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

Simple Random Walks in the Plane:

An Energy-Based Approach

Part I : Basic Facts

Bernard Beauzamy

June 2018

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 $S_N = \sum_{n=1}^{N} X_n$ be the sum of the first N variables. This random walk can be viewed as a game between two players A and B; at the n^{th} step, the first player receives 1 Euro from the second player if $X_n = +1$ and conversely if $X_n = 1$. So the sum S_N represents the increase of fortune of A compared to B at the end of N games ; this increase may of course be positive or negative. At the initial moment, we set $S_0 = 0$. Besides that, each player has an initial fortune, which is finite, or infinite in a theoretical setting. The game may stop when one of the players is ruined (his fortune becomes equal to 0). The general question is to study the behavior of S_N (possible values, with their probabilities), the duration of the game, depending upon the initial fortunes, and the asymptotic behavior, when $N \to +\infty$.

We observe that the behavior of S_N is determined by laws of Nature: one may repeat the experiment and check the results. But, at the same time, these laws are axiomatically defined, as we just did. Such random walks are probably the only example of laws of Nature which may be axiomatically defined: all laws in Physics are otherwise empirical.

Siège social et bureaux : 111, Faubourg Saint Honoré, 75008 Paris. Tel : 01 42 89 10 89. Fax : 01 42 89 10 69. www.scmsa.eu
Société Anonyme au capital de 56 200 Euros. RCS : Paris B 399 991 041. SIRET : 399 991 041 00035. APE : 7219Z

Among the many existing results on this topic, let us mention in particular:

- Feller's "Gambler's ruin" ; see [Feller]. The problem may be stated as follows : given an initial fortune and a barrier, what is the probability to reach the barrier without having first reached the barrier $y=0$ (which means ruin) ? The gambler's ruin does not care about a specific time, whereas we compute the probability for each specific time. We thank Doron Zeilberger for useful discussions about this comparison.

- Asymptotic results : Khintchin's law of the iterated logarith (1924); see [Khintchin]: 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 will here present a new approach to such problems, which is "energy based" and not probabilistic in nature. This will allow us to develop a unified framework, and to obtain quantitative estimates which were not known previously.

Indeed, the probabilistic appearance of Khinchin's laws 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 $y = \pm\sqrt{2x\, Log\left(Log\left(x\right)\right)}$, of course). 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 $n \to +\infty$.

Our approach relies upon a concept derived from "energy absorption". Our aim is to obtain quantitative estimates, of the following form:

Given a curve $y=\varphi(n)$, or possibly a couple of curves $y=\left(\varphi_1(n),\varphi_2(n)\right)$ with $\varphi_1>0, \varphi_2<0$, what is the proportion of paths which never touched the curve(s) before the instant N ?

II. Basic settings

A. Preliminary tools

At any time $n,$ we have of course $|S_n| \le n.$ The values of S_n are even if n is even, and are odd if n is odd.

The following Lemma is well-known (see for instance [1]) ; it simply reflects the combinatorics:

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

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

Proof of Lemma 1

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

The following Lemma is very intuitive: it indicates that, half of the time, the random walk is above the $x-$axis:

Lemma 2. – *For all* $n \ge 1$*:*

$$P(S_n \ge 0) > \frac{1}{2}.$$

Proof of Lemma 2

This is clear, since $P(S_n>0)=P(S_n<0)$, $P(S_n<0)+P(S_n=0)+P(S_n>0)=1$ and $P(S_n \ge 0)=P(S_n>0)+P(S_n=0)$.

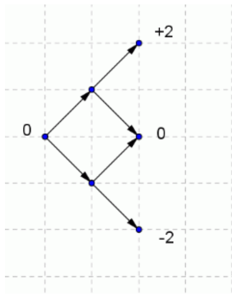
B. Introducing the energy

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 $(1,1)$ and one at the point $(1,-1)$. 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 $(2,2)$

receives 1/4, $(2,0)$ receives 1/2, $(2,2)$ receives 1/4. At any step, in this configuration, the sum is always 1.

In what follows, we will almost always restrict ourselves to the case where n is even. This means that the elementary game consists in two repetitions, $X_1 + X_2$, with :

$$P(X_1 + X_2 = -2) = \frac{1}{4}, P(X_1 + X_2 = 0) = \frac{1}{2}, P(X_1 + X_2 = 2) = \frac{1}{4} \qquad (1)$$



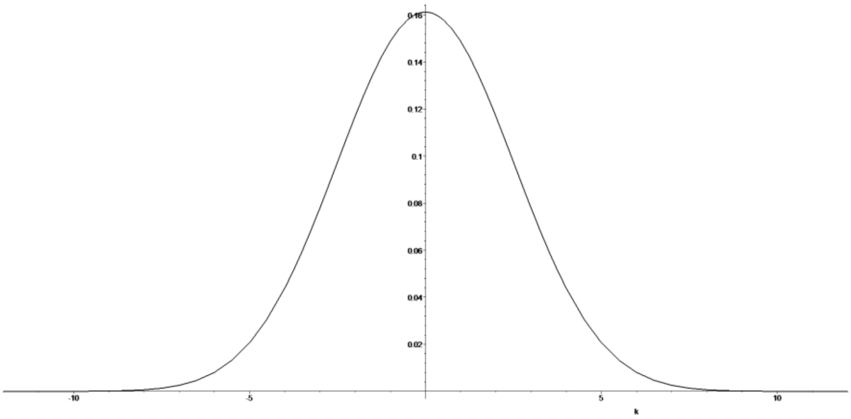
So an energy put at any point will divide into four : one fourth 2 steps above, one half at the same level, one fourth 2 steps below: see picture.

We write $N(A) = N(0 \to A)$ for the total number of paths, starting at 0, finishing at A and, more generally, $N(A \to B)$ for the number of paths starting at A, finishing at B.

In this basic setting, since the energy 1 is put at O and since there is a total of 2^{2n} possible paths $N(A_{2n,2k})$ at time $2n$, each point $A_{2n,2k}$ receives an amount of energy equal to:

$$P(S_{2n} = 2k) = e(A_{2n,2k}) = \frac{1}{2^{2n}} \binom{2n}{n+k} \qquad (2)$$

We see that 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 $A_{2n,-2n}$).



Example of energy distribution for $n = 12$

We consider, for simplicity, only the even values of n, k ; let $f(n,k) = e(A_{2n,2k})$ be the energy put at the point of coordinates $2n, 2k$. It satisfies for any k :

$$f(n,k) = \frac{1}{4} f(n-1,k-1) + \frac{1}{2} f(n-1,k) + \frac{1}{4} f(n-1,k+1) \qquad (1)$$

Now, we observe, using the symmetry of the process, that:

$$f(n,0)=\frac{1}{2}f(n-1,0)+\frac{1}{2}f(n-1,1) \tag{2}$$

C. Horizontal lines

We first study the decrease of probability on each horizontal line. The probability to reach $(2n,2k)$ is $p_n=\frac{1}{2^{2n}}\binom{2n}{n+k}$ and the probability to reach $(2n+2,2k)$ is $p_{n+1}=\frac{1}{2^{2n+2}}\binom{2n+2}{n+k+1}$. The condition $p_{n+1}\le p_n$ is equivalent to:

$$\frac{1}{4}\binom{2n+2}{n+k+1}\le\binom{2n}{n+k}$$

which, after simplification, is equivalent to:

$$n+1\ge 2k^2$$

So, for fixed $k,$ the probability first increases and then decreases.

The probability to reach $A_{2n,2k}$ is, as we saw, $e(A_{2n,2k})=\frac{1}{2^{2n}}\binom{2n}{n+k}$. For a given $k,$ this quantity tends to 0 when $n\to+\infty$.

D. Diagonals

We investigate the probability to reach a point $(2n+2k,2k)$, that is the $2k^{th}$ point on the $2n^{th}$ diagonal. We use only even values, as before. The 0^{th} diagonal, denoted by D_0, contains 1 at the origin an then $\frac{1}{4^k}$ at the $2k^{th}$ place. So the values are decreasing. The probability to reach $A_{2n+2k,2k}$ is $e(A_{2n+2K,2k})=\frac{1}{2^{2n+2k}}\binom{2n+2k}{n+2k}=\frac{1}{2^{2n+2k}}\binom{2n+2k}{n}$, which is decreasing in $k,$ for fixed n.

E. Further changes of variables

We introduce new notation, which will be useful in Part II.

We set, for any $n\ge 1$ and $k\le n$:

$$x(n,k)=\frac{1}{2}\big(f(n,k-1)+f(n,k)\big)$$

Lemma 2. - *We have, for any* $n,\ k \le n$ *:*

$$x(n,k) = \frac{1}{2^{2n+1}}\binom{2n+1}{n+k}$$

Proof of Lemma 2

Indeed, we have:

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

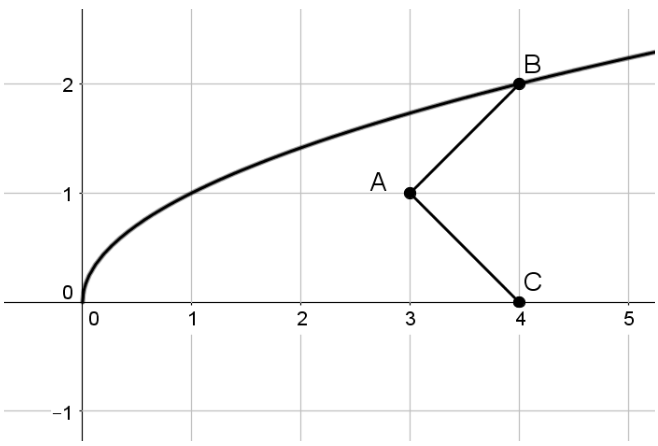
using Pascal's formula. This proves Lemma 2.

III. Introducing a barrier

A. General definition

In the preliminary approach, the total amount of energy remained the same at each time step. 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 $j=1,...,n$. Later, we will investigate the probability to remain between the curve and its symmetric, which means $|S(j)|<\varphi(j)$, or, more generally, to remain between two curves : $-\varphi_1(j)<S(j)<\varphi_2(j)$.

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.



Example of energy absorption

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

The curve we introduce 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 3. - *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 $k<\varphi(n)$, without ever touching the curve at any time before ($j \le n$).*

Proof of Proposition 3

This is a mere rephrasing of the disappearance of energy. Any time a path touches the curve, it is annihilated, so what remains is the set of 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 $p_{ad}(A)=\frac{N_{ad}(A)}{2^n}$ the probability to reach A by an admissible path. Proposition 3 states that:

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

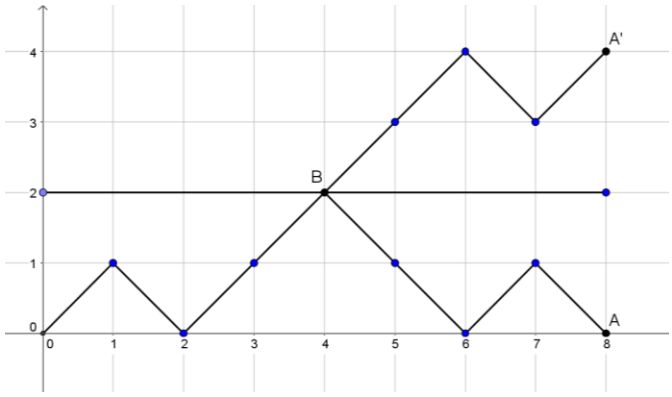
B. The case of an horizontal line

We now compute the number of admissible paths when the critical curve is a simple horizontal line segment. As we already said, we restrict ourselves to the even case.

Proposition 4. - *Let $y=2\xi$ ($\xi \ge 0$) be an horizontal line segment. Let $A_{2n,2k}$, with coordinates $(2n,2k)$, be any point that the random walk may reach, with $k<\xi$. The number of paths, starting at 0, finishing at $A_{2n,2k}$, which touch the horizontal segment at a time before $2n$ is $N\left(A_{2n,2\xi-2k}\right)$, where $A_{2n,2\xi-2k}$ is the symmetric of $A_{2n,2k}$ with respect to the line segment.*

Proof of Proposition 4

This property is well-known (see for instance [1]), under the name of "reflection principle" :



The reflection 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', symmetric of A with respect to the barrier.

The symmetric of $A_{2n,2k}$ is $A_{2n,4\xi-2k}$. So the number of paths which touch the segment $y=2\xi$ at any time before n is, by Lemma 1 :

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

This proves Proposition 4.

From this Lemma follows that the number of paths which reach $A_{2n,2k}$ without ever touching the segment $y=2\xi$ is:

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

Corollary 5. – *For any real x and any $n\geq 1$, we have:*

$$P\left(\exists k\leq n, S_k>x\right)\leq 2P\left(S_n>x\right)$$

Proof of Corollary 5

Fix any level x and time n. The probability to be above x at time n is, by definition:

$$P\left(S_n>x\right)=\frac{1}{2^n}\sum_{k>x}N\left(A_{n,k}\right)$$

Consider now the paths which touched x before time n. There are two types:

- either such a path finishes at a point $A_{n,k}$ $k>x$
- or they finish at a point $A_{n,k}$ $k\leq x$

But, by the reflection principle, the number of the second type is equal to the number of the first type ; this proves the Corollary.

We may indicate another proof, which is not of combinatorial type, but purely probabilistic. It comes from [Velenik], §2.3.

For any fixed x real, we set:

$$\sigma_x = \inf \{k \geq 0, S_k > x\}$$

This is the first time when the sequence S_k is above the value x. The events $\{\sigma_x = k\}$ are mutually disjoint, and we have:

$$P\{\exists k \leq n, S_k > x\} = \sum_{k=1}^{n} P\{\sigma_x = k\}$$

For $k = 1, ..., n,$ we introduce the event:

$$U_k = \{S_k \leq S_n\}.$$

The events $\{\sigma_x = k\} \cap U_k$, $k = 1, ..., n,$ are a partition of the event $\{S_n > x\}$; indeed, they are disjoint and their union is the set $\{S_n > x\}$: if $S_n > x,$ there is a k, $1 \leq k \leq n$ such that $S_k > x$.

Therefore:

$$P\{S_n > x\} = \sum_{k=1}^{n} P\left(\{\sigma_x = k\} \cap U_k\right) = \sum_{k=1}^{n} P\left(\sigma_x = k\right) P\left(U_k\right)$$

Indeed, the event $\sigma_x = k$ depends upon $X_1, ..., X_k,$ and the event U_k can be writtten $S_n - S_k > 0,$ that is $X_{k+1} + \cdots + X_n > 0$; so, it is independent from $X_1, ..., X_k$.

We have:

$$\sum_{k=1}^{n} P\left(\sigma_x = k\right) P\left(U_k\right) \geq \min_{k=1,...,n} P\left(U_k\right) \sum_{k=1}^{n} P\left(\sigma_x = k\right)$$

But $P\left(U_k\right) = P\left(X_{k+1} + \cdots + X_n > 0\right) > \frac{1}{2}$ by Lemma 2 above: all partial sums have the same law. Therefore:

$$\sum_{k=1}^{n} P\left(\sigma_x = k\right) P\left(U_k\right) \geq \frac{1}{2} \sum_{k=1}^{n} P\left(\sigma_x = k\right) = \frac{1}{2} P\left(\bigcup\left(\sigma_x = k\right)\right)$$

But the set $\bigcup\left(\sigma_x = k\right)$ can be described by the fact that there is a $k,$ $1 \leq k \leq n,$ such that $S_k > x.$ This proves Corollary 5.

In the sequel, we denote by W_{2n} the "vertical" at time $2n$. This is the set of points $A_{2n,2k}$, $k=-n,\dots,n$. We also denote by E_{2n} the total energy on this vertical : $E_{2n}=\sum_{k=-n}^{n} e\left(A_{2n,2k}\right)$. In this setting, E_{2n} is the probability that the game reaches time $2n$, or, in other words, did not stop earlier.

Proposition 6. - *Assume that our critical curve is the line segment $y=2\xi$, $\xi \geq 1$. The energy left at time $2n$ is:*

$$E_{2n}=1-\frac{2}{2^{2n}}\sum_{k=\xi+1}^{n}\binom{2n}{n+2\xi-k}=\frac{1}{2^{2n}}\sum_{j=-\xi}^{\xi}\binom{2n}{n+j}$$

Proof of Proposition 6

The critical line segment $y=2\xi$ 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 W_{2n}, is $\frac{2}{2^{2n}}\sum_{k=\xi+1}^{n}\binom{2n}{n+2\xi-k}$. This proves Proposition 5.

C. *Game with bounded initial fortune*

We start with a simple case:

1. No initial fortune

As a simple example, let us consider the following situation: The player A has no initial fortune, and there is no restriction on B. Our question is: what is the probability that the game lasts at least $2n$ moves ? There is a ruin of A if the random walk touches the x axis. Of course, in order that the game initially starts, the player A must win the first two games. So the starting point is $A_{2,2}$ which is reached with probability $1/4$. Let us denote by $A_{2n,2k}$ the point with coordinates $\left(2n,2k\right)$, $k=1,\dots,n$ and $B_{2n,2k}$ the point with coordinates $\left(2n,-2k\right)$, $k=1,\dots,n$.

Proposition 7. - *Assume that A has no initial fortune. The probability that the game lasts at least $2n$ moves is:*

$$p = \frac{1}{2^{2n+2}}\left(\binom{2n-2}{n-1} + \binom{2n-2}{n}\right)$$

Proof of Proposition 7

The probability to reach $A_{2n,2k}$ starting from $A_{2,2}$ is:

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

The probability to reach $B_{2n,2k}$ starting from $A_{2,2}$ is:

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

And the probability to reach $A_{2n,2k}$ without touching the x axis is:

$$P_+\left(A_{2,2} \to A_{2n,2k}\right) = P\left(A_{2,2} \to A_{2n,2k}\right) - P\left(A_{2,2} \to B_{2n,2k}\right)$$

The probability that the game did not stop at time $2n$ is the sum of the probabilities to reach one of the points $A_{2n,2k}$, $k = 1,\dots,n$, without touching the x axis. Let V_{2n} be the upper part of the vertical at time $2n$, that is $V_{2n} = \left\{A_{2n,2k}\,;\, k = 1,\dots,n\right\}$. We have:

$$\begin{aligned} P_+\left(A_{2,2} \to V_{2n}\right) &= \sum_{k=1}^{n} P_+\left(A_{2,2} \to A_{2n,2k}\right) = \frac{1}{2^{2n-2}}\left(\sum_{k=1}^{n}\binom{2n-2}{n+k-2} - \sum_{k=1}^{n-2}\binom{2n-2}{n+k}\right) \\ &= \frac{1}{2^{2n-2}}\left(\binom{2n-2}{n-1} + \binom{2n-2}{n}\right) \end{aligned}$$

which proves Proposition 7. Direct computation shows that, for $n \geq 31$, $p \leq 0.05$ and for $n \geq 795$, $p \leq 0.01$.

Using Stirling's formula, we may easily compute asymptotic estimates, when $n \to +\infty$:

$$P_+\left(0 \to V_{2n}\right) = \frac{1}{2\sqrt{\pi n}} + O\left(\frac{1}{n^{3/2}}\right) \text{ when } n \to +\infty.$$

We now turn to a simple generalization of the previous case:

2. Bounded initial fortune

We assume now that the player A has an initial fortune equal to $2m$ euros, $m \geq 1$; again, there is no limit upon B.

Proposition 8. – *Assume that A has an initial fortune of $2m$ Euros, $m \geq 1$. The probability that the game lasts at least until time $2n$ is:*

$$p = \frac{1}{2^{2n}} \sum_{k=-m+1}^{m} \binom{2n}{n+k}$$

Proof of Proposition 8

The random walk should not go below the line D with equation $y = -2m$; the starting point is O. The possible targets for the random walk are the points $A_{2n,2k}$, $k = -m+1, \dots, n$. The symmetric $B_{2n,2k}$ of $A_{2n,2k}$ with respect to D has coordinates $(2n, -2k-4m)$.

The probability to reach $A_{2n,2k}$ starting from O is:

$$P\left(0 \to A_{2n,2k}\right) = \frac{1}{2^{2n}} \binom{2n}{n+k}$$

The probability to reach $B_{2n,2k}$ starting from O is:

$$P\left(0 \to B_{2n,2k}\right) = \frac{1}{2^{2n}} \binom{2n}{n-k-2m}$$

As before:

$$P_+\left(A_0 \to A_{2n,2k}\right) = P\left(A_0 \to A_{2n,2k}\right) - P\left(A_0 \to B_{2n,2k}\right)$$

which gives:

$$P_+\left(0 \to V_{2n}\right) = \frac{1}{2^{2n}} \left(\sum_{k=-m+1}^{n} \binom{2n}{n+k} - \sum_{k=m+1}^{n-2m} \binom{2n}{n-k-2m} \right)$$

But the paths starting at O, arriving at $-2m-2, -2m-4, \dots, -2n$ have the same probability as the paths arriving at $2m+2, 2m+4, \dots, 2n$. In the formula above, we keep only $2m$ terms, namely:

$$P_+\left(0 \to V_{2n}\right) = \frac{1}{2^{2n}} \sum_{k=-m+1}^{m} \binom{2n}{n+k}$$

This proves Proposition 8. Direct computations show that, if our initial fortune is 1000 Euros, we have probability $\frac{1}{2}$ to stay in the game for $n = 5\,092\,958$ time steps and probability 0.95 to last at least $n = 1\,410\,791$ time steps.

It is clear that $E_{2n} \to 0$ when $n \to +\infty$. Indeed, in the expression $E_{2n} = \frac{1}{2^{2n}} \sum_{j=-\xi}^{\xi} \binom{2n}{n+j}$, there is a fixed number of terms and each term tends to 0 when $n \to +\infty$. But we want to make this statement quantitative.

IV. Gaussian interpretation

We have, using the approximation of the binomial law by the normal law, for fixed ξ_1, ξ_2:

$$P\left(2\xi_1 \le S_{2n} \le 2\xi_2\right) = \frac{1}{2^{2n}} \sum_{j=\xi_1}^{\xi_2} \binom{2n}{n+j} \approx \int_{\xi_1}^{\xi_2} \exp\left(-\frac{t^2}{2\sigma^2}\right) \frac{dt}{\sigma\sqrt{2\pi}}, \quad \text{with } \sigma^2 = 2n.$$

We want to make this approximation precise.

Lemma 1. – *For any* n *and any* x, $0 \le x \le n$, *we have:*

$$P\left(S_n \ge x\right) \le e^{-\frac{x^2}{2n}}$$

Proof of Lemma 1

We know that $E\left(S_n\right) = 0$ and $\operatorname{var}\left(S_n\right) = n$. Using Markov's Inequality $P\left(X \ge \varepsilon\right) \le \frac{E\left(X\right)}{\varepsilon}$, we write, for any $\lambda > 0$:

$$P\left(S_n \ge x\right) = P\left(e^{\lambda S_n} \ge e^{\lambda x}\right) \le e^{-\lambda x} E\left(e^{\lambda S_n}\right)$$

We have also:

$$e^{-\lambda x} E\left(e^{\lambda S_n}\right) = e^{-\lambda x} E\left(\prod_1^n e^{\lambda X_k}\right) = e^{-\lambda x}\left(E e^{\lambda X_1}\right)^n$$

But:

$$E\left(e^{\lambda X_1}\right) = \frac{e^{-\lambda} + e^{\lambda}}{2} \le e^{\lambda^2/2} \tag{1}$$

Indeed,

$$\frac{e^{-\lambda} + e^{\lambda}}{2} = \sum_{k=0}^{+\infty} \frac{\lambda^{2k}}{(2k)!}, \quad e^{\lambda^2/2} = \sum_{k=0}^{+\infty} \frac{\lambda^{2k}}{2^k k!}$$

and $2^k k! \le (2k)!$ which proves (1). We deduce from (1), for any λ:

$$P(S_n \ge x) \le e^{-\lambda x} e^{n\lambda^2/2}$$

and if we take $\lambda = \frac{x}{n}$, we obtain the required estimate. This proves Lemma 1.

In order to obtain reverse inequalities, we use Berry-Esseen Theorem [Berry-Esseen], which may be stated as follows:

For all x and all n:

$$\left| P\left(\frac{S_n}{\sqrt{n}} \le x \right) - \int_{-\infty}^{x} \exp\left(\frac{-t^2}{2} \right) \frac{dt}{\sqrt{2\pi}} \right| \le \frac{1}{\sqrt{n}}$$

which we write under the form:

$$\left| P\left(S_{2n} \le x\sqrt{2n} \right) - \int_{-\infty}^{x\sqrt{2n}} \exp\left(\frac{-t^2}{2\sigma^2} \right) \frac{dt}{\sigma\sqrt{2\pi}} \right| \le \frac{1}{\sqrt{2n}}, \text{ with } \sigma = \sqrt{2n}$$

or, with $\xi_1 = \frac{x\sqrt{2n}}{2}$:

$$\left| P\left(S_{2n} \le 2\xi_1 \right) - \int_{-\infty}^{2\xi_1} \exp\left(\frac{-t^2}{2\sigma^2} \right) \frac{dt}{\sigma\sqrt{2\pi}} \right| \le \frac{1}{\sqrt{2n}}$$

But, by (2):

$$P\left(S_{2n} \le 2\xi_1 \right) = \frac{1}{2^{2n}} \sum_{k \le \xi_1} \binom{2n}{n+k}$$

Therefore:

$$\left| \frac{1}{2^{2n}} \sum_{k \le \xi_1} \binom{2n}{n+k} - \int_{-\infty}^{2\xi_1} \exp\left(\frac{-t^2}{2\sigma^2} \right) \frac{dt}{\sigma\sqrt{2\pi}} \right| \le \frac{1}{\sqrt{2n}} \tag{3}$$

and also with ξ_2 :

$$\left| \frac{1}{2^{2n}} \sum_{k \le \xi_2} \binom{2n}{n+k} - \int_{-\infty}^{2\xi_2} \exp\left(\frac{-t^2}{2\sigma^2} \right) \frac{dt}{\sigma\sqrt{2\pi}} \right| \le \frac{1}{\sqrt{2n}}$$

Taking the difference, we obtain:

$$\left| \frac{1}{2^{2n}} \sum_{k=\xi_1+1}^{\xi_2} \binom{2n}{n+k} - \int_{2\xi_1}^{2\xi_2} \exp\left(\frac{-t^2}{2\sigma^2} \right) \frac{dt}{\sigma\sqrt{2\pi}} \right| \le \sqrt{\frac{2}{n}} \tag{4}$$

Formula (4) has an interpretation, namely that the energy, on any vertical W_{2n}, between the levels $2\xi_1$ and $2\xi_2$, may be viewed as a gaussian integral between these two levels, the variance of the law being the distance between 0 and the vertical (this distance is $2n$). The error in this approximation is smaller than $\sqrt{\frac{2}{n}}$.

From (4), we immediately deduce an estimate for the sum $\frac{1}{2^{2n}} \sum_{k=-\xi}^{\xi} \binom{2n}{n+k}$, valid for all n :

Proposition 8. - *For all ξ and n, we have the estimate:*

$$E_{2n} \le \frac{2\xi}{\sqrt{\pi n}} + \sqrt{\frac{2}{n}}$$

Proof of Proposition 8

Indeed, by (4):

$$E_{2n} \le \int_{-2\xi}^{2\xi} \exp\left(\frac{-t^2}{2\sigma^2} \right) \frac{dt}{\sigma\sqrt{2\pi}} + \sqrt{\frac{2}{n}} = \int_{-\frac{2\xi}{\sqrt{2n}}}^{\frac{2\xi}{\sqrt{2n}}} \exp\left(\frac{-t^2}{2} \right) \frac{dt}{\sqrt{2\pi}} + \sqrt{\frac{2}{n}} \le \frac{2\xi}{\sqrt{\pi n}} + \sqrt{\frac{2}{n}}$$

which proves Proposition 8.

We now turn to lower estimates for $P(S_n > k)$, in terms of Gaussian integrals.

Proposition 9. - *If* $n > 32\pi e$ and $k < \sqrt{n}$, *we have, with* $c = \frac{1}{4\sqrt{2\pi}}$:

$$P(S_n \geq k) \geq c \exp\left(-\frac{k^2}{2n}\right)$$

Proof of Proposition 9

We write Berry-Essen Theorem under the form:

For all x and all n:

$$\left| P\left(\frac{S_n}{\sqrt{n}} > x\right) - \int_x^{+\infty} \exp\left(\frac{-t^2}{2}\right)\frac{dt}{\sqrt{2\pi}} \right| \leq \frac{1}{\sqrt{n}}$$

With $x = \frac{k}{\sqrt{n}}$, it gives:

$$P(S_n > k) \geq \int_{\frac{k}{\sqrt{n}}}^{+\infty} \exp\left(\frac{-t^2}{2}\right)\frac{dt}{\sqrt{2\pi}} - \frac{1}{\sqrt{n}}$$

Let $f(x)$ be the density of Gauss Law and $F(x)$ be the repartition function; we have the estimate, for all $x > 0$ ([Komatsu]):

$$F(x) > \frac{2f(x)}{\sqrt{x^2+4}+x}$$

which gives here:

$$\int_{\frac{k}{\sqrt{n}}}^{+\infty} \exp\left(\frac{-t^2}{2}\right)\frac{dt}{\sqrt{2\pi}} > \frac{2f\left(\frac{k}{\sqrt{n}}\right)}{\sqrt{\frac{k^2}{n}+4}+\frac{k}{\sqrt{n}}}$$

But, if $k \leq \sqrt{n}$ then $\sqrt{\frac{k^2}{n}+4}+\frac{k}{\sqrt{n}} \leq \sqrt{5}+1 < 4$ and:

$$\int_{\frac{k}{\sqrt{n}}}^{+\infty} \exp\left(\frac{-t^2}{2}\right)\frac{dt}{\sqrt{2\pi}} > \frac{1}{2\sqrt{2\pi}} \exp\left(-\frac{k^2}{2n}\right)$$

Moreover, $\frac{1}{\sqrt{n}} < \frac{1}{4\sqrt{2\pi}} \exp\left(-\frac{k^2}{2n}\right)$ is satisfied since

$$\frac{1}{\sqrt{n}} < \frac{1}{4\sqrt{2\pi}} \exp\left(-\frac{1}{2}\right) \le \frac{1}{4\sqrt{2\pi}} \exp\left(-\frac{k^2}{2n}\right)$$

which is realized, since we assumed $n > 32\pi e$.

So we obtain:

$$P(S_n > k) \ge \int_{\frac{k}{\sqrt{n}}}^{+\infty} \exp\left(\frac{-t^2}{2}\right)\frac{dt}{\sqrt{2\pi}} - \frac{1}{\sqrt{n}} > \frac{1}{2\sqrt{2\pi}} \exp\left(-\frac{k^2}{2n}\right) - \frac{1}{4\sqrt{2\pi}} \exp\left(-\frac{k^2}{2n}\right) = \frac{1}{4\sqrt{2\pi}} \exp\left(-\frac{k^2}{2n}\right)$$

which proves our claim.

For $\xi = 0,$ we find the estimate $E_{2n} \le \sqrt{\frac{2}{n}}$, whereas a direct application of Stirling's formula gives $E_{2n} \le \frac{1}{\sqrt{\pi n}}$, so the estimate in Proposition 8 is not best possible.

Corollary 10. – *If the initial fortune of* A *is* $2m$, *the probability that the game lasts at least until time* $2n$ *satisfies the asymptotic estimate :*

$$E_{2n} \le \frac{2m}{\sqrt{\pi n}} + \sqrt{\frac{2}{n}}$$

This is an immediate consequence of Proposition 8.

The setting in terms of Gaussian integrals is much easier to handle, since these integrals are simpler to manipulate than binomial sums. Let us give a complete reinterpretation of the previous paragraph: energy absorption in case of a barrier at ξ.

In this continuous setting, there is no need to differentiate between the odd and even cases, which is also a simplification.

The symmetric of a point $A_{n,t}$ with respect to the barrier $y = \xi$ is $A_{n,2\xi-t}$. We have:

Proposition 11. - *The density of energy sent by 0 to the point* $A_{n,t}$, *taking into account the annihilation by the barrier, is :*

$$f_n(t) = \frac{1}{\sigma\sqrt{2\pi}}\left(\exp\left(-\frac{t^2}{2\sigma^2}\right) - \exp\left(-\frac{(2\xi - t)^2}{2\sigma^2}\right)\right), \textit{ for } t \le \xi \textit{, 0 if } t > \xi$$

with $\sigma = \sqrt{n}$.

Proof of Proposition 11

This is a mere rephrasing of the previous results, but we see that the function is simply the difference of two gaussian functions with same variance.

From Proposition 11, we easily obtain the profile of energy, on the vertical W_n, that is the position of the point of maximal energy:

Proposition 12. - *The point of maximal energy on the vertical* W_n *has coordinate:*

$$t_n \approx \frac{3\xi}{2} - \frac{1}{2}\sqrt{\xi^2 + 4n}$$

Proof of Proposition 12

Indeed, we have to find the maximum of the function f_n. But:

$$f_n'(t) = \frac{1}{\sigma\sqrt{2\pi}}\left(-\frac{t}{\sigma^2}\exp\left(-\frac{t^2}{2\sigma^2}\right) + \frac{t-2\xi}{\sigma^2}\exp\left(-\frac{(t-2\xi)^2}{2\sigma^2}\right)\right)$$

So, the condition $f_n' = 0$ is equivalent to:

$$\frac{t}{t-2\xi} = \exp\left(\frac{t^2}{2\sigma^2} - \frac{(t-2\xi)^2}{2\sigma^2}\right) = \exp\left(\frac{-2\xi(\xi-t)}{n}\right) \tag{1}$$

Consider first the function $h(t) = \frac{t}{t-2\xi}$; the derivative is $h'(t) = \frac{-2\xi}{(t-2\xi)^2} < 0$, so the function is decreasing, has the limit 1 at $-\infty$ and takes the value -1 for $t = \xi$. The function $g(t) = \exp\left(\frac{-2\xi(\xi-t)}{n}\right)$ is increasing, has limit 0 when $t \to -\infty$ and takes the value 1 for $t = \xi$.

Therefore, a unique solution t_n of equation (1) exists. When $n \to +\infty$, we have the rough estimate:

$$\frac{t}{t-2\xi} \sim 1 - \frac{2\xi(\xi-t)}{n}$$

that is:

$$t_n \approx \frac{3\xi}{2} - \frac{1}{2}\sqrt{\xi^2 + 4n}$$

which implies that $t_n \to -\infty$ and proves Proposition 12.

V. References

[Berry-Esseen]
https://en.wikipedia.org/wiki/Berry%E2%80%93Esseen_theorem

[Feller] W. Feller : An Introduction to Probability Theory and Its Applications, Vol. 1, 3rd Edition, Wiley series in Probabilities.

[Khinchine] A. Khinchine. "Über einen Satz der Wahrscheinlichkeitsrechnung", Fundamenta Mathematica, 6:9-20, 1924.

[Komatsu] Yûsaku Komatu, *Elementary inequalities for Mills' ratio*, Rep. Statist. Appl. Res. Un. Jap. Sci. Engrs. **4** (1955), 69–70. MR 0079844 (18,155f)

[Velenik] Y. Velenik, Université de Genève "Chapitres choisis de Théorie des Probabilités" : http://www.unige.ch/math/folks/velenik/papers/LN-CC1.pdf