Société de Calcul Mathématique SA
Mathematical Modelling Company, Corp
Tools for decision help
since 1995

Random Walks in the Plane: An Energy-Based Approach

Three prizes offered by SCM

Bernard Beauzamy
August 2016

I. Introduction

We consider a simple random walk in the plane: a sequence of random variables X_n with values ± 1, probability $1/2$ in each case. Let $S_n = \sum_{j=1}^{n} X_j$ be the sum of the first n variables ; the initial value (at time $n=0$) is 0. We are interested in quantitative estimates of the behavior of S_n when n becomes large.

A well-known result is Khintchine's law of the iterated logarith (1924): almost surely, when $n \to +\infty$:

$$\limsup \frac{S_n}{\sqrt{2nLog\left(Log\left(n\right)\right)}} = +1 \text{ and } \liminf \frac{S_n}{\sqrt{2nLog\left(Log\left(n\right)\right)}} = -1$$

We make two comments:

1. Such estimates are not quantitative at all

A first attempt to obtain quantitative estimates, using the original probabilistic proof, was made by the author in :
http://www.scmsa.eu/archives/BB_paradoxes_probabilistes_2016_02.pdf

2. Their probabilistic appearance is misleading

Looking at such a statement, everyone has the impression that, for a given player, there are some unknown forces which will, sooner or later, bring his fortune close to Khintchin's curves (a Khintchin curve is of the form $y = \pm\sqrt{2x\,Log\left(Log\left(x\right)\right)}$). This is completely wrong; at any time, the game is only governed by the ± 1 rule, with equal probability.

What Khintchin's laws say, and, more generally, what any result about random walks says, is that there are more paths with some properties than paths with other properties. They are not individual results about each path; they are results about the number of paths with a given characteristic. Such results are in fact of combinatorial nature. For instance, at time n, the proportion of paths which never touched the curve $y = \sqrt{n}$ tends to 0 when $n \to +\infty$.

In order to prove such results, we take here a completely different approach, which is not probabilistic anymore, but relies upon a concept derived from "energy absorption". Our aim is also to obtain quantitative estimates, of the form:

Given a curve $y = \varphi\left(x\right)$, what is the proportion of paths which never touched the curve before the instant n ?

II. Basic settings

We consider that, at time 0, a unit of energy is put at the origin. This unit will then divide itself in two halves, at time $n = 1$, one at the point $\left(1,1\right)$ and one at the point $\left(1,-1\right)$. More generally, every time a division point is met, the available energy divides equally into the two possible paths. So, for instance, at the time $n = 2$, 3 points will receive some energy, namely $\left(2,2\right)$ receives 1/4, $\left(2,0\right)$ receives 1/2, $\left(2,-2\right)$ receives 1/4. At any step, in this configuration, the sum is always 1.

As it is well-known, the values of S_n are even if n is even, and are odd if n is odd. We observe that, obviously, at any time n, we have $\left|S_n\right| \le n$. The following Lemma is well-known (in what follows, for simplicity, we consider only the even values):

Lemma 1. - *Let $A_{2n,2k}$ be the point of coordinates $\left(2n,2k\right)$, with $k = -n,\ldots,n$. The number of paths from 0 to $A_{2n,2k}$ is:*

$$N\left(2n,2k\right) = \binom{2n}{n+k}$$

Proof of Lemma 1

If we want to reach this point in $2n$ steps, we need x times the value 1 and y times the value -1, with $x+y=2n$ and $x-y=2k$, which gives $x=n+k,\ \ y=n-k$. So there are $\binom{2n}{x}$ possible paths, which proves the result.

We write $N(A)$ for the total number of paths, starting at 0, finishing at A.

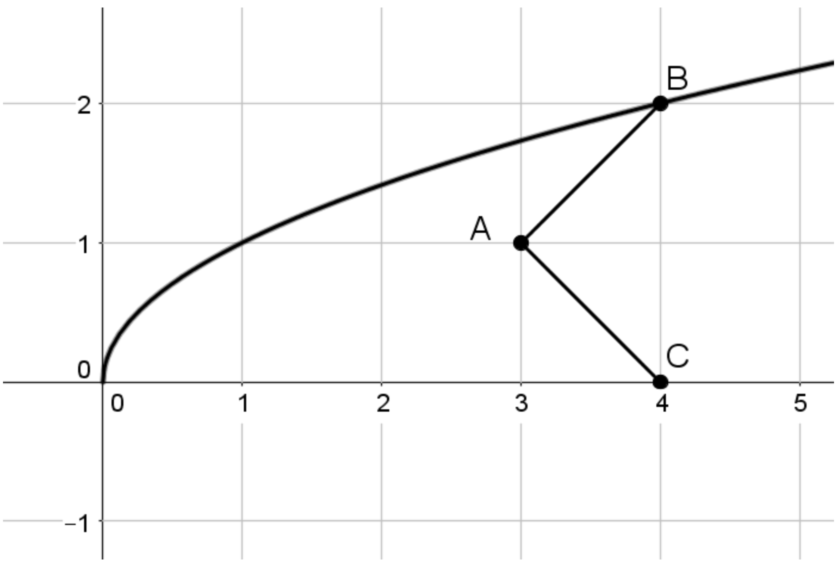
Since there is a total of 2^{2n} possible paths at time $2n,$ each point $A_{2n,2k}$ receives an amount of energy equal to:

$$e\left(A_{2n,2k}\right)=\frac{1}{2^{2n}}\binom{2n}{n+k}$$

The repartition of energy is given by a binomial law: there is more energy at the central points and very little energy at the extreme points ($A_{2n,2n}$ and $A_{2n,-2n}$).

In this preliminary approach, the total amount of energy remains the same. Now, we introduce a curve, $y=\varphi(x),$ located in the upper half-plane (the same holds for the lower half-plane, of course), and we want to investigate the probability that the random walk, up to time $n,$ remains constantly below this curve, which means that $S(j)<\varphi(j)$ for all $j=1,\ldots,n.$

Our representation, in order to investigate this phenomenon, will be the fact that the curve φ absorbs the energy. This means that, for any path which touches the curve, the corresponding energy disappears.



In this example, the point A sends its energy to both B and C, but B is on the curve we want to investigate, so this part of the energy disappears, and we are left with $e(C)=\frac{1}{2}e(A).$

The curve we want to investigate will be called the critical curve. It may be considered as a "black frontier" (in the sense of a black hole), meaning that it absorbs all energy it receives, and sends back nothing.

We have:

Proposition 2. - *Let* $y=\varphi(x)$ *be any critical curve, in the upper half-plane. The total energy left, at time* n*, is equal to the total probability to reach any of the points* $A_{n,k}$ *below the curve, that is satisfying* $k<\varphi(n)$*, without ever touching the curve at any time before (* $j \le n$ *).*

Proof of Proposition 2

This is a mere rephrasing of the disparition of energy. Any time a path touches the curve, it is annihilated, so what remains is the paths which never touched the curve.

If a time n is fixed, and a curve φ is fixed, we will call admissible a path with never touches it (at any time $j \le n$). For any point A in the plane, let $N_{ad}(A)$ be the number of admissible paths which reach A, and $p_{ad}(A)=\dfrac{N_{ad}(A)}{2^n}$ the probability to reach A by an admissible path. Proposition 2 states that:

$$\sum_{k=-n}^{n} e\left(A_{n,k}\right)=\sum_{k<\varphi(n)} p_{ad}\left(A_{n,k}\right)$$

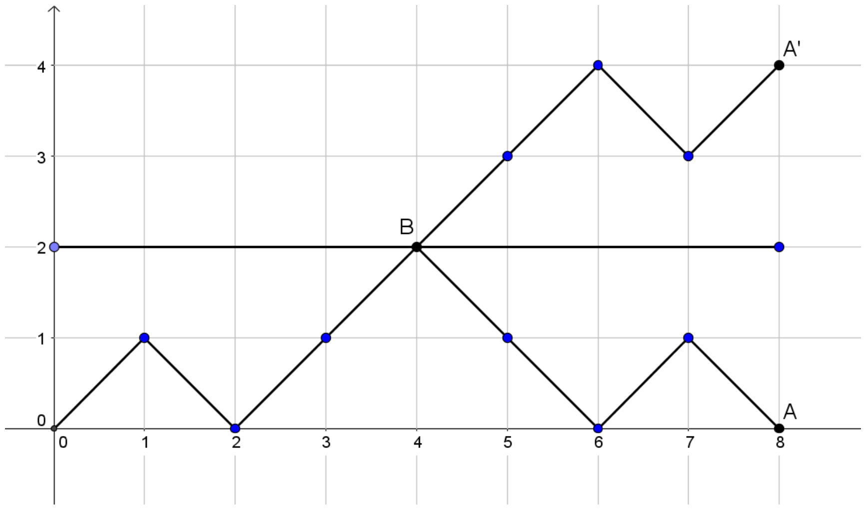
III. The case of an horizontal line

We now compute the number of admissible paths when the critical curve is an horizontal line segment:

Lemma 3. - *Let* $y=2k_0$ $(k_0 \ge 0)$ *be an horizontal line segment. Let* $A_{2n,2k}$*, with coordinates* $(2n,2k)$*, be any point that the random walk may reach, with* $k<k_0$*. The number of paths, starting at 0, finishing at* $A_{2n,2k}$*, which touch the horizontal segment at a time before* $2n$ *is* $N\left(A'_{2n,2k}\right)$*, where* $A'_{2n,2k}$ *is the symmetric of* $A_{2n,2k}$ *with respect to the line segment.*

Proof of Lemma 3

This property is well-known, under the name of "reflexion principle" :



Let B be the first time a path touches the segment (there may be several). There are as many paths from B to A than from B to A'.

Since the coordinates of $A'_{2n,2k}$ are $(2n, 4k_0 - 2k)$, the symmetric of $A_{2n,2k}$ is $A_{2n,4k_0-2k}$. So the number of paths which touch the segment $y = 2k_0$ at any time before n is, by Lemma 1:

$$N\left(A_{2n,4k_0-2k}\right) = \binom{2n}{n+2k_0-k}$$

Therefore, the number of paths which reach $A_{2n,2k}$ without ever touching the segment $y = 2k_0$ is:

$$N_{ad}\left(A_{2n,2k}\right) = \binom{2n}{n+k} - \binom{2n}{n+2k_0-k}$$

Proposition 4. - *Assume that our critical curve is the line segment $y = 2k_0$. The distribution of energy at time $2n$ is:*

$$e\left(A_{2n,2k}\right) = \frac{1}{2^{2n}}\left(\binom{2n}{n+k} - \binom{2n}{n+2k_0-k}\right)$$

Indeed, this follows immediately from the previous Lemma.

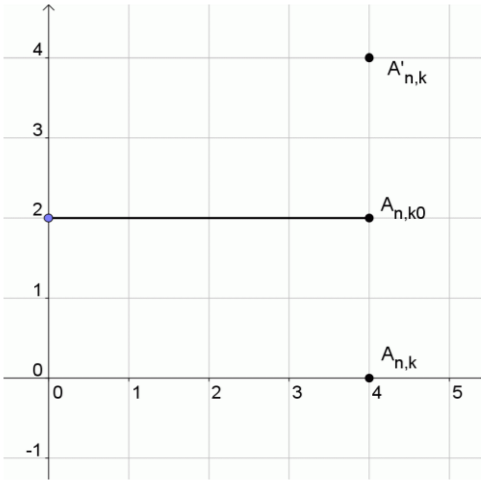
We observe that, due to the absorption, this distribution of energy is not symmetric anymore. Starting at $k = k_0$ and moving downwards, it first increases, reaches its maximum for some $k_1 < 0$ and then decreases.

Proposition 5. - *Assume that our critical curve is the line segment $y = 2k_0$. The energy left at time $2n$ is:*

$$e(2n) = 1 - \frac{2}{2^{2n}} \sum_{j=k_0+1}^{n} \binom{2n}{n+j} - \frac{1}{2^{2n}} \binom{2n}{n+k_0}$$

Proof of Proposition 5

Let us look at the picture below:



The critical line segment $y = 2k_0$ has two effects :

- No point $A_{2n,2k}$ above this segment receives any energy at all; there is a drop of total energy equal to the probability to reach this point;

- For every point strictly below this segment, there is a drop of energy equal to the probability to reach its symmetric.

Since both terms are equal, the total drop of energy (that is the total energy "swallowed" by the segment), instead of reaching the points $A_{2n,2k}$ not on the segment, is $\frac{2}{2^{2n}} \sum_{j=k_0+1}^{n} \binom{2n}{n+j}$.

Now, there is the single point $A_{2n,2k_0}$ which is its own symmetric and should be counted only once; this proves Proposition 5.

Corollary 6. - *Let* $2k_0$ *be a given threshold. The probability that the random walk never reaches this threshold at any time* $t \le 2n$ *is :*

$$p_{ad}(2n) = 1 - \frac{2}{2^{2n}} \sum_{j=k_0+1}^{n} \binom{2n}{n+j} - \frac{1}{2^{2n}} \binom{2n}{n+k_0}$$

It tends to 0 when $n \to +\infty$.

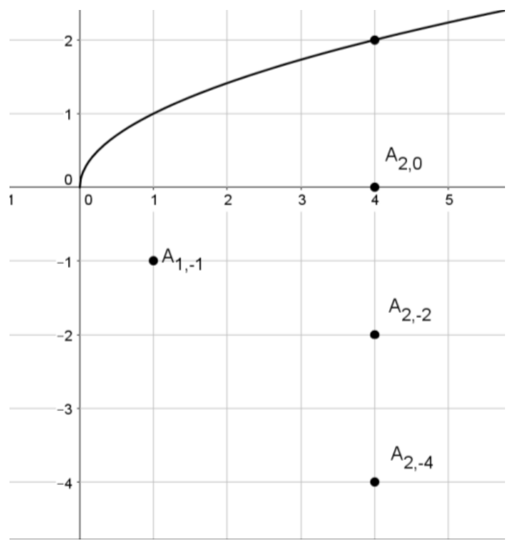
Indeed, when $n \to +\infty$, $\frac{1}{2^{2n}} \binom{2n}{n+k_0} \to 0$ and $\frac{1}{2^{2n}} \sum_{j=k_0+1}^{n} \binom{2n}{n+j} \to \frac{1}{2}$, since $\frac{1}{2^{2n}} \sum_{j=0}^{k_0} \binom{2n}{n+j} \to 0$.

IV. The critical curve $y=\sqrt{x}$

We now investigate the critical curve $y=\sqrt{x}$.

A. *Initial step*

At the beginning, we put an energy equal to 1 at the origin. The only admissible move (that is, a path which does not touch the critical curve) is $X_1=-1$, which leads us to the point $A_{1,-1}$; this point receives an energy equal to 1/2 (the other half is lost).



Now, we compute the energy which arrives on the vertical V_4. Only the points with even coordinates may be reached, namely $A_{2,0}, A_{2,-2}, A_{2,-4}$ (see picture above).

We have :

$$N\left(A_{1,-1} \to A_{2,2j}\right)=\binom{3}{\frac{3+2j+1}{2}}=\binom{3}{2+j}$$

The symmetric $A'_{2,2j}$ of $A_{2,2j}$ with respect to the line segment $y=2$ has coordinates $(4,4-2j)$. Therefore :

$$N\left(A_{1,-1} \to A'_{2,2j}\right)=\binom{3}{\frac{3+4-2j+1}{2}}=\binom{3}{4-j}$$

The energy sent by $A_{1,-1}$ to any of the points $A_{2,2j}$ is, taking into account the attenuation:

$$e\left(A_{2,2j}\right)=\frac{1}{2^4}\left(\binom{3}{2+j}-\binom{3}{4-j}\right)$$

We find:

$$e\left(A_{2,0}\right)=\frac{3}{16},\ e\left(A_{2,-2}\right)=\frac{3}{16},\ e\left(A_{2,-4}\right)=\frac{1}{16}.$$

B. General step

We now investigate the general step of the induction.

We discretize the x axis using the points of coordinate n^2 and the y axis using the points n. Let A_{n^2,j_n} be the point of coordinates $\left(n^2, j_n\right)$, with $j_n=-n^2,\ldots,n^2$ (so there are $2n^2+1$ such points). We want to compute the energy received by any of the points A_{n^2,j_n} knowing the energy received by all the $A_{(n-1)^2,j_{n-1}}$. In order to simplify our notation, we set $A_j=A_{(n-1)^2,j}$ and $B_k=A_{n^2,k}$. The points A_j and B_k are below the curve if, respectively, $j<n-1$ and $k<n.$

Let $e\left(A_j\right)$ be the energy received by each point A_j at the $n-1^{st}$ induction step (supposed to be known) ; we want to compute $e\left(B_k\right)$, the energy available at the n^{th} induction step.

The number of paths from A_j to B_k is :

$$N\left(A_j\to B_k\right)=\binom{n^2-(n-1)^2}{\dfrac{n^2-(n-1)^2+k-j}{2}}=\binom{2n-1}{\dfrac{2n-1+k-j}{2}}$$

Let B_k' be the symmetric of B_k with respect to the curve $y=n$; the coordinates of B_k' are: $B_k'=\left(n^2,2n-k\right)$. The number of paths from A_j to B_k' is:

$$N\left(A_j\to B_k'\right)=\binom{2n-1}{\dfrac{2n-1+2n-k-j}{2}}=\binom{2n-1}{\dfrac{4n-1-k-j}{2}}$$

We observe, in these formulas, that j and k cannot have the same parity: the sum $k+j$ or the difference $k-j$ must be odd, otherwise there is no path. Indeed, the squares n^2 are alternatively odd and even.

Therefore, for $j<n-1$ and $k<n,$ the amount of energy sent by A_j to B_k is:

$$e\left(A_j \to B_k\right) = \frac{e\left(A_j\right)}{2^{2n-1}}\left(\binom{2n-1}{\frac{2n-1+k-j}{2}} - \binom{2n-1}{\frac{4n-1-k-j}{2}}\right)$$

The total amount of energy available at each point B_k is obtained, summing upon j :

$$e\left(B_k\right) = \sum_{j=-(n-1)^2}^{n-2} \frac{e\left(A_j\right)}{2^{2n-1}}\left(\binom{2n-1}{\frac{2n-1+k-j}{2}} - \binom{2n-1}{\frac{4n-1-k-j}{2}}\right)$$

And the total amount of energy available at time n is the sum of these quantities, summing upon k :

$$e\left(V_n\right) = \sum_{k=-n^2}^{n-1} \sum_{j=-(n-1)^2}^{n-2} \frac{e\left(A_j\right)}{2^{2n-1}}\left(\binom{2n-1}{\frac{2n-1+k-j}{2}} - \binom{2n-1}{\frac{4n-1-k-j}{2}}\right)$$

C. Complete formula at step n

Let us set, for $n \geq 1$

$$w\left(n\,; j_{n-1}, j_n\right) = \frac{1}{2^{2n-1}}\left(\binom{2n-1}{\frac{2n-1+j_n-j_{n-1}}{2}} - \binom{2n-1}{\frac{4n-1-j_n-j_{n-1}}{2}}\right)$$

we have, for every j_n:

$$e\left(A_{n^2, j_n}\right) = \sum_{j_{n-1}=-(n-1)^2}^{n-2} \cdots \sum_{j_1=-2}^{0} w\left(1; j_0, j_1\right) w\left(2; j_1, j_2\right) \cdots w\left(n; j_{n-1}, j_n\right)$$

and:

$$e\left(V_{n^2}\right) = \sum_{j_n=-n^2}^{n-1} \sum_{j_{n-1}=-(n-1)^2}^{n-2} \cdots \sum_{j_1=-2}^{0} w\left(1; j_0, j_1\right) w\left(2; j_1, j_2\right) \cdots w\left(n; j_{n-1}, j_n\right)$$

Here, V_{n^2} denotes the vertical at time n^2, below the curve, that is the set of all points A_{n^2, j_n} with $j_n < n$. We want to prove that $e\left(V_{n^2}\right) \to 0$ when $n \to +\infty$.

D. Analysis of a single term

Let us look at the final sum in the above expression, that is :

$$D(n\,;j_{n-1})=\frac{1}{2^{2n-1}}\sum_{j_n=-n^2}^{n-1}\left(\binom{2n-1}{\frac{2n-1+j_n-j_{n-1}}{2}}-\binom{2n-1}{\frac{4n-1-j_n-j_{n-1}}{2}}\right)$$

Simple computations show that it may be written:

$$D(n\,;j_{n-1})=\frac{1}{2^{2n-1}}\left(\binom{2n-1}{\frac{n+j_{n-1}}{2}}+\binom{2n-1}{\frac{n+1+j_{n-1}}{2}}+\cdots+\binom{2n-1}{\frac{3n-1-j_{n-1}}{2}}\right)$$

Assume that n is even, $n=2m,$ so j_{n-1} is odd, $j_{n-1}=2j+1.$ The number of terms in the above sum is $2m-2j-1=n-j_{n-1}$; this number is minimal for $j=n-1$ (we have 1 term). It is maximal (namely $2n-1$) if we have $n+j_{n-1}\le 0$ and $\frac{3n-1-j_{n-1}}{2}\ge 2n-1$, that is $j_{n-1}\le -n.$ In this case, $D(n;j_{n-1})=1.$

We also observe that the sum extends on both sides of the median term $n.$

As we already sait, let V_{n^2} denote the vertical at $n^2,$ that is the set of points A_{n^2,j_n}. We have:

$$e\left(V_{n^2}\right)=\sum_{j_{n-1}=-(n-1)^2}^{n-2}e\left(A_{(n-1)^2,j_{n-1}}\right)D(n\,;j_{n-1})$$

The quantity $D(n\,;j_{n-1})$ may ve viewed as the proportion of energy sent by $A_{(n-1)^2,j_{n-1}}$ to the whole vertical V_{n^2}. It is an "attenuation coefficient", which is $\le 1.$ We observe that the total quantity of energy at step $n,$ $e\left(V_{n^2}\right)$ satisfies:

$$e\left(V_{n^2}\right)=\sum_{j_{n-1}=-(n-1)^2}^{n-2}e\left(A_{(n-1)^2,j_{n-1}}\right)D(n\,;j_{n-1})<\sum_{j_{n-1}=-(n-1)^2}^{n-2}e\left(A_{(n-1)^2,j_{n-1}}\right)=e\left(V_{(n-1)^2}\right)$$

so it is decreasing at each step (this was obvious, since some energy disappears and no energy is created). Moreover, the maximum value of the energy is decreasing at each step:

$$\max_{j_n} e\left(A_{n^2,j_n}\right)<\max_{j_{n-1}} e\left(A_{(n-1)^2,j_{n-1}}\right)$$

We also observe that the introduction of the critical curve leads to the fact that some energy disappears, because some paths do not exist anymore. It does not create any new path. Therefore, all "blocks" of terms which had their energy tending to zero in the normal scheme will have the same property in the new scheme.

V. Prizes offered

After this presentation of the subject, and of some basic properties, we would like to offer three prizes.

A. *Prize 1 : 500 Euros*

In the case of the critical curve $y = \sqrt{x}$ (see above), prove that the energy received by the n^{th} vertical, after attenuation due to the curve, tends to 0.

Recall that, in the above notation, V_n is the set of points on the n^{th} vertical which lie below the curve. So we want to show that $e(V_n) \to 0,$ when $n \to +\infty$.

Of course, since this quantity is decreasing, it is enough to do it for a subsequence, but we want quantitative estimates, namely, for any $\varepsilon > 0$, we want to know explicitly the value of some n_0 such that for all $n \geq n_0$, $e(V_n) < \varepsilon$.

B. *Prize 2 : 1 000 Euros*

Fix any $\varepsilon > 0$ and consider Khintchin's critical curve $y = (1-\varepsilon)\sqrt{2xLog(Log(x))}$. Prove that $e(V_n) \to 0,$ when $n \to +\infty$; same as above : we want explicit quantitative estimates.

C. *Prize 3 : 1 000 Euros*

Fix any $\varepsilon > 0$ and consider Khintchin's critical curve $y = (1+\varepsilon)\sqrt{2xLog(Log(x))}$. Prove that $\lim_{n \to +\infty} e(V_n) > 0$; same as above : we want explicit quantitative estimates.

D. *General rules for participation*

Each prize will be given only once, to the best contribution. A contribution will be considered only if it is well written and complete. All details should be given.

Please send the contributions (in English or in French) to contact@scmsa.com
no later than June 30th, 2017.

Everyone may participate (individuals, institutions, and so on).