Lecture 16: Simple Random Walk

In 1950 William Feller published *An Introduction to Probability Theory and Its Applications* [10]. According to Feller [11, p. vii], at the time "few mathematicians outside the Soviet Union recognized probability as a legitimate branch of mathematics." In 1957, he published a second edition, "which was in fact motivated principally by the unexpected discovery that [Chapter III's] enticing material could be treated by elementary methods." Here is an elementary treatment of some of these fun and possibly counterintuitive facts about random walks, the subject of Chapter III. It is based primarily on Feller [11, Chapter 3] and [12, Chapter 12], but I have tried to make it even easier to follow.

16.1 ⋆ What is the simple random walk?

Let $X_1, \dots, X_t, \dots$ be a sequence of independent Rademacher random variables,

$$X_t = \begin{cases} 1 & \text{with probability } 1/2 \\ -1 & \text{with probability } 1/2 \end{cases}$$

so

$$\boldsymbol{E}\, X_t = 0, \qquad \text{and} \qquad \boldsymbol{Var}\, X_t = 1 \qquad (t = 1, 2, \dots)$$

(Imagine a game played between Hetty and Taylor, in which a fair coin is tossed repeatedly. When Heads occurs, Hetty wins a dollar from Taylor, and when Tails occurs Taylor wins a dollar from Hetty. Then X_t is the change in Hetty's net winnings on the t^{th} coin toss.)

The index t indicates a point in time. Feller uses the term **epoch** to denote a particular moment t, and reserves the use of the word "time" to refer to a duration or interval of time, rather than a point in time, and I shall adhere to his convention. The set of epochs is the set

$$\mathbb{Z}_+ = \{0, 1, 2, \dots\}$$

of nonnegative integers. The epoch 0 is the moment before any coin toss.

For each t, define the running sums

$$S_t = X_1 + \cdots + X_t.$$

For convenience, we define $S_0 = 0$.

(The random variable S_t is Hetty's total net winnings at epoch t, that is, after t coin tosses.)

It follows that for each S_t,

$$\boldsymbol{E}\, S_t = 0 \qquad \text{and} \qquad \boldsymbol{Var}\, S_t = t.$$

The sequence of random variables $S_0, \dots, S_t, \dots$, $t \in \mathbb{Z}_+$ is a discrete-time stochastic process known as the **simple random walk on the integers**. It is both a martingale ($\boldsymbol{E}(S_{t+s} \mid S_t) = S_t$) and a stationary Markov chain (the distribution of $S_{t+s} \mid S_t = k_t, \dots, S_1 = k_1$ depends only on the value k_t).

16.1.1 Remark The walk $S_t = X_1 + \cdots + X_t$ can be "restarted" at any epoch n and it will have the same probabilistic properties. By this I mean that the process defined by

$$\hat{S}_t = S_{n+t} - S_n = X_{n+1} + \cdots + X_{n+t},$$

is also a simple random walk.

16.2 ⋆ Asymptotics

The Strong Law of Large Numbers tells us that

$$\frac{S_t}{t} \xrightarrow[t\to\infty]{} 0 \text{ a.s.},$$

and the Central Limit Theorem tells us that

$$\frac{S_t}{\sqrt{t}} \xrightarrow[t\to\infty]{\mathcal{D}} N(0,1).$$

Recall that the probability that the absolute value of a mean-zero Normal random variable exceeds its standard deviation is $2\big(1-\Phi(1)\big) = 0.317$, where Φ is the standard normal cdf. The standard deviation of S_t is $\sqrt{t}$, so there is about a two-thirds chance that S_t lies in the interval $[-\sqrt{t}, \sqrt{t}]$. See Figure 16.1. At each largish t, about two-thirds of the paths cross the vertical

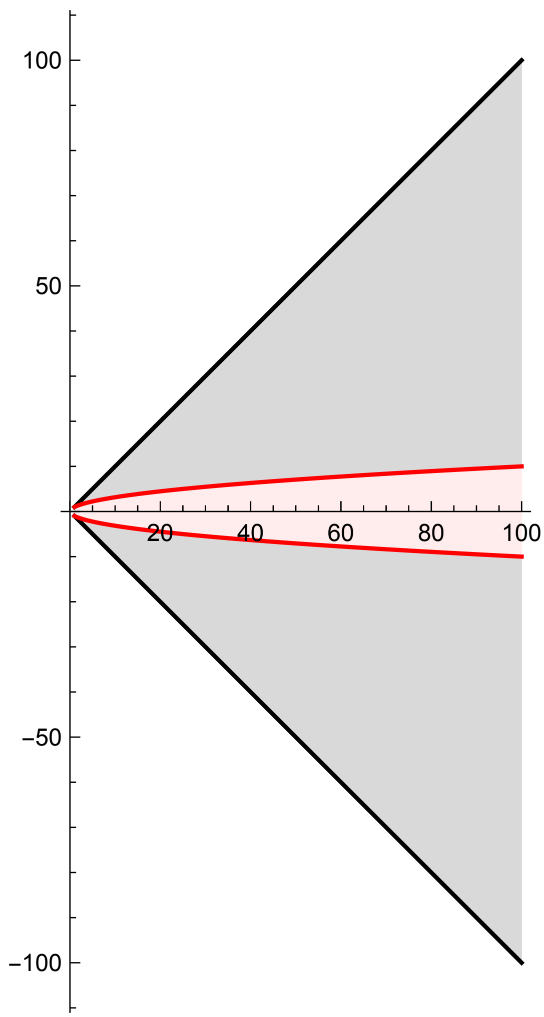


Figure 16.1. The areas bounded by $\pm t$ and by $\pm\sqrt{t}$.

line at t in the red area.

But the Strong Law of Large Numbers and the Central Limit Theorem are rather coarse and may mislead us about behavior of the random walk. They do not address interesting questions such as, Which values can the walk assume?, What are the waiting times between milestones?, or What does a "typical" path look like?

16.3 ⋆ Paths as the sample space

A natural way to think about the random walk is in term of **paths**. The outcome path $s = (s_1, s_2, \dots)$ can be identified with the sequence (t, s_t), $t = 0, 1, \dots$ of ordered pairs, or better yet with the graph of the piecewise linear function that connects the points (t, s_t). See Figure 16.2.

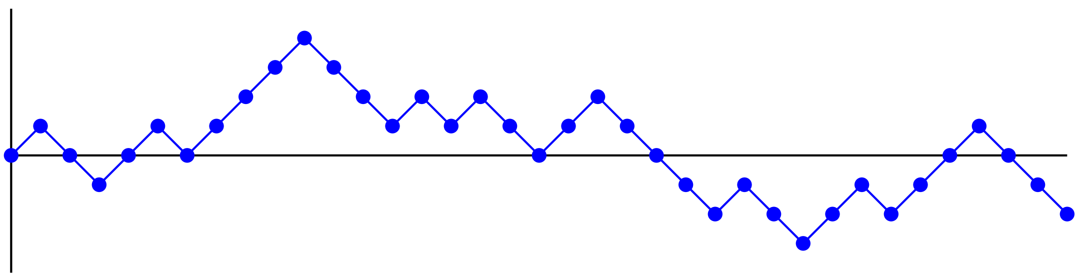


Figure 16.2. A sample path of a random walk.

There are infinitely many paths, but it is also convenient to refer to an initial segment of a path as a path. (Technically, the initial segment defines a set, or equivalence class, of paths that agree through some epoch.)

There are 2^t paths the walk may take through epoch t, and each one has equal probability, namely $1/2^t$.

Let us say that the **path s visits k at epoch t** if

$$s_t = k.$$

If there is a path s with $s_t = k$, we say that the **path s reaches (t, k)** and that (t, k) is **reachable** from the origin. More generally, if (t_0, k_0) and (t_1, k_1), where $t_1 > t_0$, are on the same path, then we say that **(t_1, k_1) is reachable from (t_0, k_0)**.

16.3.1 Characterization of reachable points

Which of the lattice points $(t, k) \in \mathbb{Z}_+ \times \mathbb{Z}$ can belong to a path? Or in other words, which points (t, k) are reachable? Not all lattice points are reachable. For instance, the points $(1, 0)$ and $(1, 2)$ are not reachable since S_1 is either 1 or -1.

16.3.1 Proposition (Criterion for reachability) *In order for (t, k) to be reachable, there must be nonnegative integers p and m, where p is the number of plus ones and m is the number of minus ones such that*

$$\begin{aligned} p + m = t \quad &\text{and} \quad p - m = k, \\ p = \frac{t+k}{2} \quad &\text{and} \quad m = \frac{t-k}{2}. \end{aligned} \tag{1}$$

Reachability implies that both $t + k$ and $t - k$ must be even, so that

t and k must have the same parity.

We must also have $t \geqslant |k|$. But those are the only restrictions.

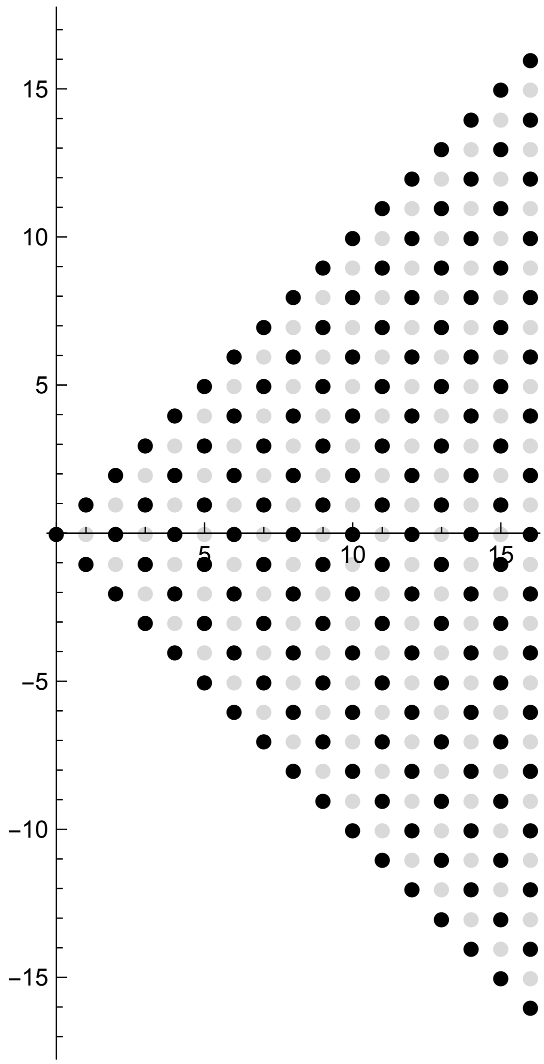


Figure 16.3. Reachable points are the big dots.

Many points can be reached by more than one path from the origin.

16.3.2 Definition *The number of initial segments of paths that reach the reachable point (t,k) is denoted*

$$N_{t,k}.$$

If (t,k) is not reachable, then $N_{t,k} = 0$.

16.3.3 Proposition (Number of paths that reach (t,k)) *If (t,k) is reachable, then*

$$N_{t,k} = \binom{t}{\frac{t+k}{2}} = \binom{t}{\frac{t-k}{2}}. \tag{2}$$

Proof: By Proposition 16.3.1, if (t,k) is reachable, there must be nonnegative integers p and m, where p is the number of plus ones and m is the number of minus ones such that (1) is satisfied.

Since the p (1)s and m (-1)s can be arranged in any order, there are

$$N_{t,k} = \binom{p+m}{p} = \binom{p+m}{m} = \binom{t}{\frac{t+k}{2}} = \binom{t}{\frac{t-k}{2}}$$

paths with this property. ∎

Since there are 2^t paths of length t, the probability is given by:

Define

$$p_{t,k} = P\left(S_t = k\right).$$

16.3.4 Corollary ($p_{t,k}$) *If (t,k) is reachable, then*

$$p_{t,k} = \binom{t}{\frac{t+k}{2}} 2^{-t}. \tag{3}$$

16.3.5 Corollary *If (t_1,k_1) is reachable from (t_0,k_0), then the number of sample paths connecting them is*

$$N_{t_1-t_0,k_1-k_0}. \tag{4}$$

16.3.2 The Reflection Principle

Feller referred to "elementary methods" that simplified the analysis of the simple random walk. The procedure is this: Treat paths as piecewise linear curves in the plane. Use the simple geometric operations of cutting, joining, sliding, rotating, and reflecting to create new paths. Use this technique to demonstrate the one-to-one or two-to-one correspondence between events (sets of paths). If we can find a one-to-one correspondence between a set we care about and a set we can easily count, then we can compute its probability.

The first example of this geometric manipulation approach is called the Reflection Principle.

16.3.6 The Reflection Principle *Let (t_1,k_1) be reachable from (t_0,k_0) and on the same side of the time axis. Then there is a one-to-one correspondence between the set of paths from (t_0,k_0) to (t_1,k_1) that meet (touch or cross) the time axis and the set of all paths from $(t_0,-k_0)$ to (t_1,k_1).*

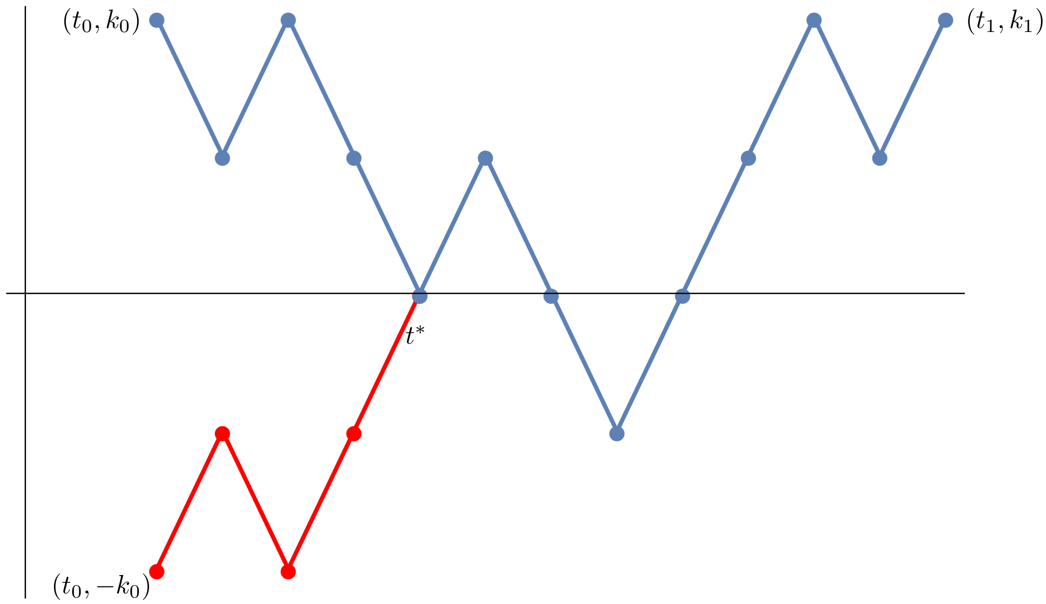


Figure 16.4. The red path is the reflection of the blue path up until the first epoch t^* where the blue path touches the time axis. This establishes a one-to-one correspondence between paths from $(t_0, -k_0)$ to (t_1, k_1) and paths from (t_0, k_0) to (t_1, k_1) that touch the time axis at some $t_0 < t < t_1$. q.e.d.

Proof: One picture is worth a thousand words, so Figure 16.4 should suffice for a proof. ∎

A consequence is the following.

16.3.7 The Ballot Theorem *If $k > 0$, then there are exactly*

$$\frac{k}{n} N_{n,k}$$

paths from the origin to (n, k) satisfying $s_t > 0$, $t = 1, \dots, n$.

Proof:

- If $s_t > 0$ for all $t = 1, \dots, n$, then $s_1 = 1$.
- By Corollary 16.3.5, the total number of paths from $(1, 1)$ to (t_1, k) is $N_{n-1,k-1}$.
- Some of these paths though touch the time axis, and when they do, they do not satisfy $s_t > 0$. How many of these paths touch the time axis? By the Reflection Principle, it is as many as the paths from $(1, -1)$ to (n, k), which by Corollary 16.3.5 is $N_{n-1,k+1}$.
- Thus the number of paths from $(1, 1)$ to (n, k) that do not touch the time axis is

$$N_{n-1,k-1} - N_{n-1,k+1}.$$

- Let p and m be as defined by (1). Then $p + m = n$, $p - m = k$, $n + k = 2p$, so the following "trite calculation," as Feller puts it, yields

$$\begin{aligned} N_{n-1,k-1} - N_{n-1,k+1} &= \binom{n-1}{(n+k-2)/2} - \binom{n-1}{(n+k)/2} \\ &= \binom{m+p-1}{p-1} - \binom{m+p-1}{p} \\ &= \frac{(m+p-1)!}{m!(p-1)!} - \frac{(m+p-1)!}{p!(m-1)!} \\ &= \frac{p(m+p-1)!}{m!p!} - \frac{m(m+p-1)!}{p!m!} \\ &= (p-m)\frac{(m+p-1)!}{m!p!} \\ &= \frac{p-m}{p+m}\frac{(m+p)!}{m!p!} \\ &= \frac{k}{n} N_{n,k}. \end{aligned}$$

∎

Why is this called the Ballot Theorem?

16.3.8 The Ballot Theorem, version 2 *Suppose an election with n ballots cast has one candidate winning by k votes. Count the votes in random order. The probability the winning candidate always leads is*

$$\frac{k}{n}.$$

16.3.9 The Ballot Theorem, version 3 *Suppose an election has one candidate getting p votes and the other getting m votes with $p > m$. Count the votes in random order. The probability the winning candidate always leads is*

$$\frac{p-m}{p+m}.$$

16.4 ⋆ Returns to zero

16.4.1 Definition *We say that the walk* ***equalizes*** *or* ***returns to zero*** *at epoch t if $S_t = 0$.*

Epoch t must be even for equalization to occur, so let $t = 2m$. The number of paths from the origin to $(2m, 0)$ is $N_{2m,0}$, so the probability u_{2m} defined by

$$u_{2m} = P\left(S_{2m} = 0\right),$$

satisfies

$$u_{2m} = \frac{N_{2m,0}}{2^{2m}} = \binom{2m}{m}\frac{1}{2^{2m}} \qquad (m \geqslant 0). \tag{5}$$

Recall

16.4.2 Proposition (Stirling's formula)

$$n! = e^{-n}n^n\sqrt{2\pi n}(1+\varepsilon_n)$$

where $\varepsilon_n \to 0$ as $n \to \infty$.

For a proof, see, e.g., Feller [11, p. 52].

Stirling's formula applied to (5) implies that

$$u_{2m} \sim \frac{1}{\sqrt{\pi m}}, \tag{6}$$

where the notation $a_m \sim b_m$ means $a_m/b_m \to 1$ as $m \to \infty$.

16.5 ⋆ The Main Lemma

The next application of the geometric manipulation of paths approach is to prove the following mildly surprising result. (It is the distillation of Lemma 3.1 and the discussion following on pp. 76–77, and problem 3.10.7, p. 96 in Feller [11].)

16.5.1 Main Lemma *The following are equal:*

$$P(S_{2m} = 0) \ \ (= u_{2m}) \tag{7}$$
$$P(S_1 \neq 0, \dots, S_{2m} \neq 0) \tag{8}$$
$$P(S_1 \geqslant 0, \dots, S_{2m} \geqslant 0) \tag{9}$$
$$P(S_1 \leqslant 0, \dots, S_{2m} \leqslant 0) \tag{10}$$
$$2P(S_1 > 0, \dots, S_{2m} > 0) \tag{11}$$
$$2P(S_1 < 0, \dots, S_{2m} < 0) \tag{12}$$

Proof: Start with the easy cases.

- By symmetry, (9) = (10) and (11) = (12).
- In order to have $(S_1 \neq 0, \dots, S_{2m} \neq 0)$, either $(S_1 > 0, \dots, S_{2m} > 0)$ or $(S_1 < 0, \dots, S_{2m} < 0)$. Both are equally likely. So (8) = (11) = (12).

Let

$$\begin{aligned} &\mathcal{Z}_t \text{ denote the set of paths satisfying } s_t = 0, \\ &\mathcal{P}_t \text{ denote the set of paths satisfying } (s_1 > 0, \dots, s_t > 0), \\ &\mathcal{N}_t \text{ denote the set of paths satisfying } (s_1 \geqslant 0, \dots, s_t \geqslant 0). \end{aligned}$$

(The mnemonic is zero, positive, and nonnegative.)

16.5.2 Lemma *There is a one-to-one correspondence between $\mathcal{P}_{2m}$ and $\mathcal{N}_{2m-1}$:*

Proof: Every path s in $\mathcal{P}_{2m}$ passes through $(1,1)$ and satisfies $s_t \geqslant 1$ for $t = 1, \dots, 2m$, so shifting the origin to $(1,1)$ creates a path s' of length $2m-1$ that satisfies $s'_t \geqslant 0$, $t = 1, \dots, 2m-1$. That is, $s' \in \mathcal{N}_{2m-1}$. See Figure 16.5. ∎

- Lemma 16.5.2 establishes a one-to-one correspondence between $\mathcal{P}_{2m}$ and $\mathcal{N}_{2m-1}$, so

$$|\mathcal{N}_{2m-1}| = |\mathcal{P}_{2m}|.$$

Thus

$$P(\mathcal{N}_{2m-1}) = \frac{|\mathcal{N}_{2m-1}|}{2^{2m-1}} = 2\frac{|\mathcal{N}_{2m-1}|}{2^{2m}} = 2\frac{|\mathcal{P}_{2m}|}{2^{2m}} = 2P(\mathcal{P}_{2m}).$$

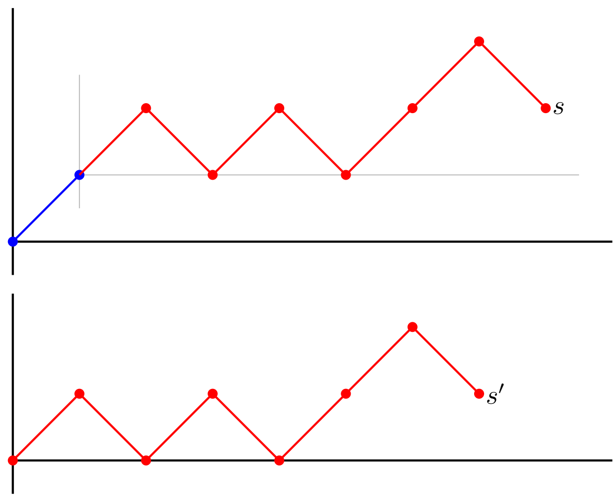


Figure 16.5. The paths $s \in \mathcal{P}_{2m}$ and $s' \in \mathcal{N}_{2m-1}$.

Since $2m-1$ is odd, and equalization occurs only in even epochs, we must have $s'_{2m-1} > 0$ for any $s' \in \mathcal{N}_{2m-1}$. There are two possible continuations of s' and both of them will satisfy $s'_{2m} \geqslant 0$. That is,

$$|\mathcal{N}_{2m}| = 2|\mathcal{N}_{2m-1}|.$$

Thus

$$P(\mathcal{N}_{2m}) = 2P(\mathcal{N}_{2m-1}) = 2P(\mathcal{P}_{2m}).$$

- So far we have shown (8) = (9) = (10) = (11) = (12).
- We complete the Main Lemma by showing (7) = (9), or

$$P\left(S_{2m} = 0\right) = P\left(S_1 \geqslant 0, \dots, S_{2m} \geqslant 0\right).$$

We shall establish this with the following lemma. Feller attributes this construction used in the proof to Edward Nelson.

16.5.3 Nelson's Lemma *There is a one-to-one correspondence between $\mathcal{Z}_{2m}$ and $\mathcal{N}_{2m}$. Moreover, a path in $\mathcal{Z}_{2m}$ with minimum value $-k$ corresponds to a path in $\mathcal{N}_{2m}$ with terminal value $2k$.*

Proof: Let s be a path in $\mathcal{Z}_{2m}$. It achieves a minimum value $-k^* \leqslant 0$ at some $t \leqslant 2m$, perhaps more than once. Let t^* be the smallest t for which $s_t = -k^*$.

If s also belongs to $\mathcal{N}_{2m}$, that is, if $s_t \geqslant 0$ for all $t = 0, \dots, 2m$, then $k^* = 0$ and $t^* = 0$, and we leave the path alone. If s does not belongs to $\mathcal{N}_{2m}$, that is, if $s_t < 0$ for some $0 < t < 2m$, then $k^* > 0$ and $0 < t^* < 2m$. See Figure 16.6.

We create a new path s' in $\mathcal{N}_{2m}$ as follows: Take the path segment from $(0, 0)$ to $(t^*, -k^*)$, and reflect it about the vertical line $t = t^*$. Slide this reversed segment to the point $(2m, 0)$. See Figure 16.7.

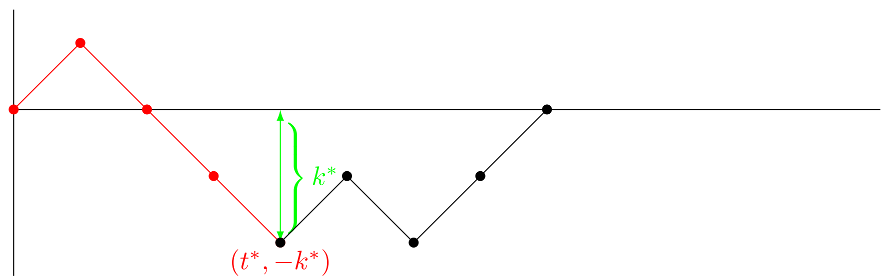


Figure 16.6. The path $s \in \mathcal{Z}_{2m}$. Epoch t^* is the first epoch at which the minimum $-k^*$ occurs.

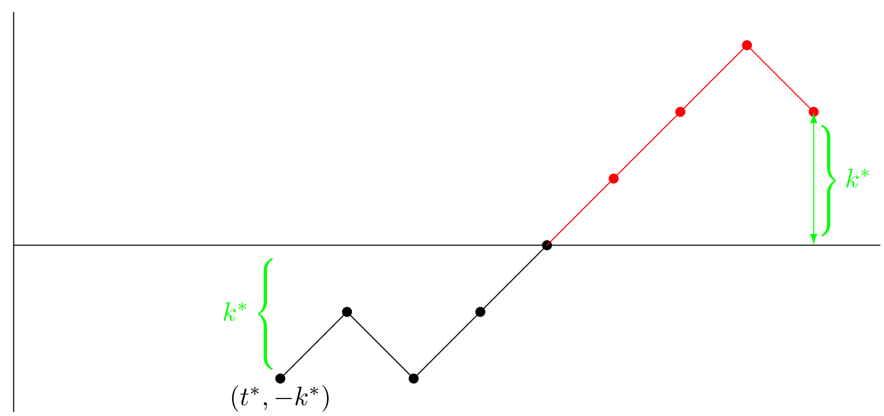


Figure 16.7. Reflect the initial segment around $t = t^*$, and slide it to the end.

Now slide the whole thing so that $(t^*, -k^*)$ becomes the new origin. The path now ends at $(2m, 2k^*)$, where $k^* > 0$. See Figure 16.8.

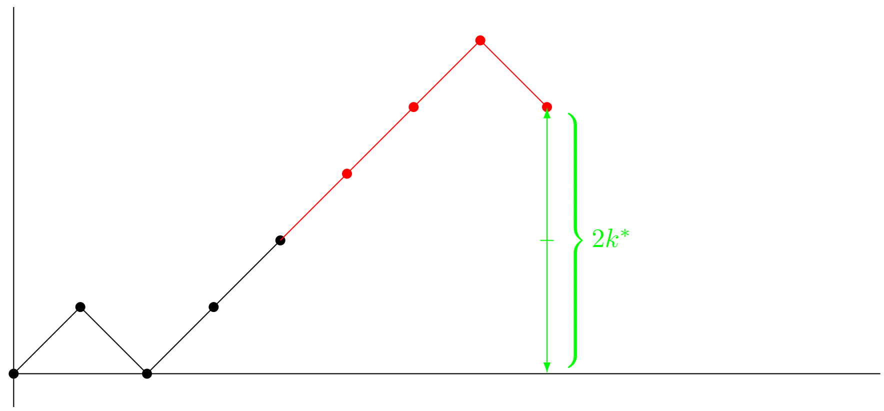


Figure 16.8. Now slide $(t^*, -k^*)$ to the origin to get the path $s' \in \mathcal{N}_{2m}$.

This process is invertible: Let s be a path in $\mathcal{N}_{2m}$. If $S_{2m} = 0$, leave it alone. If $s_{2m} > 0$, since s_{2m} is even, write $s_{2m} = 2\bar{k} > 0$. Let $\bar{t}$ be the last time $s_t = \bar{k}$. See Figure 16.9.

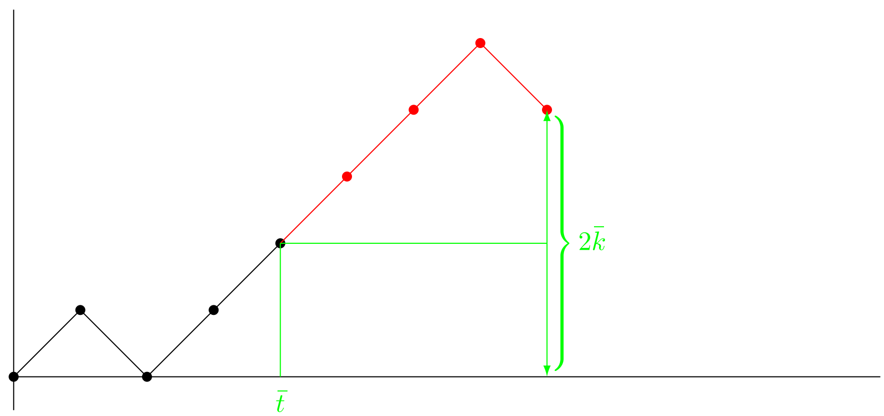


Figure 16.9. The path $s \in \mathcal{N}_{2m}$ satisfies $s_{2m} = 2\bar{k} > 0$. Epoch $\bar{t}$ is the last epoch for which $s_t = k$.

Take the segment of the path from $(\bar{t}, \bar{k})$ to $(2m, 2\bar{k})$, reflect it about the vertical line $t = \bar{t}$, slide it to the origin, (so it juts out up and to the left). See Figure 16.10.

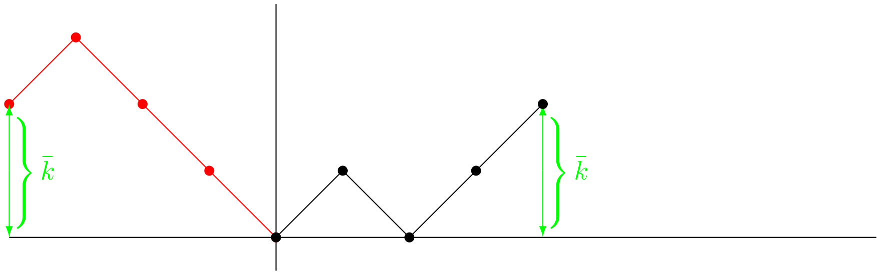


Figure 16.10. Take the segment of the path from $(\bar{t}, \bar{k})$ to $(2m, 2\bar{k})$, reflect it about the vertical line $t = \bar{t}$, slide it to the origin.

Now make the beginning the new origin. See Figure 16.11. This new path s'' satisfies $s_{2m} = 0$, and has a strictly negative minimum.

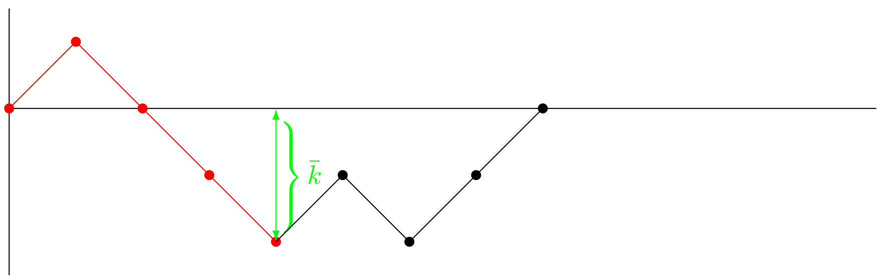


Figure 16.11. The path $s'' \in \mathcal{Z}_{2m}$.

In fact the procedure above inverts the first procedure. This establishes a one-to-one correspondence between $\mathcal{Z}_{2m}$ and $\mathcal{N}_{2m}$.

∎

- The proof of the Main Lemma is now finished.

∎

16.6 ⋆ First return to zero

16.6.1 Definition *The* ***first return to zero*** *happens at epoch $t = 2m$ if $s_1 \neq 0$, ...$s_{2m-1} \neq 0$ and $s_{2m} = s_t = 0$. Let $f_t = f_{2m}$ denote the probability of this event, and define $f_0 = 0$.*

The next result is equation (3.7), p. 78, [11].

16.6.2 Corollary *The explicit formula for f_{2m} is*

$$f_{2m} = u_{2m-2} - u_{2m} = \frac{1}{2m-1} u_{2m} = \frac{1}{2m-1} \binom{2m}{m} \frac{1}{2^{2m}} \qquad (m = 1, 2, \dots). \tag{13}$$

Proof: The event that the first return to zero occurs at epoch $2m$ is

$$(S_1 \neq 0, \dots, S_{2m-2} \neq 0, S_{2m} = 0) = (S_1 \neq 0, \dots, S_{2m-2} \neq 0) \setminus (S_1 \neq 0, \dots, S_{2m} \neq 0)$$

Since $(S_1 \neq 0, \dots, S_{2m} \neq 0) \subset (S_1 \neq 0, \dots, S_{2m-2} \neq 0)$

$$P\left((S_1 \neq 0, \dots, S_{2m-2} \neq 0) \setminus (S_1 \neq 0, \dots, S_{2m} \neq 0)\right) = P\left(S_1 \neq 0, \dots, S_{2m-2} \neq 0\right) - P\left(S_1 \neq 0, \dots, S_{2m} \neq 0\right),$$

which by the Main Lemma is

$$u_{2m-2} - u_{2m}.$$

Equation (5) implies

$$u_{2m-2} = \frac{(2m-2)!}{(m-1)!(m-1)!}\frac{1}{2^{2m-2}} = \left(\frac{4m^2}{2m(2m-1)}\right)\frac{(2m)!}{m!m!}\frac{1}{2^{2m}}$$

so $u_{2m-2} - u_{2m} = \left(\frac{2m}{2m-1} - 1\right)u_{2m}$, which implies the second and third equalities of (13). ∎

16.6.3 Corollary *With probability 1, the random walk returns to zero. Consequently, with probability 1 it returns to zero infinitely often.*

Proof: The event that the walk returns to zero is the union over m of the disjoint events that the first return occurs at epoch $2m$. Its probability is just the sum of the first return probabilities. According to (13), we have the telescoping series

$$\sum_{m=1}^{\infty} f_{2m} = \sum_{m=1}^{\infty} (u_{2(m-1)} - u_{2m}) = u_0 = 1. \tag{14}$$

∎

While the walk is certain to return to zero again and again, you wouldn't want to hold your breath waiting for it. Cf. Theorem 1, p. 395, [12].

16.6.4 Proposition *(Waiting time for first return to zero) Let W denote the epoch of the first return to zero. Then*

$$\boldsymbol{E}\, W = \infty.$$

Proof: From Corollary 16.6.2,

$$\boldsymbol{E}\, W = \sum_{m=1}^{\infty} 2m\, f_{2m} = \sum_{m=1}^{\infty} \frac{2m}{2m-1} u_{2m}. \tag{15}$$

From (6),

$$\frac{\frac{2m}{2m-1}u_{2m}}{\frac{1}{\sqrt{\pi m}}} \xrightarrow[m\to\infty]{} 1.$$

But

$$\sum_{m=1}^{\infty} \frac{1}{\sqrt{\pi m}} \geqslant \frac{1}{\sqrt{\pi}} \sum_{m=1}^{\infty} \frac{1}{m} \to \infty.$$

So by the Limit Comparison Test [2, Theorem 10.9, p. 396], the series (15) also diverges to ∞. ∎

16.7 ⋆ Recurrence

16.7.1 Definition *The value k is **recurrent** if*

$$P\left(S_t = k \text{ infinitely often}\right) = 1.$$

Corollary 16.6.3 proved that 0 was a recurrent value. An astonishingly simple consequence of this that *every* value is recurrent.

16.7.2 Corollary *For every integer k, with probability 1 the random walk visits k. Consequently, with probability 1 it visits k infinitely often.*

Proof: The result relies heavily on symmetry, and so does this proof.

- For every k, the probability that the walk visits k is greater than zero. Indeed, for $k \geqslant 0$, we have $P\left(S_k = k\right) = 2^{-k} > 0$; and for $k \leqslant 0$, we have $P\left(S_k = k\right) = 2^k > 0$.
- Once the walk visits k, the probability that it later visits zero must be one, otherwise the probability of visiting zero infinitely often could not be one.
- *But the probability of reaching* 0 *from k is the same as reaching k from* 0*!*
- Therefore the probability the walk visits k is one.
- Once at k, the probability of visiting k again is the same as the probability of revisiting zero from the origin, which is one. Therefore k is recurrent.

∎

This fact is another illustration of the difference between impossibility and probability zero. The path $s_t = t$ for all t never returns to zero or anything else, and it is certainly a possible path. But it has probability zero of occurring.

16.8 ⋆ The Arc Sine Law

The next result appears as [11, Theorem 1, p. 79].

For each m, define the random variable

$$\begin{aligned} L_{2m} &= \text{the epoch of the last visit to zero, up to and including epoch } 2m \\ &= \max\{t : 0 \leqslant t \leqslant 2m \;\&\; S_t = 0\}. \end{aligned}$$

Note that $L_{2m} = 0$ and $L_{2m} = 2m$ are allowed. For convenience, define

$$\alpha_{2k,2m} = P\left(L_{2m} = 2k\right).$$

16.8.1 The Arc Sine Law for Last Returns *The probability mass function for L_{2m} is given by*

$$P\left(L_{2m} = 2k\right) = \alpha_{2k,2m} = u_{2k}u_{2(m-k)}. \qquad (k = 0, \dots, m) \tag{16}$$

Proof: The event $(L_{2m} = 2k)$ can be written

$$\underbrace{S_{2k} = 0}_{A},\ \underbrace{S_{2k+2} \neq 0, \dots, S_{2m} \neq 0}_{B}.$$

Recall that $P(AB) = P(B|A)P(A)$. Now $P(A)$ is just u_{2k} and $P(B|A)$ is just the probability that starting at 0, the next $2(m-k)$ values of S_t are nonzero, which is the same as the probability that $S_t \neq 0$, $t = 1, \dots, 2(m-k)$. By the Main Lemma 16.5.1 this is equal to $u_{2(m-k)}$. ∎

Why is this called the Arc Sine Law?: From (6), $u_{2k} \sim \frac{1}{\sqrt{\pi k}}$, so for large enough k, m,

$$\alpha_{2k,2m} = u_{2k}u_{2(m-k)} \approx \frac{1}{\pi\sqrt{k(m-k)}} = \frac{1}{m}\frac{1}{\pi\sqrt{\frac{k}{m}\left(1-\frac{k}{m}\right)}}. \tag{17}$$

See Figure 16.12. As you can see from the figure, the approximation is rather good for even modest

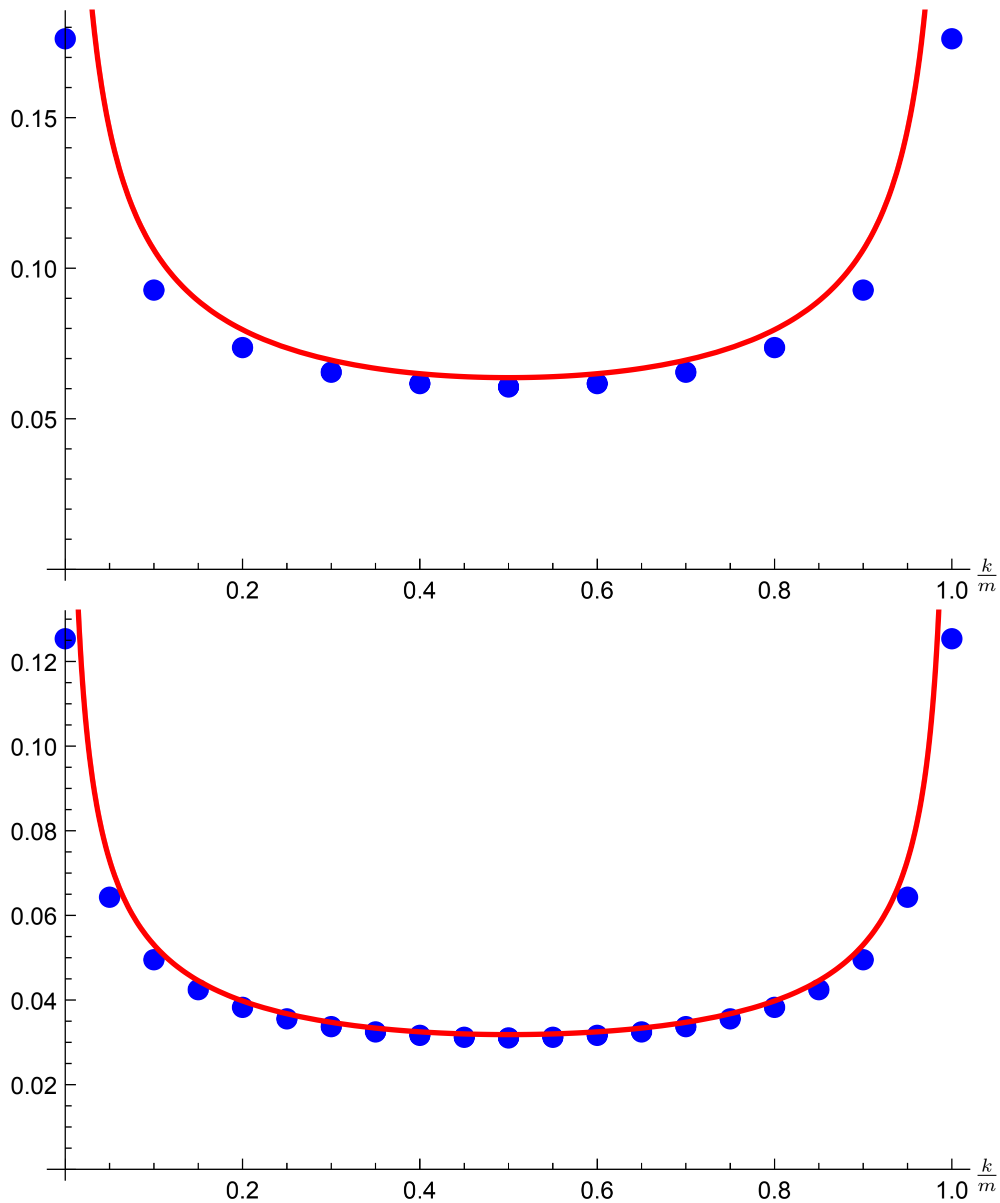


Figure 16.12. Plots of the points $(k/m, \alpha_{2k,2m})$, $k = 0, \ldots, m$ and the function $f(x) = \frac{1}{m}\frac{1}{\pi\sqrt{x(1-x)}}$, $x \in [0,1]$ for $m = 10, 20$.

values of k and m, and the highest probabilities are for $k = 0$ and $k = 2m$, with the minimum occurring around m.

The function

$$f(x) = \frac{1}{\pi\sqrt{x(1-x)}}.$$

is a probability density on the unit interval, and the cumulative distribution function involves the arc sine function: For $0 \leqslant \rho \leqslant 1$,

$$\int_0^\rho f(x)\,dx = \frac{2}{\pi}\arcsin\sqrt{\rho}.$$

So for every $0 < \rho < 1$, for m large enough,

$$P\left(L_{2m} \leqslant \rho 2m\right) = \sum_{k<\rho m} \alpha_{2k,2m} \approx \int_0^\rho f(x)\,dx = \frac{2}{\pi}\arcsin\sqrt{\rho}. \tag{18}$$

The Arc Sine Law has the following remarkable consequence, [11, p. 78].

16.8.2 Corollary *For every m,*

$$P\left(\textit{the latest return to } 0 \textit{ through epoch } 2m \textit{ occurs no later than epoch } m\right) = P\left(L_{2m} \leqslant m\right) = \frac{1}{2}.$$

In other words, the probability that no equalization has occurred in the last half of the history is 1/2 regardless of the length of the history.

Proof: For t even, $P\left(L_{2m} = t\right) = u_t u_{2m-t}$, which is symmetric about m, so m is the median value of L_{2m}. ∎

16.9 ⋆ Dual walks

We would like to know about the probabilities of visiting points other than zero. To do that, we shall make use of the **dual** of a random walk. Recall that the walk S is given by

$$S_t = X_1 + \cdots X_t,$$

where the X_ts are independent and identically distributed Rademacher random variables.

16.9.1 Definition *Fix a length n, and create a new random walk S^* of length n by reversing the order of the X_ts. That is, define*

$$X_1^* = X_n,\ \ldots, X_n^* = X_1,$$

and

$$S_t^* = X_1^* + \cdots + X_t^* = X_n + \cdots + X_{n-t+1} = S_n - S_{n-t}, \qquad (t = 1, \ldots, n). \tag{19}$$

*This walk is called the **dual** of S.*

Since S_n and S_n^* are sums of the same X_ts, they have the same terminal points, that is, $S_n = S_n^*$. More importantly, *every event related to S has a dual event related to S^* that has same the probability.* Given a path s for S, the dual path s^* for S^* is gotten by rotating the path s one hundred eighty degrees around the origin (time reversal), so the left endpoint has a negative time coordinate, and then sliding the left endpoint to the origin to get s^*. See Figures 16.13 and 16.14.

For instance, it follows from (19) that

$$P\left(S_n = k, S_1 > 0, \ldots, S_{n-1} > 0\right) = P\left(S_n^* = k, S_n^* > S_1^*, \ldots, S_n^* > S_{n-1}^*\right). \tag{20}$$

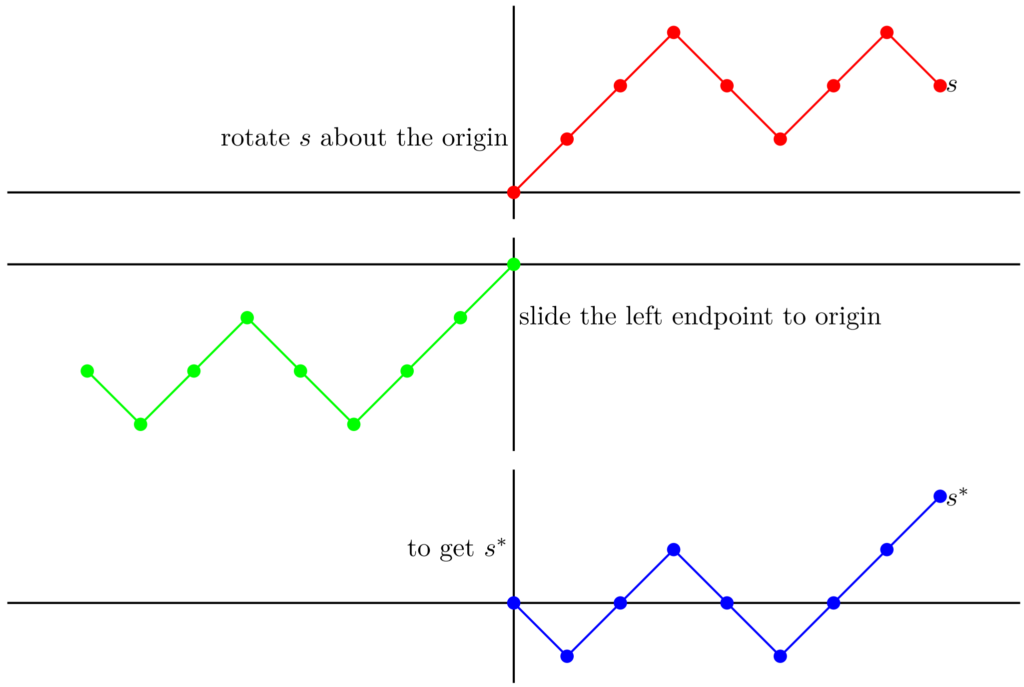


Figure 16.13. Transforming s to s^*.

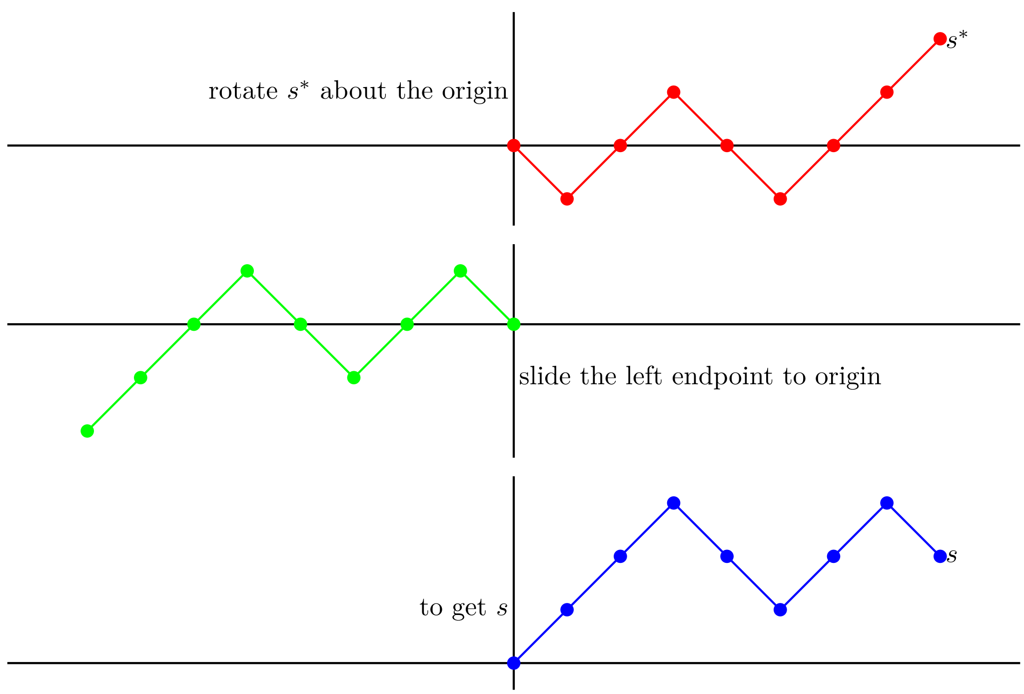


Figure 16.14. Transforming s^* back to s by the same method, $(s^*)^* = s$. (This figure also demonstrates a one-to-one correspondence that proves (20).)

16.10 ⋆ First visits

This argument is taken from Feller [11, pp. 92–93]. We know from Proposition 16.3.3 that the probability that $S_t = k$ is

$$P(S_t = k) = \frac{N_{t,k}}{2^t} = \binom{t}{\frac{t-k}{2}}\frac{1}{2^t},$$

provided (t,k) is reachable from the origin. (Otherwise, it is zero.)

Let k be greater than zero. Assume that (n,k) is reachable from the origin. That is, $n-k \geqslant 0$ and $n-k$ is even. What is the probability that the first visit to k happens at epoch n? This is the probability of the event

$$(S_1 < S_n, \dots, S_{n-1} < S_n,\ S_n = k). \tag{21}$$

Now consider the dual walk S^*. In terms of S^*, it follows from (19) that the event (21) is the same as

$$\left(S_1^* > 0, \dots, S_{n-1}^* > 0,\ S_n^* = k\right). \tag{22}$$

We already know the probability of this dual event. There are 2^n paths of length n, and according to the Ballot Theorem 16.3.7, $\frac{k}{n} N_{n,k}$ of these paths satisfy $s_t^* > 0$ for $t = 1, \dots, n$. Thus the probability of event (22), and hence of event (21) is

$$P(\text{the first visit to } k \text{ occurs at epoch } n) = \frac{k}{n}\binom{n}{\frac{n-k}{2}}\frac{1}{2^n}, \tag{23}$$

provided $n-k$ is a nonnegative even integer. (Otherwise, it is zero.)

If $n-k$ is a nonnegative even integer, write $n = 2m+k$. It follows from (23) and the fact that k is recurrent that for each $k \geqslant 1$,

$$\sum_{m=0}^{\infty} \frac{k}{2m+k}\binom{2m+k}{m}\frac{1}{2^{2m+k}} = 1, \tag{24}$$

but don't ask me to prove it directly.

Aside: You may have noticed a similarity between the values of f_{2m} given in (13) and the terms in (24) for $k = 1$. (When I first noticed it, it kept me awake until I could verify the following equalities.)

$$\begin{aligned}
1 &= \sum_{m=1}^{\infty} f_{2m} && \text{equation (14)}\\
&= \sum_{m=1}^{\infty} \frac{1}{2m-1}\binom{2m}{m}\frac{1}{2^{2m}} && \text{equation (13)}\\
&= \sum_{n=0}^{\infty} \frac{1}{2n+1}\binom{2n+2}{n+1}\frac{1}{2^{2n+2}} && \text{substitute } n = m-1\\
&= \sum_{n=0}^{\infty} \frac{1}{2n+1}\binom{2n+1}{n}\frac{1}{2^{2n+1}} && \text{since } \binom{2n+2}{n+1} = 2\binom{2n+1}{n}\\
&= 1 && \text{from (24) with } k = 1.
\end{aligned}$$

16.11 ⋆ The number of visits to k before equalization

The following highly counterintuitive result is Example (b) on p. 395 of Feller [12].

Let k be nonzero, and let

$$M_k = \text{the count of epochs } n \text{ for which } S_n = k \text{ before the first return to zero.}$$

16.11.1 Proposition *For each k,*

$$\boldsymbol{E}\, M_k = 1.$$

Proof: Since k and $-k$ are symmetric, it suffices to consider $k > 0$. Let V_n^k be the event that a visit to k occurs at epoch n before the first return to zero. That is,

$$V_n^k = (S_n = k, S_1 > 0, \dots, S_{n-1} > 0)\,.$$

Then

$$M_k = \sum_{n=1}^{\infty} \mathbf{1}_{V_n^k},$$

where $\mathbf{1}_{V_n^k}$ is the indicator function of the event V_n^k.

By the Monotone Convergence Theorem (oops, I never told you about that one, but see, for instance, [1, Theorem 11.19, p. 414]) we have

$$\begin{aligned}\boldsymbol{E}\, M_k &= \sum_{n=1}^{\infty} \boldsymbol{E}\, \mathbf{1}_{V_n^k} \\ &= \sum_{n=1}^{\infty} P(V_n^k)\end{aligned}$$

Now we need to find $P(V_n^k)$. Consider the dual walk $S_1^*, \dots, S_n^*$. By (20), we have

$$\begin{aligned}P\left(S_n = k, S_1 > 0, \dots, S_{n-1} > 0\right) &= P\left(S_n^* = k, S_n^* > S_1^*, \dots, S_n^* > S_{n-1}^*\right) \\ &= P\left(\text{first visit to } k \text{ occurs at epoch } n\right)\end{aligned}$$

Therefore

$$\begin{aligned}\boldsymbol{E}\, M_k &= \sum_{n=1}^{\infty} P(V_n^k) \\ &= \sum_{n=1}^{\infty} P\left(\text{first visit to } k \text{ occurs at epoch } n\right) \\ &= P\left(\text{walk visits } k\right) \\ &= 1.\end{aligned}$$

The last equality holds because k is recurrent. ∎

16.12 ⋆ Sign changes

There is a **sign change** at epoch t if S_{t-1} and S_{t+1} have opposite signs. This requires that $S_t = 0$, and that t be even.

16.12.1 Theorem *Let $t = 2m + 1$ be odd. Then*

$$P\left(\textit{there are exactly } c \textit{ sign changes before epoch } t\right) = 2P\left(S_t = 2c + 1\right). \tag{25}$$

Proof: To save space, let

$$\mathcal{C}_{t,c} = (\text{there are exactly } c \text{ sign changes before epoch } t)\,.$$

Now

$$P(\mathcal{C}_{t,c}) = P(\mathcal{C}_{t,c} \mid S_1 = 1)P(S_1 = 1) + P(\mathcal{C}_{t,c} \mid S_1 = -1)P(S_1 = -1) = P(\mathcal{C}_{t,c} \mid S_1 = 1),$$

where the last equality follows from symmetry. That is, the probability of $\mathcal{C}_{t,c}$ is independent of the value of S_1, so we may assume that $S_1 = 1$.

Now $P(\mathcal{C}_{t,c} \mid S_1 = 1)$ is the number of paths starting at $(1, 1)$ that have exactly c sign changes before epoch $t = 2m+1$ divided by the number of paths starting at epoch 1 and ending at epoch $t = 2m + 1$. Thus the theorem reduces to the following Lemma. ∎

16.12.2 Lemma *For every odd $t = 2m + 1$, there is a one-to-one correspondence between the sets of paths $\{s : s_1 = 1$ and s has exactly c sign changes before $t\}$ and $\{s : s_t = 2c + 1\}$.*

I don't have time to write up the proof of the lemma, but you can find it in Feller [11, Section III.5, pp. 84–86].

16.12.3 Corollary *The probability of c sign changes decreases with c. Consequently the most likely number of sign changes is zero!*

16.13 ⋆ More remarkable facts

Blackwell, Deuel, and Freedman [7] discovered the following theorem while validating some code for an IBM 7090 computation.

16.13.1 Theorem *Let $V_{m,n}$ be the event that there exists t satisfying $m \leqslant t < m + n$ and $S_{2t} = 0$. That is, $V_{m,n}$ is the event that there is an equalization between epochs $2m - 1$ and $2(m + n) - 1$. Then for all $m, n \geqslant 1$,*

$$P(V_{m,n}) + P(V_{n,m}) = 1.$$

Note that the Corollary 16.8.2 of the Arcsine Law can be rewritten as the special case $P(V_{m,m}) = 1/2$.

16.14 ⋆ Asymmetry

The remainder of this discussion relies heavily on the exposition by Frederick Mosteller [13, pp. 51–55].

The preceding results made heavy use of the symmetry that arose from the fact that upticks and downticks were exactly equally likely. What happens when that is not the case?

Let $X_1, \dots, X_t, \dots$ be a sequence of independent Rademacher(p) random variables,

$$X_t = \begin{cases} 1 & \text{with probability } p \\ -1 & \text{with probability } 1 - p \end{cases}$$

so

$$\boldsymbol{E}\, X_t = 2p - 1, \qquad \text{and} \qquad \boldsymbol{Var}\, X_t = 4p(1 - p) \qquad (t = 1, 2, \dots)$$

It is convenient to consider starting walks at arbitrary integers, so let

$$S_0 = m, \qquad S_t = S_0 + X_1 + \cdots + X_t \qquad (t \geqslant 1)$$

denote the **asymmetric random walk starting at m with uptick probability p**. It is no longer a martingale, but it is a stationary Markov chain.

16.15 ⋆ Reaching zero

For a symmetric random walk, every state is recurrent. This fails for the asymmetric random walk. Let's calculate the probability z_m of reaching zero starting at $m > 0$.

Let's start with $S_0 = m$. In order to reach 0 from m, you must first reach $m-1$. Then from $m-1$, you must reach $m-2$, etc., all the way to reaching 0 from 1. Each of these steps looks exactly like the last. I also claim that the independence of the X_ts means that the probability of all of these steps happening is the product of their probabilities.[1] Thus

$$z_m = z_1^m, \qquad (m > 1).$$

Now the trick is to calculate z_1. Well, starting at 1, with probability $1-p$ we reach 0 on the first step. With probability p we reach 2, and then with probability $z_2 = z_1^2$ we reach 0. Thus

$$z_1 = 1 - p + pz_1^2.$$

This is a quadratic that has two solutions,

$$z_1 = 1, \qquad z_1 = \frac{1-p}{p}.$$

This makes sense, because z_1 really depends on p. For $p = 0$ (never gain), clearly $z_1 = 1$. And when $p = 1$ (never lose), $z_1 = 0$. When $p = 1/2$ both roots agree and $z_1 = 1$.

Figure 16.15 shows a plot of 1 and $(1-p)/p$ against p. These are the candidates for $z_1(p)$. We know $z_1(p)$ at three points $p = 0, 1/2, 1$. So if $z_1(p)$ is a continuous function, we must have

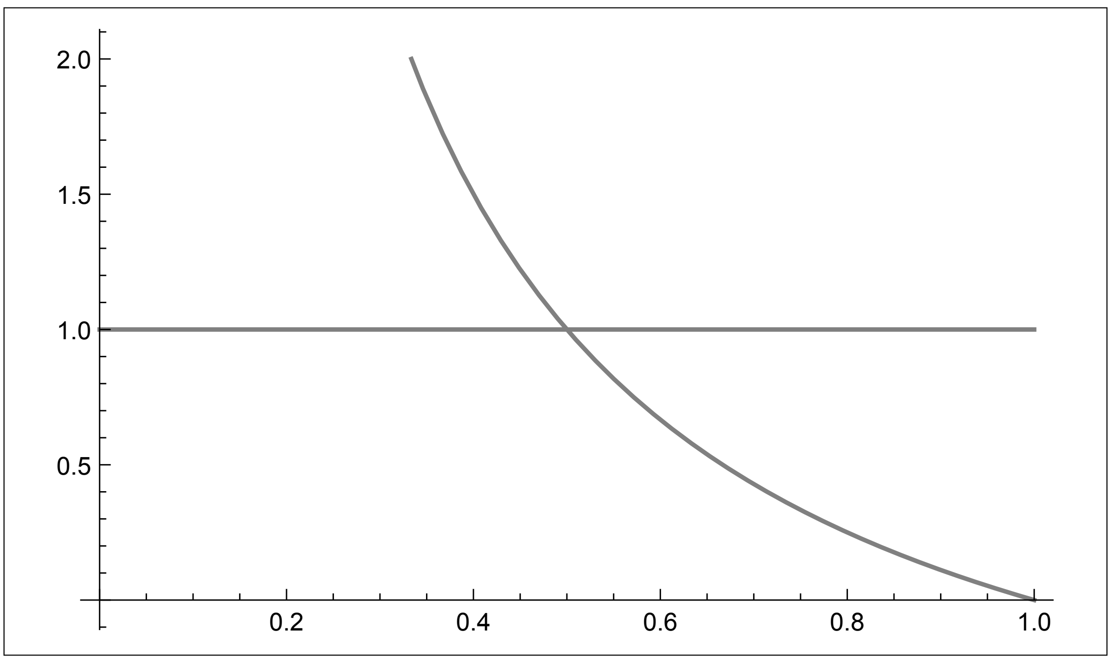


Figure 16.15. Plot of 1 and $(1-p)/p$ against p.

$$z_1 = \begin{cases} 1 & p \leqslant 1/2 \\ \frac{1-p}{p} & p \geqslant 12 \end{cases}$$

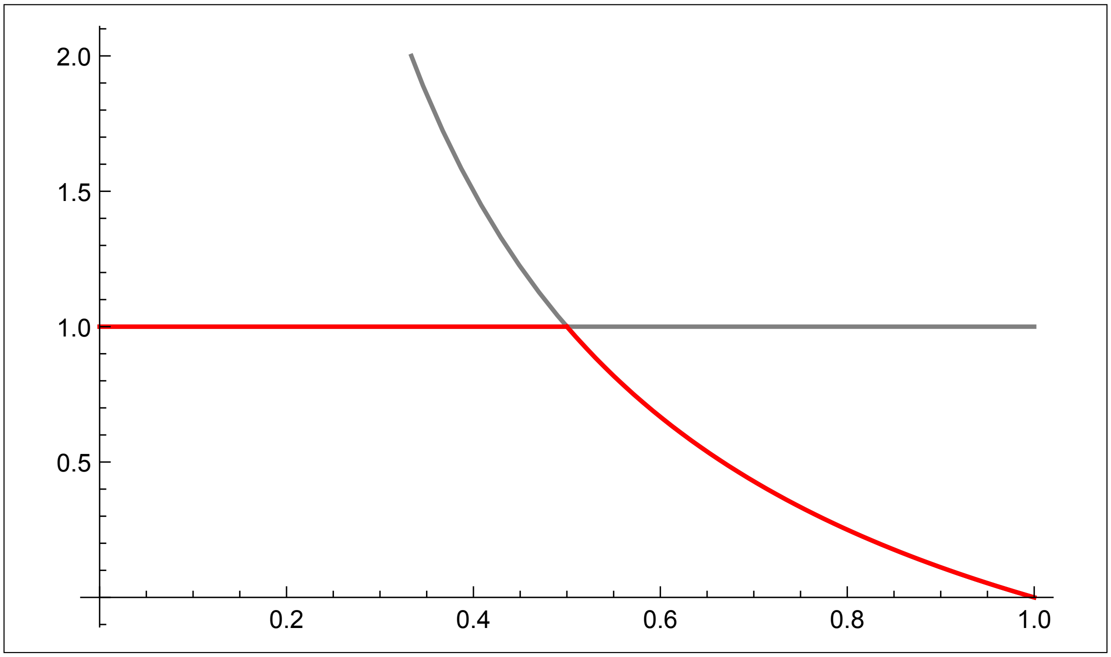


Figure 16.16. $z_1(p)$.

See Figure 16.16.

So if $p > 1/2$, the probability of reaching zero from m is

$$z_m(p) = \left(\frac{1-p}{p}\right)^m \to 0 \qquad \text{as } m \to \infty.$$

But for $p = 1/2$, $z_m(p)$ is always 1. This is just one of the ways the simple random walk is special.

16.16 ⋆ The Gambler's Ruin problem

The random walk is a model of the fortunes of a gambler who always make the same size bet. We saw that if the probability of winning is $p > 1/2$, then there is a positive probability that the gambler may never go bankrupt.[2] What happens when the gambler plays against a casino that has limited resources? What is the probability that the gambler "breaks the bank?" That is, the casino goes bankrupt before the gambler does?

- One gambler, call him Bond, starts with fortune b.
- The other, call him Goldfinger, start with fortune g.
- They play until one is bankrupted. (How do we know this happens with probability 1?)
- Let p be the probability Bond wins each bet. As Bond is the better gambler, assume that

$$p \geqslant 1/2,$$

[1] You should not let me get away with that assertion without more work.

[2] The term bankrupt, meaning almost literally "broken bank," is derived from the ancient Greek practice of punishing debtors who cannot repay their debts by breaking (rupturing) their workbench (bank).

and to simplify notation, let

$$q = 1 - p.$$

- What is the probability that Bond (the stronger player) breaks Goldfinger?

Aside: This is an example of a Markov chain with two **absorbing states**. A state in a Markov chain is absorbing if the probability of leaving it is zero. The two absorbing states are 0 (Bond is broken) and $m + n$ (Goldfinger is broken).

- Let

$$B \text{ denote the probability that Bond breaks Goldfinger.}$$

- Consider first the counterfactual that Bond is not playing against Goldfinger, but is playing against the Federal Reserve Bank, which can create money at will. We have seen that the probability the Fed breaks Bond is just $(q/p)^b$.

There two ways the Fed can break Bond:

- ◦ One is that Bond never attains $b + g$ on his way to bankruptcy,
- ◦ the other is that he does attain $b + g$ before bankruptcy, but is still broken by the Fed.

- The event that Bond never attains $b+g$ on his way to bankruptcy has the same probability that Goldfinger breaks Bond, namely $1 - B$.

- The event that Bond attains $b+g$ before bankruptcy is the event that Bond breaks Goldfinger, which happens with probability B.

- The probability that the Fed breaks Bond upon reaching $b + g$ is just $(q/p)^{b+g}$.

- Thus

$$\underbrace{\left(\frac{1-p}{p}\right)^b}_{\text{prob Fed breaks Bond}} = \underbrace{(1-B)}_{\text{prob Goldfinger breaks Bond}} + \underbrace{B}_{\text{prob Bond reaches } b+g} \underbrace{\left(\frac{1-p}{p}\right)^{b+g}}_{\text{prob Fed breaks Bond from } b+g}$$

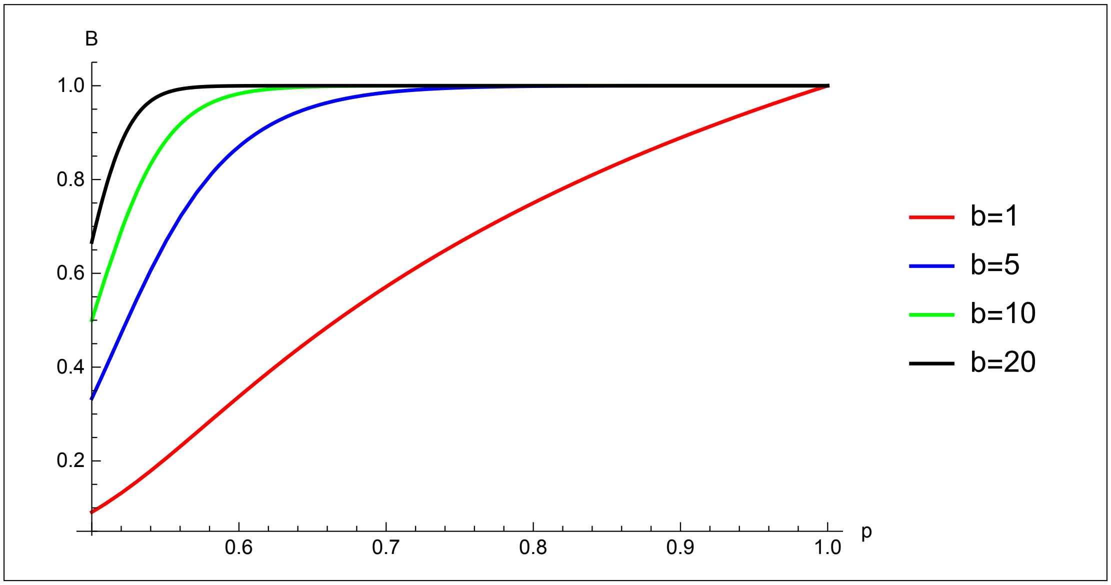
- Solving for B gives

$$\text{Probability } B \text{ that Bond breaks Goldfinger} = \frac{1 - \left(\frac{1-p}{p}\right)^b}{1 - \left(\frac{1-p}{p}\right)^{b+g}}.$$

- For the case $p = 1/2$, the formula gives $0/0$, but evaluating it with l'Hôpital's rule gives

$$\frac{b}{b+g}.$$

- Here is a chart showing the effect of p and b on the probability that Bond breaks Goldfinger, holding g fixed at 10.



16.17 ⋆ Brownian motion

Brownian motion is a continuous time stochastic process on a time interval, say $T = [0, 1]$ that is an idealized version of the simple random walk in which with "infinitely short" steps being taken "infinitely often" in any time interval. The summary is this. There is an uncountable probability space (Ω, Σ, P) and a real function

$$B : T \times \Omega \to \boldsymbol{R}.$$

For each $t \in T$ there is a random variable

$$B_t : \Omega \to \boldsymbol{R} \text{ defined by } B_t(\omega) = B(t, \omega),$$

and for each $\omega \in \Omega$, there is a function

$$B_\omega : T \to \boldsymbol{R} \text{ defined by } B_\omega(t) = B(t, \omega).$$

The key properties are

- For each $t \in T$, the random variable B_t has a Normal$(0, t)$ distribution.
- Increments are stochastically independent. That is, whenever

$$0 \leqslant t_0 \leqslant t_1 \leqslant \cdots \leqslant t_k \leqslant 1,$$

the random variables

$$B_{t_1} - B_{t_0},\ B_{t_2} - B_{t_1},\ \ldots,\ B_{t_k} - B_{t_{k-1}}$$

are stochastically independent.

- For every $\omega \in \Omega$, the function B_ω is continuous.

These properties pin Brownian motion. But how do we know such a stochastic process exists? One way to approach it is take as the sample space Ω the set of continuous function on $[0, 1]$. The random experiment consists of drawing a function $\omega \in C[0, 1]$ at random according to some probability measure W on $C[0, 1]$, where W has the desired properties. The measure W is known as **Wiener measure** after Norbert Wiener [15, pp. 214–234]. Now we just have find W.

Let's start by considering a random walk with many small steps per time interval. Rescale the simple random walk by taking n steps per epoch, but shortening them to have length $1/\sqrt{n}$. Now

$$X_k^{(n)} = X_k/\sqrt{n}$$

has mean 0 and variance $1/n$. The sum of nt of these scaled Rademachers thus has variance t. Define the partial sums

$$S_k^{(n)} = X_1^{(n)} + \cdots + X_k^{(n)},$$

We now squash these down so that every time interval of length k has nk steps in it, and linearly interpolate to get a continuous function like this:

$$B^{(n)}(t,\omega) = S_{[nt]}(\omega) + (nt - [nt])X_{[nt]+1}(\omega), \qquad (t \geqslant 0),$$

Figure 16.17 show the graph of the function $B_\omega^{(n)}$ for a fixed ω and $n = 1, 4, 16$. When nt is an

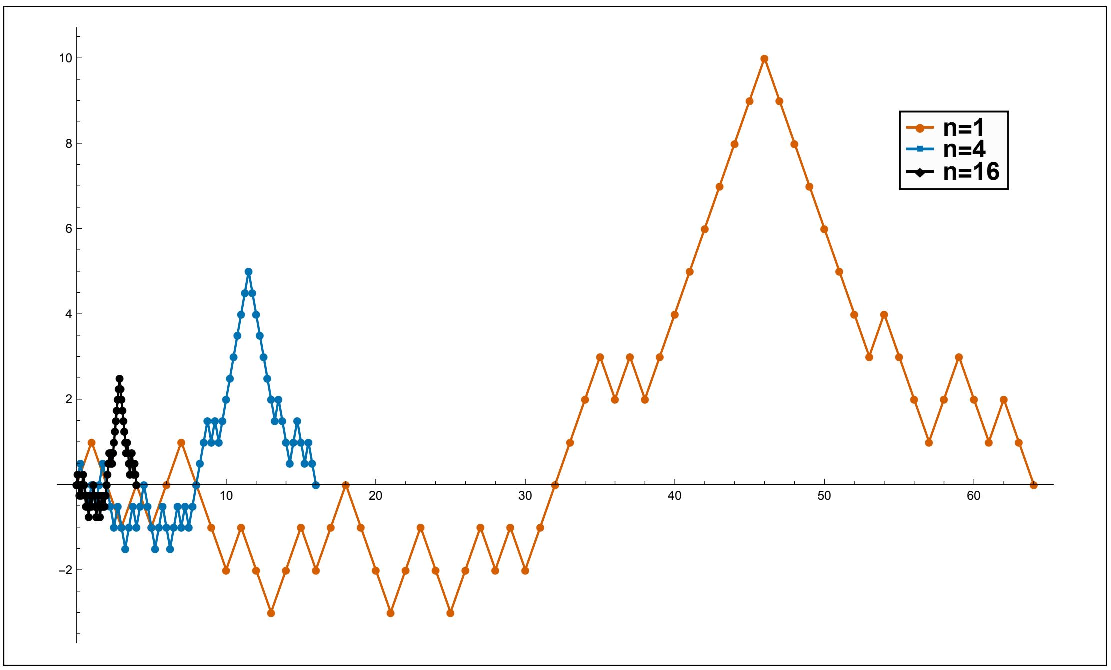


Figure 16.17. Scaled sample paths of a random walk.

integer, then

$$B_t^{(n)} = S_{nt}/\sqrt{n}$$

has mean 0 and variance t. For such t, the Central Limit Theorem implies

$$YB_t^{(n)} \xrightarrow[n\to\infty]{\mathcal{D}} N(0,t).$$

Also note that for each n, for $0 \leqslant t_0 \leqslant t_1 \leqslant \cdots \leqslant t_k \leqslant 1$, the random variables

$$B_{t_1}^{(n)} - B_{t_0}^{(n)},\ B_{t_2}^{(n)} - B_{t_1}^{(n)},\ \ldots,\ B_{t_k}^{(n)} - B_{t_{k-1}}^{(n)}$$

are stochastically independent (being the sums of disjoint sets of independent scaled Rademachers).

At this point I am going to make a lot of claims that are well beyond the scope of this course. For each n, $B^{(n)}$ defines a probability distribution on the set $C[0,1]$ of continuous functions on the interval as follows. For each subset A,[3] define its probability by

$$\text{Prob}(A) = P\{\omega \in \Omega : B^{(n)}_\omega \in A\}.$$

It is possible to define convergence in distribution for probability distributions on spaces such as $C[0,1]$ (see for instance Aliprantis and Border [1, Chapter 15], Billingsley [4], or Parthasarathy [14]).[4] Monroe Donsker [9] proved that the probability distribution on $C[0,1]$ associated with $B^{(n)}$ converges in distribution to Weiner measure W as $n \to \infty$. This result is known as **Donsker's Theorem**. See also Billingsley [4, § 10, p. 68].

Informally this means that properties of Brownian motion can be deduced as limiting results from scaled random walks.

As an example, Billingsley [4, pp. 80–83] uses the Arc Sine Law for random walks to prove the Arc Sine Law for Brownian motion on $[0,1]$: If T is the last time $B_t = 0$, then $P(T \leqslant t) = \frac{2}{\pi} \arcsin \sqrt{t}$. Compare this with (18) on page 16–16 of these notes.

16.18 What is Brownian motion a model of?

Brownian motion is named after Robert Brown, a 19th century botanist who observed the motion of pollen particles in water [5, p. 443]. We can generalize Brownian motion to several dimensions by considering random vector functions, where the components are independent Brownian motions. To cover such things as the two or three dimensional motion of particles colliding molecules of air or water. Brownian motion was the first good model of random diffusion.

In 1900, Louis Bachelier [3, 8] developed a mathematical model for Brownian motion and argued that stock prices must follow a Brownian motion stochastic process. In 1973 Fischer Black and Myron Scholes [6] independently rediscovered some of Bachelier's results.

Bibliography

[1] C. D. Aliprantis and K. C. Border. 2006. *Infinite dimensional analysis: A hitchhiker's guide*, 3d. ed. Berlin: Springer–Verlag.

[2] T. M. Apostol. 1967. *Calculus*, 2d. ed., volume 1. Waltham, Massachusetts: Blaisdell.

[3] L. J.-B. A. Bachelier. 1900. Théorie de la spéculation. *Annales Scientifiques de l'École Normale Supérieure, 3 Série* 17:21–86.
http://www.numdam.org/item?id=ASENS_1900_3_17__21_0

[4] P. Billingsley. 1968. *Convergence of probability measures*. Wiley Series in Probability and Mathematical Statistics. New York: Wiley.

[5] ——— . 1979. *Probability and measure*. Wiley Series in Probability and Mathematical Statistics. New York: Wiley.

[3] Actually we want to restrict A to be a "nice" subset of $C[0,1]$. First turn $C[0,1]$ into a normed space with norm $\|f\| = \max_{x\in[0,1} f(x)$. This gives rise to the notion of open sets in $C[0,1]$. The smallest σ-algebra that includes the open sets is called the **Borel σ-algebra**. We restrict the discussion to sets B that belong to the Borel σ-algebra.

[4] Recall Fact 11.3.4 which states that if each X_n is a random variable with probability distribution P_n, X has probability distribution P, and if g is a bounded continuous function, then $P_n \xrightarrow{\mathcal{D}} P$ implies $\boldsymbol{E}\, g(X_n) \to \boldsymbol{E}\, g(X)$. We can turn this around *define* $P_n \xrightarrow{\mathcal{D}} P$ to mean that for every such g, we have $\boldsymbol{E}\, g(X_n) \to \boldsymbol{E}\, g(X)$.

[6] F. Black and M. Scholes. 1973. The pricing of options and corporate liabilities. *Journal of Political Economy* 81(3):637–654. http://www.jstor.org/stable/1831029

[7] D. Blackwell, P. Deuel, and D. Freedman. 1964. The last return to equilibrium in a coin-tossing game. *Annals of Mathematical Statistics* 35:1344.
http://www.jstor.org/stable/2238263

[8] M. Davis and A. Etheridge. 2011. *Louis Bachelier's theory of speculation: The origins of modern finance*. Princeton, New Jersey: Princeton University Press.

[9] M. D. Donsker. 1951. An invariance priciple for certain probaility limit theorems. *Memoirs of the American Mathematical Society* 6.

[10] W. Feller. 1950. *An introduction to probability theory and its applications*, 1st. ed., volume 1. New York: Wiley.

[11] ——— . 1968. *An introduction to probability theory and its applications*, 3d. ed., volume 1. New York: Wiley.

[12] ——— . 1971. *An introduction to probability theory and its applications*, 2d. ed., volume 2. New York: Wiley.

[13] F. Mosteller. 1987. *Fifty challenging problems in probability with solutions*. Mineola, New York: Dover. Reprint of the 1965 Addison-Wesley edition.

[14] K. R. Parthasarathy. 1967. *Probability measures on metric spaces*. Probability and Mathematical Statistics. New York: Academic Press.

[15] N. Wiener. 1930. Generalized harmonic analysis. *Acta Mathematica* 55(1):117–258.
DOI: 10.1007/BF02546511