j + \ldots + X_{i+1} \neq 0] \\ &\to \; P(S_i \neq S_0 : \forall i > 1)^2 = c^2 \text{ (by Markov)} \end{aligned}$$

Thus, returning to our decomposition of $(R_n)^2$ and by linearity of expectation, we have:

$$\begin{aligned} E\left[\left(\frac{R_n}{n}\right)^2\right] &\leq \frac{E[R_n]}{n^2} + \text{a Cesàro mean of terms close to } c^2. \\ &= O(n^{-1}) + \text{a Cesàro mean of terms close to } c^2. \end{aligned}$$

Therefore,

$$E\left[\left(\frac{R_n}{n}\right)^2\right] \leq \; c^2 + o(1).$$

By subtracting the square of the expectation, we thus obtain $\mathrm{Var}\left(\frac{R_n}{n}\right) = o(1)$, and the Bienaymé–Chebyshev inequality now gives

$$P\left(\left|\frac{R_n}{n} - E\left[\frac{R_n}{n}\right]\right| \geq \varepsilon\right) \leq \frac{\mathrm{Var}\left(\frac{R_n}{n}\right)}{\varepsilon^2} \to_{n\to\infty} 0$$

for any $\varepsilon > 0$. We conclude using the fact that $E\left[\frac{R_n}{n}\right] \to_{n\to\infty} c$.

3 Intersection of Random Walks

The question of the intersection of two independent random walks with the same distribution is a major topic in probability and stochastic processes theory.
Let's start by providing an intuitive result regarding the number of intersections of a random walk depending on the dimension. For this purpose, we conduct the following simulation:
For random walks in $\mathbb{Z}^d$, we aim to estimate the expectation $E[I_n]$, where I_n represents the number of intersections between the paths after n steps.
The following protocol is thus implemented:

1. Select a dimension d.
2. Perform two independent random walks, each with a maximum of 100 steps.
3. Calculate the number of intersections between the paths after n steps, for $n = 1, 2, \dots, 100$.
4. Repeat this procedure 1000 times.
5. Calculate the average of the 1000 intersection counts for each n, thereby providing an approximate estimation of $E[I_n]$ for $n = 1, 2, \dots, 100$.

The calculated values of $E[I_n]$ are then depicted in the corresponding figure.

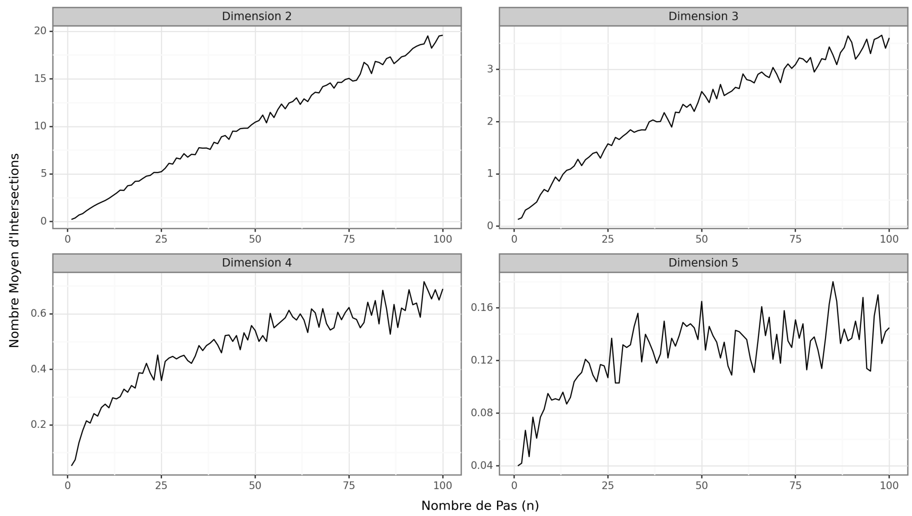


Figure 2: Nombre Moyen d'Intersections suivant la dimension

We observe that as the dimension of the random walks increases, the number of intersections grows at a slower rate. This trend becomes so pronounced that from the fifth dimension onwards, we can already conjecture that the number of intersections tends towards a constant as the number of steps increases. This observation was generalized by Erdős and Taylor, who demonstrated that up to a dimension equal to 4, the number of intersections is almost certainly infinite, while for all dimensions greater than 5, the number of intersections is almost certainly finite.

3.1 Erdős-Taylor Theorem

First, let's start by defining the function that gives the number of intersections of k random walks for dimension d.
We denote $I^{(d)}(i)$ as the number of points in the intersection of the paths of i independent random walks in $\mathbb{Z}^d$.

More precisely, suppose we have i independent random walks in $\mathbb{Z}^d$, each represented by a path in the space $\mathbb{Z}^d$. The number of points in the intersection of these paths is counted, i.e., the points that are common to all paths.

Formally, for $k > 1$ and $d > 1$, we define $I^{(d)}(k)$ as follows:

$$I^{(d)}(k) = \sum_{x\in\mathbb{Z}^d} \prod_{i=1}^{k} 1_{x\in S^{(d)}(i)}$$

where:

- $S^{(d)}(i)$ represents the path of the i-th random walk in $\mathbb{Z}^d$.
- $1_{x\in S^{(d)}(i)}$ is an indicator function that is 1 if point x is in the path of the i-th random walk, and 0 otherwise.

- $\sum_{x\in\mathbb{Z}^d}$ is a sum over all points in $\mathbb{Z}^d$.

In other words, $I^{(d)}(k)$ counts the number of points in the intersection of the paths of k random walks in $\mathbb{Z}^d$.

From there, we can establish the Erdős and Taylor Theorem:

Erdős & Taylor Theorem:

Let d be the dimension of the ambient space.

(i) If $d \leq 2$, then the probability that the number of intersections between the paths of i independent random walks is infinite is almost surely zero for all $i > 2$:

$$\mathbb{P}(I^{(d)}(i) = \infty) = 1 \quad \text{for all} \quad i > 2$$

(ii) If $d = 3$, then the probability that the number of intersections between the paths of three independent random walks is infinite is almost surely nonzero, but the probability that the number of intersections between the paths of four independent random walks is infinite is almost surely zero:

$$\mathbb{P}(I^{(3)}(3) = \infty) > 0 \quad \text{and} \quad \mathbb{P}(I^{(3)}(4) < \infty) = 1$$

(iii) If $d = 4$, then the probability that the number of intersections between the paths of two independent random walks is infinite is almost surely nonzero, but the probability that the number of intersections between the paths of three independent random walks is infinite is almost surely zero:

$$\mathbb{P}(I^{(4)}(2) = \infty) > 0 \quad \text{and} \quad \mathbb{P}(I^{(4)}(3) < \infty) = 1$$

(iv) If $d \geq 5$, then the probability that the number of intersections between the paths of two independent random walks is infinite is almost surely zero:

$$\mathbb{P}(I^{(d)}(2) < \infty) = 1$$

3.2 Proofs of the theorem

To prove this theorem, we first need to recall the Local Limit Theorem, which will allow us to establish the Green's function, our main tool for proving the Erdős & Taylor Theorem.

Local Limit Theorem:

Let S_n be a symmetric random walk in $\mathbb{Z}^d$. For a positive integer n and a point $x \in \mathbb{Z}^d$, according to the local limit theorem:

$$P\{S_n = x\} \approx 2\left(\frac{d}{2\pi n}\right)^{d/2} e^{-\frac{d|x|^2}{2n}}.$$

Green's Function:

We start by defining the generating function of Green as the power series in ξ:

Let $p \in P \cup P^*$ and $x, y \in \mathbb{Z}^d$, we have:

$$G(x, y; \xi) = \sum_{n=0}^{\infty} \xi^n \mathbb{P}^x\{S_n = y\} = \sum_{n=0}^{\infty} \xi^n p_n(y - x).$$

So, we can define the Green's function in the case of a transient random walk:

$$G(x, 0) = G(x) = \sum_{n=0}^{\infty} p_n(x).$$

where $p_n(x)$ is the probability that the random walk S visits x exactly at time n, and $G(x)$ is the expected number of such visits.

Now, we can introduce a very important lemma for the proof of the Erdős & Taylor theorem:

Lemma:

Let $x \in \mathbb{Z}^d$ for $d > 3$, there exist two constants $(c_1, c_2) \in \mathbb{R}^+$ such that for all $x \in \mathbb{Z}^d$, we have:

$$c_1|x|^{2-d} \leq G(x) \leq c_2|x|^{2-d}.$$

Proof of the Lemma:

We can start by observing that:

$$\begin{aligned} G(x) &= \sum_{n=0}^{\infty} P(S_n^{(d)} = x) \\ &= \sum_{n=0}^{|x|^2} P(S_n^{(d)} = x) + \sum_{n>|x|^2}^{\infty} f\left(\frac{|x|^2}{n}\right)\, n^{\frac{-d}{2}} + o\left(n^{\frac{-d}{2}}\right) \end{aligned}$$

where f is a bounded function, and according to the Local Limit Theorem, is approximately equal to e^{-x}. From this expression of $G(x)$, we can deduce a lower and upper bound.

For the lower bound:

According to the Local Limit Theorem, we have $f\left(\frac{|x|^2}{n}\right) \approx e^{\frac{|x|^2}{n}}$. Considering that the sum $\sum_{|x|^2<n<2|x|^2}^{\infty} f\left(\frac{|x|^2}{n}\right)\ n^{\frac{-d}{2}} + o\left(n^{\frac{-d}{2}}\right)$ gives $|x|^2$ terms of order $n^{\frac{-d}{2}} \approx |x|^{-d}$, this is much larger than $c_1|x|^{2-d}$.

$$\text{Hence: } c_1|x|^{2-d} \leq G(x)$$

For the upper bound:

For the upper bound, we need to bound each sum separately:
First sum $\sum_{n=0}^{|x|^2} P(S_n^{(d)} = x)$:

$$\sum_{n=0}^{|x|^2} P(S_n^{(d)} = x) \leq \int_0^{|x|^2} f\left(\frac{|x|^2}{t}\right)\ t^{\frac{-d}{2}}\ dt \leq \int_0^{|x|^2} e^{-\frac{|x|^2}{t}} \cdot t^{-d} dt$$

By applying the change of variable $u = \frac{|x|^2}{t}$, we get $dt = -\frac{|x|^2}{u^2} du$, thus:

$$\begin{aligned}\int_0^{|x|^2} e^{-\frac{|x|^2}{t}} \cdot t^{-d} dt &= \int_1^{+\infty} e^{-u} \cdot \left(\frac{|x|^2}{u}\right)^{\frac{-d}{2}} \cdot \frac{-|x|^2}{u^2} dt \\ &= |x|^{2-d} \int_1^{+\infty} e^{-u} \cdot u^{\frac{-d}{2}-2} dt\end{aligned}$$

Now $e^{-u} \cdot u^{\frac{-d}{2}-2} \xrightarrow[u \to +\infty]{} 0$ by comparison, hence this integral converges and can be bounded by some c, giving us our first bound:

$$\sum_{n=0}^{|x|^2} P(S_n^{(d)} = x) \leq c|x|^{2-d} \tag{1}$$

For the second sum $\sum_{n>|x|^2}^{\infty} f\left(\frac{|x|^2}{n}\right)\ n^{\frac{-d}{2}}$:

$$\sum_{n>|x|^2}^{\infty} f\left(\frac{|x|^2}{n}\right)\ n^{\frac{-d}{2}} \approx \sum_{n>|x|^2}^{\infty} e^{\frac{|x|^2}{n}}\ n^{\frac{-d}{2}} \leq \sum_{n>|x|^2}^{\infty} n^{\frac{-d}{2}}$$

Since $|e^{\frac{|x|^2}{n}}| \leq 1$, and as $d \geq 3$, the series converges by the Riemann criterion. Similarly to the first bound, we bound it by taking the integral (which also converges by Riemann, since $d \geq 3$), yielding:

$$\sum_{n>|x|^2}^{\infty} n^{\frac{-d}{2}} \leq \int_{|x|^2}^{\infty} t^{-\frac{d}{2}}\ dt = \left[\frac{2}{2-d} x^{1-\frac{2}{d}}\right]_{|x|^2}^{+\infty} = \frac{-2|x|^{2-d}}{2-d} \leq |x|^{2-d}$$

For the third sum $\sum_{n>|x|^2}^{\infty} o\left(n^{\frac{-d}{2}}\right)$:

$$\sum_{n>|x|^2}^{\infty} o\left(n^{\frac{-d}{2}}\right) \leq \sum_{n>|x|^2}^{\infty} n^{\frac{-d}{2}} \leq |x|^{2-d}$$

Using what was done above. By combining our three bounds, we obtain that there exists $c_2 > 0$ such that:

$$G(x) \geq c_2|x|^{2-d}.$$

In conclusion, we have proved that $c_1|x|^{2-d} \leq G(x) \leq c_2|x|^{2-d}$ for all $x \in \mathbb{Z}^d$, where c_1 and c_2 are positive constants.

Proof of Erdős and Taylor Theorem:

Let's start with the simplest cases.
In the case where $d \leq 2$, the random walk $S^{(d)}$ is recurrent, so $S^{(d)}(i)$ is the entire $\mathbb{Z}^d$.

First, let's examine the first moment of $I^{(d)}(k)$:

$$\begin{aligned}
E(I^{(d)}(k)) &= \sum_{x\in\mathbb{Z}^d} (q(x))^k \\
&\approx \sum_{x\in\mathbb{Z}^d} (G(x))^k && \text{(Green's estimator)} \\
&\approx \sum_{x\in\mathbb{Z}^d} \left(|x|^{2-d}\right)^k && \text{(by bounding the Green function)} \\
&\approx \sum_{n>1} n^{k(2-d)} n^{d-1}
\end{aligned}$$

The series is finite in the following cases: $d \geq 5$ and $k = 2$, $d \geq 4$ and $k = 3$, $d \geq 3$ and $k = 4$. However, we see that for $d = 3$ and $k = 3$ as well as $d = 4$ and $k = 2$ the series diverges, which does not necessarily mean that $I^{(d)}(k)$ is not finite.

We will use the method of the second moment to deal with these two cases and notably the fact that if $X > 0$ is a positive random variable and if $C > 0$ is such that $E[X]^2 \leq E[X^2] \leq CE[X]^2$, then by Cauchy-Schwarz:

$$\begin{aligned}
E[X] &= E[X1_{X>E[X]/2}] + E[X1_{X<E[X]/2}] \\
&\leq E[X^2]^{1/2} P(X > \frac{E[X]}{2})^{1/2} + \frac{E[X]}{2}. \\
&\leq C^{1/2} E[X] P(X > \frac{E[X]}{2})^{1/2} + \frac{E[X]}{2}. \quad (\textit{because} \leq E[X^2] \leq CE[X]^2)
\end{aligned}$$

So,

$$\frac{E[X]}{2} \leq C^{1/2} E[X] P(X > \frac{E[X]}{2})^{1/2}$$

Hence

$$P(X > \frac{E[X]}{2})^{1/2} \geq \frac{1}{2C^{\frac{1}{2}}}$$

And finally,

$$P(X > \frac{E[X]}{2}) \geq \frac{1}{4C}$$

So we will apply this not to X but to:

$$I_n^{(d)}(k) = \sum_{x\in\mathbb{Z}^d} \prod_{j=1}^k 1_{\{S_i^{(d)}(j)=x, \text{ for certain } i\leq n\}}.$$

Noting $E[I_n^{(d)}(k)] = \sum_{x\in\mathbb{Z}^d} q_n(x)^k$ and recalling that $(a+b)^k \leq 2^k(a^k+b^k)$, we have:

$$\begin{aligned}
E\left[I_n^{(d)}(k)^2\right] &= \sum_{x,y\in\mathbb{Z}^d} P(x \text{ and } y \text{ visited by all } S_i^{(d)} \text{ at time } n) \\
&= \sum_{x,y\in\mathbb{Z}^d} P(x \text{ and } y \text{ visited by all } S_1^{(d)} \text{ at time } n)^k \\
&\leq \sum_{x,y\in\mathbb{Z}^d} (q_n(x)q_n(y-x) + q_n(y)q_n(x-y))^k \\
&\leq 2^{k+1} \sum_{x,y\in\mathbb{Z}^d} q_n(x)^k q_n(y-x)^k \\
&= 2^{k+1} \left(\sum_{x\in\mathbb{Z}^d} q_n(x)^k \right)^2 \\
&= 2^{k+1} E\left[I_n^{(d)}(k)\right]^2.
\end{aligned}$$

So, using the second moment method, we have that for all n:

$$P\left(I_n^{(d)}(k) > \frac{E[I_n^{(d)}(k)]}{2}\right) > \frac{1}{4\cdot 2^k}$$

And thus,

$$P\left(I_n^{(d)}(k) = \infty\right) > \frac{1}{2^{k+2}}$$

We conclude using Hewitt-Savage 0-1 law which ensures that this probability is either 0 or 1, and since it cannot be 0, it must be 1.

4 Calculation of Intersection Coefficients

Starting from the simulation conducted in section 3, we may wonder if it is possible to find an approximation of the number of intersections between two random walks according to the dimension. For this purpose, we will perform the same simulation as before but this time we will divide the average number of intersections by its possible approximation. Moreover, we will extend the number of steps from 100 to 1000 in order to better observe the convergences.

1. For $d = 1$ we take $n^{\frac{3}{2}}$
2. For $d = 2$ we take n
3. For $d = 3$ we take $\sqrt{n}$
4. For $d = 4$ we take $\log(n)$

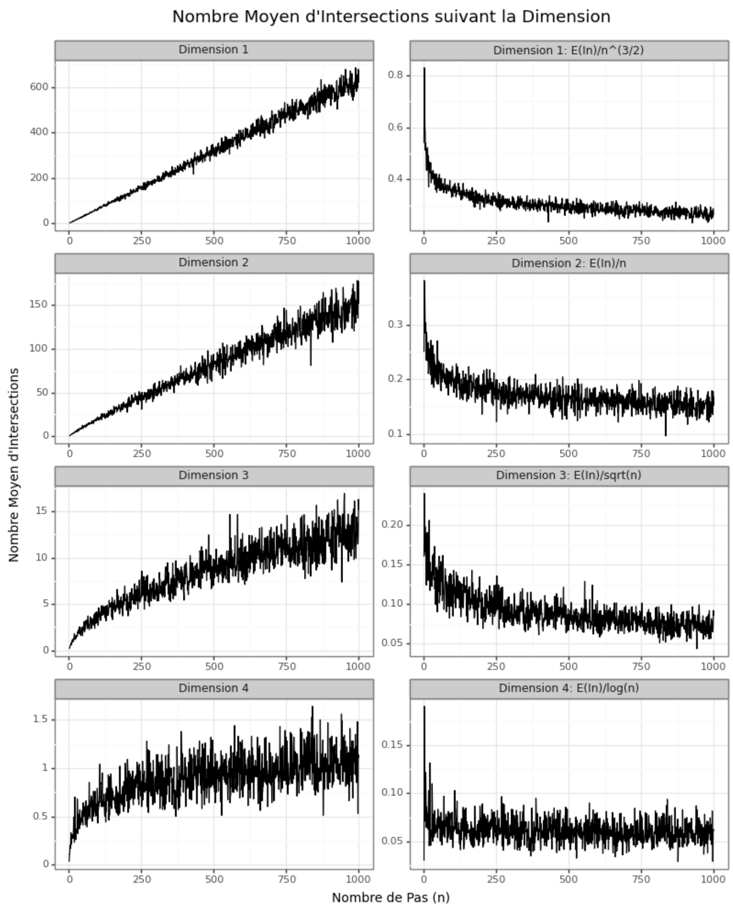


Figure 3: Average number of intersections divided by intersection coefficients

We notice from each of the graphs that dividing the number of intersections by the coefficients above leads to a convergence towards a non-zero constant, which means that these coefficients seem to be good approximations of the average number of intersections as n tends to infinity.

These results well illustrate the following theorem:

Theorem:

Let S_1 and S_2 be two simple random walks on $\mathbb{Z}^d$ starting at the origin. Let I_n be the number of intersections between the first n steps of S_1 and the first n steps of S_2. As $n \to \infty$, the expected values $E[I_n]$ are given as follows:

$$E[I_n] = \begin{cases} c_1 n^{\frac{3}{2}} + O(n^{\frac{1}{2}}), & \text{if } d = 1 \\ c_2 n + O(\ln(n)), & \text{if } d = 2 \\ c_3 n^{\frac{1}{2}} + O(1), & \text{if } d = 3 \\ c_4 \ln(n) + O(1), & \text{if } d = 4 \\ c_d + O(n^{\frac{4-d}{2}}), & \text{if } d \geq 5 \end{cases}$$

Proof:

Recalling the definition provided above of I_n and replacing k by 2, we note

$$I_n = \sum_{i=0}^{n} \sum_{j=0}^{n} 1_{\{S_{1_i} = S_{2_j}\}}.$$

By taking the expectation and using linearity,

$$\begin{aligned} E[I_n] &= \sum_{i=0}^{n} \sum_{j=0}^{n} E[1_{\{S_{1_i} = S_{2_j}\}}] \\ &= \sum_{i=0}^{n} \sum_{j=0}^{n} P\{S_{1_i} = S_{2_j}\}. \end{aligned}$$

Knowing that the probability of the random walk S^2 occurring is the same as that of its inverse occurring. Thus, the probability of both random walks S^1, after i steps, and S^2, after j steps, being at the same point is the same as the probability of a single random walk traversing the path of S^1 for the first i steps and the inverse of S^2 for the next j steps, then returning to 0 in $i + j$ steps. Hence, we obtain,

$$E[I_n] = \sum_{i=0}^{n} \sum_{j=0}^{n} P\{S_{i+j} = 0\}.$$

Replacing $i + j$ by k, we can notice that each value of k corresponds to $k + 1$ pairs of (i, j). The equation then becomes:

$$E[I_n] = \sum_{k=0}^{2n} (k+1) P\{S_k = 0\} = \sum_{k=0}^{2n} (k+1) p_k(0)$$

Moreover, since $S_0 = 0$, we have $p_0(0) = 1$, and thus:

$$E[I_n] = 1 + \sum_{k=1}^{2n} (k+1) p_k(0)$$

Using the Local Limit Theorem, we can obtain an estimate of $p_k(0)$, denoted $\overline{p}_k(0)$, such that,

$$\overline{p}_k(0) = 2\left(\frac{d}{2\pi k}\right)^{d/2} e^{-\frac{d|0|^2}{2}} = 2\left(\frac{d}{2\pi k}\right)^{d/2}.$$

With the estimation error $E(k,0) = |p_k(0) - \overline{p}_k(0)| = O\left(\frac{1}{k^{d/2+1}}\right)$.
We thus have,

$$\begin{aligned}
E[I_n] &= 1 + \sum_{k=1}^{2n}(k+1)(\overline{p}_k(0) + E(k,0)) \\
&= 1 + \sum_{k=1}^{2n}(k+1)2\left(\frac{d}{2\pi k}\right)^{\frac{d}{2}} + \sum_{k=1}^{2n}(k+1)O\left(\frac{1}{k^{\frac{d}{2}+1}}\right) \\
&= 1 + 2\left(\frac{d}{2\pi}\right)^{\frac{d}{2}}\left[\sum_{k=1}^{2n}\frac{1}{k^{\frac{d}{2}-1}} + \sum_{k=1}^{2n}\frac{1}{k^{\frac{d}{2}}}\right] + \sum_{k=1}^{2n}(k+1)O\left(\frac{1}{k^{\frac{d}{2}+1}}\right)
\end{aligned}$$

We will now estimate the values of each sum: Let,

$$\begin{aligned}
T_n &= \sum_{k=1}^{2n}(k+1)O\left(\frac{1}{k^{\frac{d}{2}+1}}\right) \\
&\leq \sum_{k=1}^{2n}(k+1)\frac{c}{k^{\frac{d}{2}+1}} \text{ for some } c \\
&\leq c\left[\sum_{k=1}^{2n}\frac{1}{k^{\frac{d}{2}}} + \sum_{k=1}^{2n}\frac{1}{k^{\frac{d}{2}+1}}\right]
\end{aligned}$$

Since, for all $k > 1$, $\frac{c}{k^{\frac{d}{2}}} \geq \frac{c}{k^{\frac{d}{2}+1}}$.

We obtain as n tends to infinity:

$$T_n = O\left(\sum_{k=1}^{2n}\frac{1}{k^{\frac{d}{2}}}\right)$$

We are left to deal with $V_n = \sum_{k=1}^{2n}\frac{1}{k^{\frac{d}{2}-1}}$ and $U_n = \sum_{k=1}^{2n}\frac{1}{k^{\frac{d}{2}}}$
Let's start with V_n:
For $d = 1$**:**

$$V_n = \sum_{k=1}^{2n}\frac{1}{k^{\frac{1}{2}-1}} = \sum_{k=1}^{2n}\sqrt{k}$$

So,

$$\int_0^{2n} \sqrt{x}dx \leq \sum_{k=1}^{2n} \sqrt{k} \leq \int_0^{2n} \sqrt{x+1}dx$$

$$\left[\left(\frac{2x}{3} \right)^{\frac{2}{3}} \right]_0^{2n} \leq \sum_{k=1}^{2n} \sqrt{k} \leq \left[\frac{2(x+1)^{\frac{2}{3}}}{3} \right]_0^{2n}$$

$$2(2n)^{\frac{2}{3}} - 2^{\frac{2}{3}} \leq \sum_{k=1}^{2n} \sqrt{k} \leq \frac{2(2n)^{\frac{2}{3}}}{3} + \frac{3}{2} \times \frac{2}{3} \times (2n)^{\frac{1}{2}} + \frac{3}{2} \times \frac{1}{2} \times \frac{2}{3}(2n)^{-\frac{1}{2}} + ... = \frac{2(2n)^{\frac{2}{3}}}{3} + O(n^{\frac{1}{2}})$$

Therefore, for $d = 1$, $V_n = \frac{2(2n)^{\frac{2}{3}}}{3} + O(n^{\frac{1}{2}})$.

For $d = 2$: $V_n = \sum_{k=1}^{2n} 1 = 2n$.

For $d = 3$: $V_n = \sum_{k=1}^{2n} \frac{1}{\sqrt{k}}$

$$\int_0^{2n} \frac{1}{\sqrt{x+1}}dx \leq \sum_{k=1}^{2n} \frac{1}{k^{\frac{1}{2}}} \leq 1 + \int_1^{2n} \frac{1}{\sqrt{x}}dx$$

$$[2\sqrt{x+1}]_0^{2n} \leq \sum_{k=1}^{2n} \frac{1}{k^{\frac{1}{2}}} \leq 1 + [2\sqrt{x}]_1^{2n}$$

$$2\sqrt{2n} - 2 \leq \sum_{k=1}^{2n} \frac{1}{k^{\frac{1}{2}}} \leq 1 + 2\sqrt{2n} - 2$$

So, for $d = 3$, $V_n = 2\sqrt{(2n)} + O(1)$.

For $d = 4$: $V_n = \sum_{k=1}^{2n} \frac{1}{k}$

$$\int_0^{2n} \frac{1}{x+1}dx \leq \sum_{k=1}^{2n} \frac{1}{k} \leq 1 + \int_1^{2n} \frac{1}{x}\,dx$$

$$\int_0^{2n-1} \frac{1}{x+1}dx \leq \sum_{k=1}^{2n} \frac{1}{k} \leq 1 + [\log(x)]_1^{2n}$$

$$\log(2n) \leq \sum_{k=1}^{2n} \frac{1}{k} \leq 1 + \log(2n).$$

So, for $d = 4$, $V_n = \log(2n) + O(1)$.

For d ≥ 5 : $\quad V_n = \sum_{k=1}^{2n} \frac{1}{k^{\frac{d}{2}-1}}$, which converges by the Riemann criterion. We can therefore write:

$$\begin{aligned}\sum_{k=1}^{2n} \frac{1}{k^{\frac{d}{2}-1}} &= \sum_{k=1}^{\infty} \frac{1}{k^{\frac{d}{2}-1}} - \sum_{k=2n+1}^{\infty} \frac{1}{k^{\frac{d}{2}-1}} \\ &\geq c_d - \int_{2n}^{\infty} \frac{1}{x^{\frac{d}{2}-1}}\, dx \\ &\geq c_d - \left[\frac{2}{(4-d)x^{\frac{d}{2}-2}}\right]_{2n}^{\infty} \\ &= c_d + \frac{2}{(d-4)(2n)^{\frac{d}{2}-2}}.\end{aligned}$$

Moreover, we have:

$$\begin{aligned}\sum_{k=1}^{\infty} \frac{1}{k^{\frac{d}{2}-1}} \geq \sum_{k=1}^{2n} \frac{1}{k^{\frac{d}{2}-1}} &\geq c_d + \frac{2(d-4)(2n)}{d^{\frac{d}{2}-2}} \\ c_d \geq \sum_{k=1}^{2n} \frac{1}{k^{\frac{d}{2}-1}} &\geq c_d + \frac{2(d-4)(2n)}{d^{\frac{d}{2}-2}}\end{aligned}$$

Therefore, for $d \geq 5$, $V_n = c_d + O\left(\frac{1}{n^{\frac{d}{2}-2}}\right)$.
We proceed in the same way for U_n and we finally obtain:

$$V_n = \begin{cases} 2\left(\frac{2n}{3}\right)^{\frac{2}{3}} + O(\sqrt{n}), & \text{for } d = 1, \\ 2n, & \text{for } d = 2, \\ 2\sqrt{2n} + O(1), & \text{for } d = 3, \\ \log(2n) + O(1), & \text{for } d = 4, \\ c_d + O\left(\frac{1}{n^{\frac{d}{2}-2}}\right), & \text{for } d \geq 5. \end{cases}$$

$$U_n = \begin{cases} 2\sqrt{2n} + O(1), & \text{for } d = 1, \\ \log(2n) + O(1), & \text{for } d = 2, \\ c_d + O\left(\frac{1}{n^{\frac{d}{2}-1}}\right), & \text{for } d \geq 3. \end{cases}$$

Substituting this into our initial equation yields what we wanted to prove, i.e.,

$$E[I_n] = \begin{cases} \frac{8}{3\sqrt{\pi}} n^{\frac{3}{2}} + O(\sqrt{n}), & \text{for } d = 1, \\ \frac{4}{\pi} n + O(\log(n)), & \text{for } d = 2, \\ \frac{6}{\pi}\sqrt{\frac{3n}{\pi}} + O(1), & \text{for } d = 3, \\ \frac{8}{\pi^2}\log(n) + O(1), & \text{for } d = 4, \\ c_d + O\left(\frac{1}{n^{\frac{d}{2}-2}}\right), & \text{for } d \geq 5. \end{cases}$$

References

[1] Curien, N. *Random Walk and Graphs.*

[2] Gregory F. Lawler and Vlada Limic. *Random Walk: A Modern Introduction.*

[3] Gregory F. Lawler, *Intersection of Random Walk.*