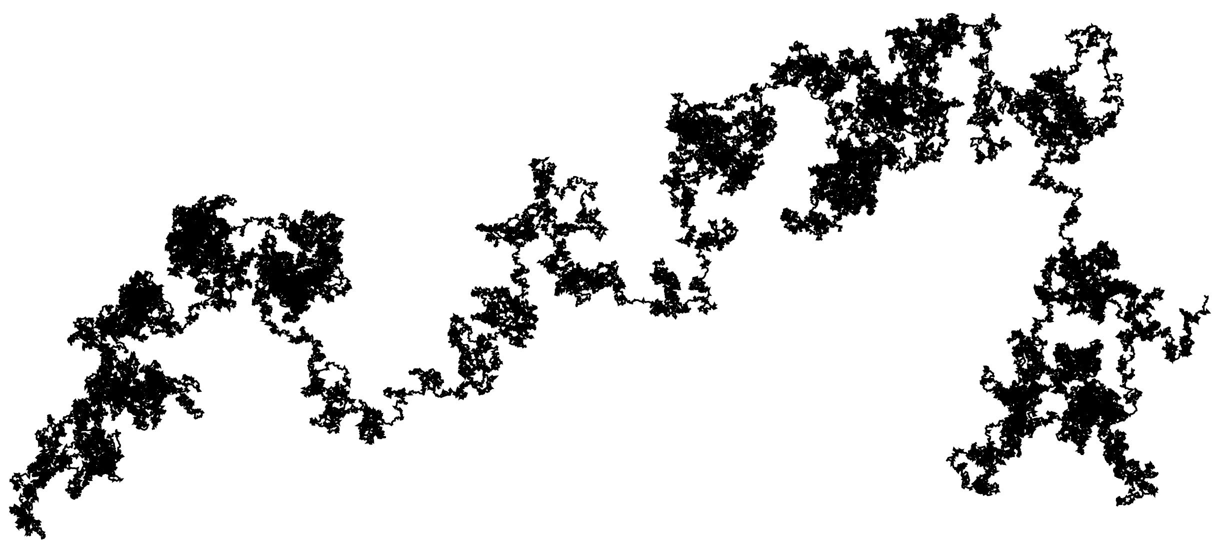
MAT 627, Spring 2022, Stony Brook University

TOPICS IN COMPLEX ANALYSIS

CONFORMAL FRACTALS, PART II: BROWNIAN MOTION

Christopher Bishop, Stony Brook



Construction of Brownian motion

Nowhere differentiable

Hausdorff dimension

Stopping times and the Markov property

Area of 2-dimensional Brownian motion

The Law of the Iterated Logarithm

Solving the Dirichlet problem with Brownian motion

2-dimensional Brownian motion is recurrent

2-dimensional Brownian motion is conformally invariant

These slides are based on the text

Fractals in probability and analysis

A more detailed discussion is found there, and can be found in many other texts on this topic, such as

Brownian motion by Mörters and Peres.

Some other standard (offline) textbooks are:

Brownian Motion and Stochastic Calculus by Karatzas and Shreve
Continuous Martingales and Brownian Motion by Revuz and Yor
Probability and Stochastics by Cinlar
An Introduction to Stochastic Processes by Schilling and Partzsch

1. Construction of Brownian motion

A real-valued random variable X on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ has a **standard Gaussian** (or **standard normal**) distribution if

$$\mathbb{P}(X > x) = \frac{1}{\sqrt{2\pi}} \int_x^{+\infty} e^{-u^2/2}\, du$$

A vector-valued random variable X has an **n-dimensional standard Gaussian** distribution if its n coordinates are standard Gaussian and independent.

A vector-valued random variable $Y : \Omega \to \mathbb{R}^p$ is **Gaussian** if there exists a vector-valued random variable X having an n-dimensional standard Gaussian distribution, a $p \times n$ matrix A and a p-dimensional vector b such that

$$Y = AX + b. \tag{1.1}$$

Lemma 1.1. *If Θ is an orthogonal $n \times n$ matrix and X is an n-dimensional standard Gaussian vector, then ΘX is also an n-dimensional standard Gaussian vector.*

Proof. As the coordinates of X are independent standard Gaussian, X has density given by

$$f(x) = (2\pi)^{-\frac{n}{2}} e^{-\|x\|^2/2},$$

where $\|\cdot\|$ denotes the Euclidean norm. Since Θ preserves this norm, the density of X is invariant under Θ. $\square$

Corollary 1.2. *Let Z_1, Z_2 be independent $\mathcal{N}(0, \sigma^2)$ random variables. Then $Z_1 + Z_2$ and $Z_1 - Z_2$ are two independent random variables having the same distribution $\mathcal{N}(0, 2\sigma^2)$.*

Proof. $\sigma^{-1}(Z_1, Z_2)$ is a standard Gaussian vector, and so, if

$$\Theta = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix},$$

then Θ is an orthogonal matrix such that

$$(\sqrt{2}\sigma)^{-1}(Z_1 + Z_2, Z_1 - Z_2)^t = \Theta\sigma^{-1}(Z_1, Z_2)^t,$$

and our claim follows from Part (i) of the Lemma. $\square$

Lemma 1.3. *Let Z be distributed as $\mathcal{N}(0,1)$. Then for all $x \geq 0$,*

$$\frac{x}{x^2+1}\frac{1}{\sqrt{2\pi}}e^{-x^2/2} \leq \mathbb{P}\left(Z > x\right) \leq \frac{1}{x}\frac{1}{\sqrt{2\pi}}e^{-x^2/2}.$$

Proof. The right inequality is obtained by the estimate

$$\mathbb{P}(Z > x) \leq \int_x^{+\infty} \frac{u}{x}\frac{1}{\sqrt{2\pi}}e^{-u^2/2}\,du$$

since, in the integral, $u \geq x$.

The left inequality is proved as follows: let us define

$$f(x) := xe^{-x^2/2} - (x^2+1)\int_x^{+\infty} e^{-u^2/2}\,du.$$

We remark that $f(0) < 0$ and $\lim_{x\to+\infty} f(x) = 0$.

$$\begin{aligned} f'(x) &= (1 - x^2 + x^2 + 1)e^{-x^2/2} - 2x\int_x^{+\infty} e^{-u^2/2}\,du \\ &= -2x\left(\int_x^{+\infty} e^{-u^2/2}\,du - \frac{1}{x}e^{-x^2/2}\right), \end{aligned}$$

so the right inequality implies $f'(x) \geq 0$ for all $x \geq 0$.

This implies $f(x) \leq 0$, proving the lemma. $\square$

A real-valued stochastic process $\{B_t\}_{t\in I}$ is a **standard Brownian motion** if it is a Gaussian process such that:

i. $B_0 = 0$;

ii. for all k natural and for all $t_1 < \cdots < t_k$ in I: $B_{t_k} - B_{t_{k-1}}, \ldots, B_{t_2} - B_{t_1}$ are independent;

iii. for all $t, s \in I$ with $t < s$, $B_s - B_t$ has $\mathcal{N}(0, s-t)$ distribution;

iv. almost surely, $t \mapsto B_t$ is continuous on I.

Recall that the **covariance matrix** of a random vector is defined as

$$\mathrm{Cov}(Y) = \mathbb{E}\left[(Y - \mathbb{E}Y)(Y - \mathbb{E}Y)^t\right].$$

Then, by the linearity of expectation, the Gaussian vector $Y = AX + b$ in (1.1) has

$$\mathrm{Cov}(Y) = AA^t.$$

As a corollary of this definition, one can already remark that for all $t, s \in I$:

$$\operatorname{Cov}(B_t, B_s) = s \wedge t,$$

(where $s \wedge t = \min(s, t)$). Indeed, assume that $t \geq s$. Then

$$\operatorname{Cov}(B_t, B_s) = \operatorname{Cov}(B_t - B_s, B_s) + \operatorname{Cov}(B_s, B_s)$$

by bi-linearity of the covariance. The first term vanishes by the independence of increments, and the second term equals s by properties (iii) and (i).

We may replace properties (ii) and (iii) in the definition by:

- for all $t, s \in I$, $\mathrm{Cov}(B_t, B_s) = t \wedge s$;
- for all $t \in I$, B_t has $\mathcal{N}(0, t)$ distribution;

or by:

- for all $t, s \in I$ with $t < s$, $B_t - B_s$ and B_s are independent;
- for all $t \in I$, B_t has $\mathcal{N}(0, t)$ distribution.

Lemma 1.4 (Borel–Cantelli)**.** *Let* $\{A_i\}_{i=0,\ldots,\infty}$ *be a sequence of events, and let*

$$A_i \;\; i.o. = \limsup_{i\to\infty} A_i = \bigcap_{i=0}^{\infty}\bigcup_{j=i}^{\infty} A_j,$$

where "i.o." abbreviates "infinitely often".

i. If $\sum_{i=0}^{\infty}\mathbb{P}(A_i) < \infty$*, then* $\mathbb{P}(A_i \; i.o.) = 0$*.*

ii. If $\{A_i\}$ *are pairwise independent, and* $\sum_{i=0}^{\infty}\mathbb{P}(A_i) = \infty$*, then* $\mathbb{P}(A_i \; i.o.) = 1$*.*

The following construction, due to Paul Lévy, consists of choosing the "right" values for the Brownian motion at each dyadic point of $[0, 1]$ and then interpolating linearly between these values.

This construction is inductive, and at each step a process is constructed that has continuous paths. Brownian motion is then the uniform limit of these processes; hence its continuity.

Theorem 1.5. *Standard Brownian motion on* $[0, \infty)$ *exists.*

Proof. We first construct standard Brownian motion on $[0, 1]$.

For $n \geq 0$, let $D_n = \{k/2^n : 0 \leq k \leq 2^n\}$, and let $D = \bigcup D_n$. Let $\{Z_d\}_{d \in D}$ be a collection of independent $\mathcal{N}(0, 1)$ random variables. We will first construct the values of B on D. Set $B_0 = 0$, and $B_1 = Z_1$.

In an inductive construction, for each n we will construct B_d for all $d \in D_n$ so that:

i. for all $r < s < t$ in D_n, the increment $B_t - B_s$ has $\mathcal{N}(0, t-s)$ distribution and is independent of $B_s - B_r$;

ii. B_d for $d \in D_n$ are globally independent of the Z_d for $d \in D \setminus D_n$.

These assertions hold for $n = 0$. Suppose that they hold for $n - 1$.

Define, for all $d \in D_n \setminus D_{n-1}$, a random variable B_d by

$$B_d = \frac{B_{d^-} + B_{d^+}}{2} + \frac{Z_d}{2^{(n+1)/2}} \tag{1.2}$$

where $d^+ = d + 2^{-n}$, and $d^- = d - 2^{-n}$, and both are in D_{n-1}.

Because $\frac{1}{2}(B_{d^+} - B_{d^-})$ is $\mathcal{N}(0, 1/2^{n+1})$ by induction, and $Z_d/2^{(n+1)/2}$ is an independent $\mathcal{N}(0, 1/2^{n+1})$, their sum and their difference, $B_d - B_{d-}$ and $B_{d+} - B_d$ are both $\mathcal{N}(0, 1/2^n)$ and independent by Corollary 1.2.

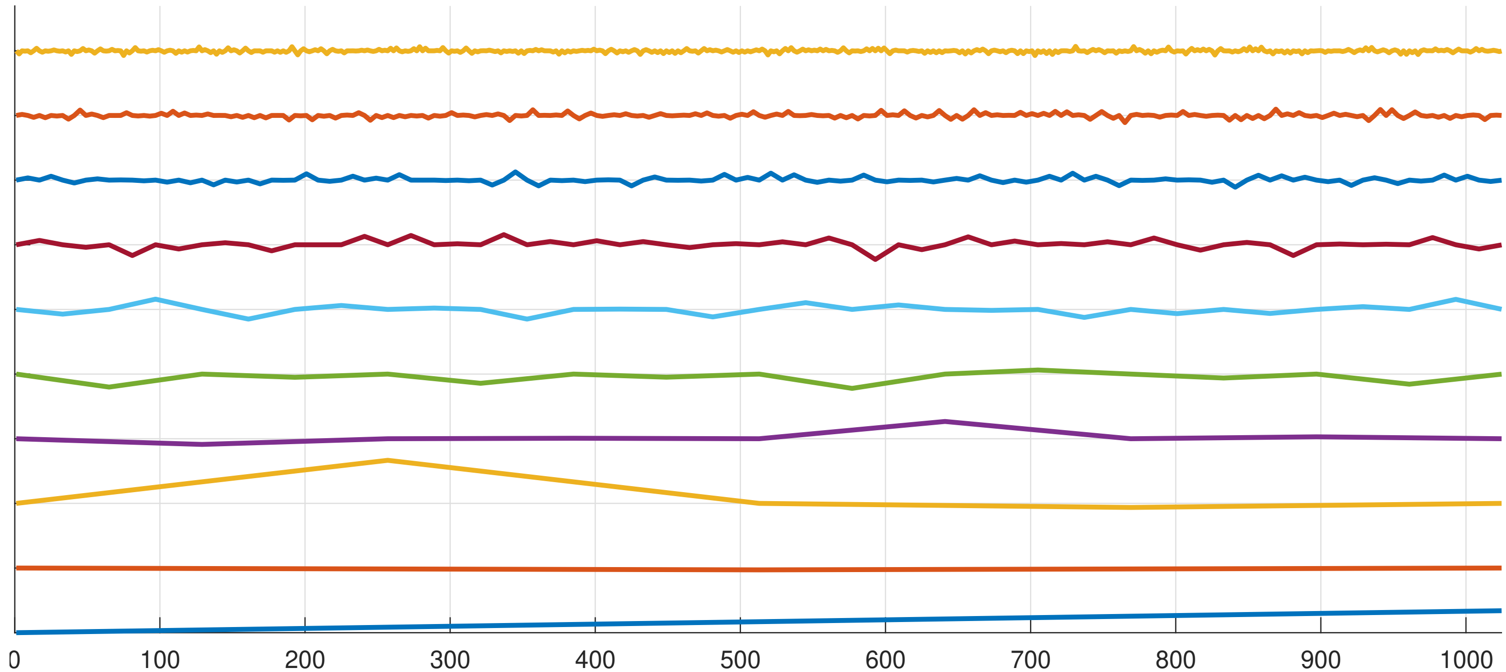
Assertion (i) follows from this and the inductive hypothesis, and (ii) is clear.

Having thus chosen the values of the process on D, we now "interpolate" between them. Formally, let $F_0(x) = xZ_1$, and for $n \geq 1$, let us introduce the function

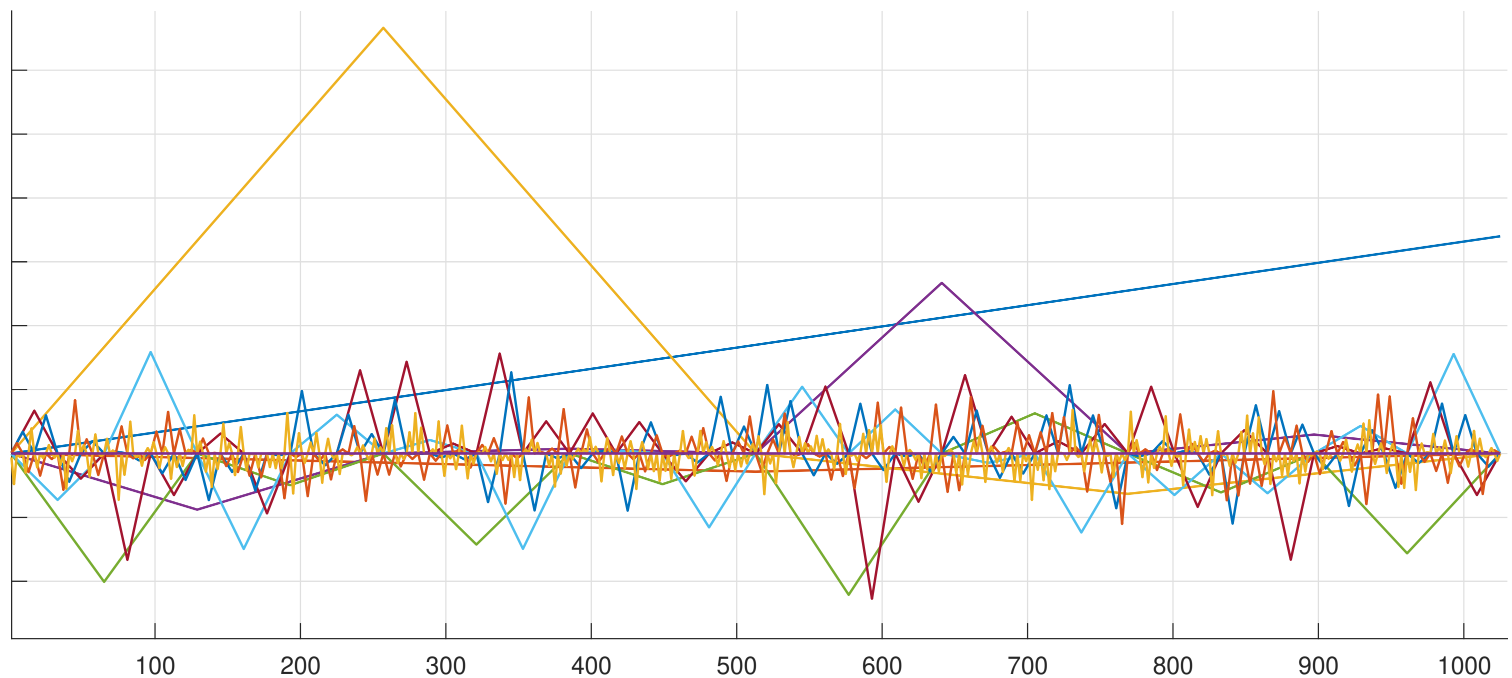
$$F_n(x) = \begin{cases} 2^{-(n+1)/2} Z_x & \text{for } x \in D_n \setminus D_{n-1}, \\ 0 & \text{for } x \in D_{n-1}, \\ \text{linear} & \text{between consecutive points in } D_n. \end{cases} \tag{1.3}$$

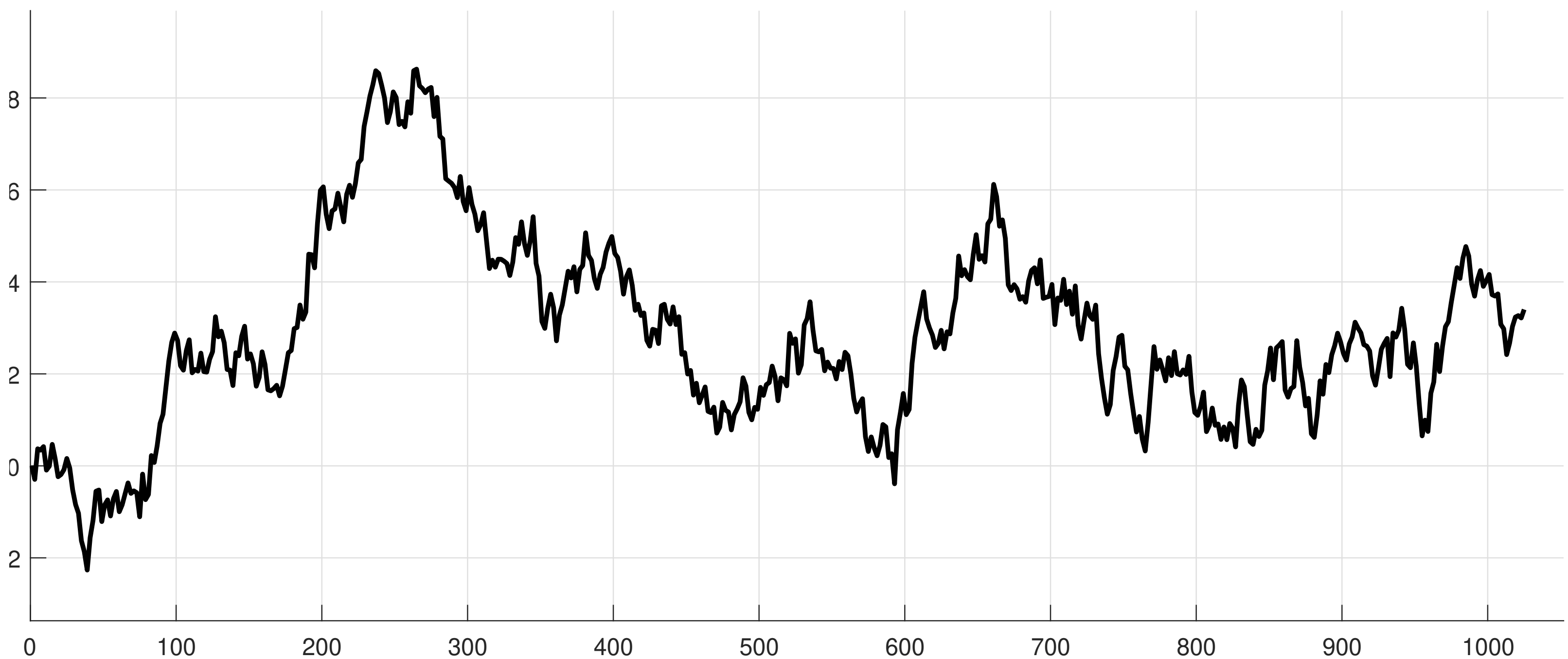
These functions are continuous on $[0, 1]$, and for all n and $d \in D_n$

$$B_d = \sum_{i=0}^{n} F_i(d) = \sum_{i=0}^{\infty} F_i(d). \tag{1.4}$$



0
100
200
300
400
500
600
700
800
900
1000


100
200
300
400
500
600
700
800
900
1000



8
6
4
2
0
2
0
100
200
300
400
500
600
700
800
900
1000

We prove this by induction. Suppose that it holds for $n-1$. Let $d \in D_n \setminus D_{n-1}$. Since for $0 \leq i \leq n-1$ the function F_i is linear on $[d^-, d^+]$, we get

$$\sum_{i=0}^{n-1} F_i(d) = \sum_{i=0}^{n-1} \frac{F_i(d^-) + F_i(d^+)}{2} = \frac{B_{d^-} + B_{d^+}}{2}. \tag{1.5}$$

Since $F_n(d) = 2^{-(n+1)/2} Z_d$, comparing (1.2) and (1.5) gives (1.4).

On the other hand, we have by definition of Z_d and by Lemma 1.3

$$\mathbb{P}\left(|Z_d| \geq c\sqrt{n}\right) \leq \exp\left(-\frac{c^2 n}{2}\right)$$

for n large enough, so the series $\sum_{n=0}^{\infty} \sum_{d \in D_n} \mathbb{P}(|Z_d| \geq c\sqrt{n})$ converges as soon as $c > \sqrt{2 \log 2}$.

Fix such a $c > \sqrt{2 \log 2}$.

By the Borel–Cantelli Lemma 1.4 we conclude that there exists a random but finite N so that for all $n > N$ and $d \in D_n$ we have $|Z_d| < c\sqrt{n}$, and so

$$|F_n|_\infty < c\sqrt{n}2^{-n/2}. \tag{1.6}$$

This upper bound implies that the series $\sum_{n=0}^{\infty} F_n(t)$ is uniformly convergent on $[0, 1]$, and so it has a continuous limit, which we call $\{B_t\}$. All we have to check is that the increments of this process have the right finite-dimensional joint distributions. This is a direct consequence of the density of the set D in $[0, 1]$ and the continuity of paths.

Indeed, let $t_1 > t_2 > t_3$ be in $[0, 1]$, then they are limits of sequences $t_{1,n}$, $t_{2,n}$ and $t_{3,n}$ in D, respectively. Now

$$B_{t_3} - B_{t_2} = \lim_{k\to\infty} (B_{t_{3,k}} - B_{t_{2,k}})$$

is a limit of Gaussian random variables, so itself is Gaussian with mean 0 and variance $\lim_{n\to\infty} (t_{3,n} - t_{2,n}) = t_3 - t_2$.

The same holds for $B_{t_2} - B_{t_1}$; moreover, these two random variables are limits of independent random variables, since for n large enough, $t_{1,n} > t_{2,n} > t_{3,n}$. Applying this argument for any number of increments, we get that $\{B_t\}$ has independent increments such that for all $s < t$ in $[0, 1]$, $B_t - B_s$ has $\mathcal{N}(0, t - s)$ distribution.

We have thus constructed Brownian motion on $[0, 1]$.

Finally, if $\{B_t^n\}_n$ for $n \geq 0$ are independent Brownian motions on $[0, 1]$, then

$$B_t = B_{t-\lfloor t \rfloor}^{\lfloor t \rfloor} + \sum_{0 \leq i < \lfloor t \rfloor} B_1^i$$

meets our definition of Brownian motion on $[0, \infty)$. $\square$

Let $\{X_i\}_{i\geq 1}$ be i.i.d. random variables with mean 0 and finite variances. By normalization, we can assume the variance $\mathrm{Var}(X_i) = 1$, for all i. Let $S_n = \sum_{i=1}^n X_i$, and interpolate it linearly to get the continuous paths $\{S_t\}_{t\geq 0}$.

Theorem 1.6 (Donsker's Invariance Principle)**.** *As $n \longrightarrow \infty$,*

$$\left\{\frac{S_{tn}}{\sqrt{n}}\right\}_{0\leq t\leq 1} \stackrel{\text{in law}}{\Longrightarrow} \{B_t\}_{0\leq t\leq 1},$$

i.e., if $\psi : \tilde{C}[0,1] \longrightarrow \mathbb{R}$, where $\tilde{C}[0,1] = \{f \in C[0,1] : f(0) = 0\}$, is a bounded continuous function with respect to the sup norm, then, as $n \longrightarrow \infty$,

$$\mathbb{E}\psi\left(\left\{\frac{S_{tn}}{\sqrt{n}}\right\}_{0\leq t\leq 1}\right) \longrightarrow \mathbb{E}\psi(\{B_t\}_{0\leq t\leq 1}).$$

Examples:

$X_n = \pm 1$ with equal probability.

X_n chosen uniformly in $[-1, 1]$

$X_n = \{-1, 0, 1\}$ with equal probability.

$X_n = \{-1, \frac{1}{2}\}$ with probabilities $\frac{1}{3}, \frac{2}{3}$.

Non-example:

X_n chosen with distribution $\frac{1}{\pi(1+x^2)}$.

Variance is not finite.

Gives the Cauchy process, a discontinuous random map $[0, \infty) \to \mathbb{R}$.

The Cauchy process.

$C(x) = y$ if $y = B_2(t)$ and t is the first time $B_1(t) = x$.

2. Nowhere differentiable

We say that two random variables Y, Z have the same distribution, and write $Y \stackrel{\mathrm{d}}{=} Z$, if $\mathbb{P}(Y \in A) = \mathbb{P}(Z \in A)$ for all Borel sets A. Let $\{B(t)\}_{t\geq 0}$ be a standard Brownian motion, and let $a \neq 0$.

The following **scaling relation** is a simple consequence of the definitions.

$$\{\tfrac{1}{a}B(a^2t)\}_{t\geq 0} \stackrel{\mathrm{d}}{=} \{B(t)\}_{t\geq 0}.$$

Proof. Continuity of the paths, independence and stationarity of the increments remain unchanged under the scaling.

It remains to observe that

$$X(t) - X(s) = \frac{1}{a}(B(a^2t) - B(a^2s)),$$

is normally distributed with expectation 0 and variance

$$(1/a^2)(a^2t - a^2s) = t - s.$$

□

Scaling invariance has many useful consequences.

For example, if $a > 0$ let

$$T(a) = \inf\{t > 0 : B(t) = a \text{ or } B(t) = b\},$$

the first exit time of a one-dimensional standard Brownian motion from the interval $[-a, a]$. Then, with $X(t) = (1/a)B(a^2 t)$ we have

$$\mathbb{E}T(a) = a^2 2\mathbb{E}\inf\{t > 0 : |X(t)| = 1\}.$$

This implies that $\mathbb{E}T(a)$ is a constant multiple of a^2.

Later we will compute the constant $c = 1$ using Wald's lemma.

Later we will prove conformal invariance, a powerful extension of scaling invariance.

Define the **time inversion** of $\{B_t\}$ as

$$W(t) = \begin{cases} 0 & t = 0; \\ tB(\frac{1}{t}) & t > 0. \end{cases}$$

Lemma 2.1. *W is standard Brownian motion.*

Proof. Recall that the finite-dimensional distributions $(B(t_1), ..., B(t_n))$ of Brownian motion are Gaussian random vectors and are therefore characterized by $\mathbb{E}[B(t_i)] = 0$ and $\mathrm{Cov}(B(t_i), B(t_j)) = t_i$ for $0 \le t_i \le t_j$.

Obviously, $\{W(t) : t > 0\}$ is also a Gaussian process and the Gaussian random vectors $(W(t_1), ..., W(t_n))$ have expectation zero. The covariances, for $t > 0, h > 0$, are given by

$\mathrm{Cov}(W(t+h), W(t)) = (t+h)t\,\mathrm{Cov}(B(1/(t+h)), B(1/t)) = t(t+h)1t+h = t.$

Hence the law of all the finite-dimensional distributions

$$(W(t_1), W(t_2), ..., W(t_n)),$$

for $0 \le t_1 \le \cdots \le t_n$, are the same as for Brownian motion.

The paths of $t \to W(t)$ are clearly continuous for all $t > 0$. To prove continuity at $t = 0$ we use the following two facts.

First, for any rational $t \geq 0$ the distribution of $\{W(t)\}$ is the same as for a Brownian motion, and hence

$$\lim_{t \to 0, t \in \mathbb{Q}} W(t) = 0$$

almost surely.

Second, the rationals are dense in the reals and $W(t)$ is almost surely continuous on $(0, \infty$ so that

$$\lim_{t \to 0} W(t) = 0$$

almost surely. Hence $\{W(t) : t > 0\}$ has almost surely continuous paths, and is a Brownian motion. $\square$

A function f is α-Hölder if for some $C < \infty$ we have

$$|f(x) - f(y)| \leq C|x - y|^{\alpha},$$

1-Hölder = Lipschitz

α-Hölder for $\alpha > 1$ is constant.

Corollary 2.2. *Brownian paths are α-Hölder a.s. for all $\alpha < \frac{1}{2}$.*

Proof. We defined Brownian motion as an infinite sum $\sum_{n=0}^{\infty} F_n$, where each F_n is a piecewise linear function given in (1.3).

The derivative of F_n exists except on a finite set, and by definition and (1.6)

$$\|F_n'\|_\infty \le \frac{\|F_n\|_\infty}{2^{-n}} \le C_1(\omega) + c\sqrt{n}\, 2^{n/2}. \tag{2.1}$$

The random constant $C_1(\omega)$ is introduced to deal with the finitely many exceptions to (1.6).

Now for $t, t+h \in [0,1]$, we have

$$\begin{aligned} |B(t+h) - B(t)| &\le \sum_n |F_n(t+h) - F_n(t)| \\ &\le \sum_{n \le \ell} h\|F_n'\|_\infty + \sum_{n > \ell} 2\|F_n\|_\infty. \end{aligned}$$

By (1.6) and (2.1) if $\ell > N$ for a random N, then the above is bounded by

$$h\left(C_1(\omega) + \sum_{n\le\ell} c\sqrt{n}\,2^{n/2}\right) + 2\sum_{n>\ell} c\sqrt{n}\,2^{-n/2}$$

$$\le C_2(\omega)h\sqrt{\ell}\,2^{\ell/2} + C_3(\omega)\sqrt{\ell}\,2^{-\ell/2}.$$

The inequality holds because each series is bounded by a constant times its dominant term.

Choosing $\ell = \lfloor \log_2(1/h) \rfloor$, and choosing $C(\omega)$ to take care of the cases when $\ell \leq N$, we get

$$|B(t+h) - B(t)| \leq C(\omega)\sqrt{h \log_2 \frac{1}{h}}. \tag{2.2}$$

$\square$

A Brownian motion is almost surely not $\frac{1}{2}$-Hölder. This will be proven later.

Lemma 2.3. *The paths of Brownian motion have no intervals of monotonicity.*

Proof. Indeed, if $[a, b]$ is an interval of monotonicity, then dividing it up into n equal sub-intervals $[a_i, a_{i+1}]$ each increment $B(a_i) - B(a_{i+1})$ has to have the same sign. This has probability $2 \cdot 2^{-n}$, and taking $n \to \infty$ shows that the probability that $[a, b]$ is an interval of monotonicity must be 0.

Taking a countable union gives that there is no interval of monotonicity with rational endpoints, but each monotone interval would have a monotone rational sub-interval. □

Consider a probability measure on the space of real sequences, and let $X_1, X_2, \ldots$ be the sequence of random variables it defines. An event, i.e., a measurable set of sequences, A is **exchangeable** if $X_1, X_2, \ldots$ satisfy A implies that $X_{\sigma_1}, X_{\sigma_2}, \ldots$ satisfy A for all finite permutations σ. Finite permutation means that $\sigma_n = n$ for all sufficiently large n.

Proposition 2.4 (Hewitt–Savage 0-1 Law)**.** *If A is an exchangeable event for an i.i.d. sequence then $\mathbb{P}(A)$ is 0 or 1.*

Sketch of Proof: Given i.i.d. variables $X_1, X_2, \ldots$, suppose that A is an exchangeable event for this sequence. Then for any $\epsilon > 0$ there is an integer n and a Borel set $B_n \subset \mathbb{R}^n$ such that the event $A_n = \{\omega : (X_1, \ldots, X_n) \in B_n\}$ satisfies $\mathbb{P}(A_n \Delta A) < \epsilon$.

Now apply the permutation σ that transposes i with $i + n$ for $1 \leq i \leq n$, i.e., it exchanges the blocks

$$[1, ..., n] \text{ and } [n + 1, ..., 2n].$$

The event A is pointwise fixed by this transformation of the measure space (since A is exchangeable) and the probability of any event is invariant (the measure space is a product space with identical distributions in each coordinate and we are simply reordering the coordinates).

Thus A_n is sent to a new event A_n^σ that has the same probability and $\mathbb{P}((A_n^\sigma \Delta A) = \mathbb{P}(A_n \Delta A) < \epsilon$, hence $\mathbb{P}(A_n^\sigma \Delta A_n) < 2\epsilon$ (since $\mathbb{P}(X\Delta Y)$ defines a metric on measurable sets).

But A_n and A_n^σ are independent, so $\mathbb{P}(A_n \cap A_n^\sigma) = \mathbb{P}(A_n)\mathbb{P}(A_n^\sigma) = \mathbb{P}(A_n)^2$. Thus

$$\mathbb{P}(A) = \mathbb{P}(A_n \cap A_n^\sigma) + O(\epsilon) = \mathbb{P}(A_n)^2 + O(\epsilon) = \mathbb{P}(A)^2 + O(\epsilon).$$

Taking $\epsilon \to 0$ shows $\mathbb{P}(A) \in \{0, 1\}$. $\square$

Proposition 2.5. *Almost surely*

$$\limsup_{n\to\infty} \frac{B(n)}{\sqrt{n}} = +\infty, \qquad \liminf_{n\to\infty} \frac{B(n)}{\sqrt{n}} = -\infty. \tag{2.3}$$

Proof. In general, the probability that infinitely many events $\{A_n\}$ occur satisfies

$$\mathbb{P}(A_n \text{ i.o.}) = \mathbb{P}(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k) = \lim_{n\to\infty} \mathbb{P}(\bigcup_{k=n}^{\infty} A_k) \geq \limsup_{n\to\infty} \mathbb{P}(A_n).$$

So, in particular,

$$\mathbb{P}(B(n) > c\sqrt{n} \text{ i.o.}) \geq \limsup_{n\to\infty} \mathbb{P}(B(n) > c\sqrt{n}).$$

By the scaling property, the expression in the lim sup equals $\mathbb{P}(B(1) > c)$, which is positive.

Let $X_n = B(n) - B(n-1)$. These are i.i.d. random variables,

$$\{\sum_{k=1}^{n} X_k > c\sqrt{n}\ i.o.\} = \{B(n) > c\sqrt{n}\ i.o.\}$$

is exchangeable and has positive probability, so the Hewitt–Savage 0–1 law says it has probability 1.

Taking the intersection over all natural numbers c gives the first part of Proposition 2.5, and the second is proved similarly. $\square$

The two claims of Proposition 2.5 together mean that $B(t)$ crosses 0 for arbitrarily large values of t. If we use time inversion $W(t) = tB(\frac{1}{t})$, we get that Brownian motion crosses 0 for arbitrarily small values of t.

Letting $Z_B = \{t : B(t) = 0\}$, this means that 0 is an accumulation point from the right for Z_B. But we get even more.

For a function f, define the upper and lower right derivatives

$$\begin{aligned} D^* f(t) &= \limsup_{h \downarrow 0} \frac{f(t+h) - f(t)}{h}, \\ D_* f(t) &= \liminf_{h \downarrow 0} \frac{f(t+h) - f(t)}{h}. \end{aligned}$$

Corollary 2.6. *Fix $t_0 \geq 0$. Brownian motion W almost surely satisfies $D^*W(t_0) = +\infty$, $D_*W(t_0) = -\infty$, and t_0 is an accumulation point from the right for the level set $\{s : W(s) = W(t_0)\}$.*

Proof. We have

$$D^*W(0) \geq \limsup_{n\to\infty} \frac{W(\frac{1}{n}) - W(0)}{\frac{1}{n}} \geq \limsup_{n\to\infty} \sqrt{n}\, W(\tfrac{1}{n}) = \limsup_{n\to\infty} \frac{B(n)}{\sqrt{n}}$$

which is infinite by Proposition 2.5.

Similarly, $D_*W(0) = -\infty$, showing that W is not differentiable at 0. $\square$

This argument shows that for any fixed time t, almost every Brownian motion is not differentiable at time t.

Using Fubini's theorem, this implies that for almost every Brownian motion, the set of times where is not differentiable has full measure in $[0, \infty)$.

This is not the same as saying Brownian motion is almost surely **nowhere** differentiable. That statement requires more work to prove.

Similarly, for almost all t the set $\{s : B(s) = B(t)\}$ has no isolated points. But this is not true of all t, e.g., times when $B(t)$ is a local maximum.

We leave it to the reader to show that for all t,

$$\mathbb{P}(t \text{ is a local maximum}) = 0,$$

but almost surely local maxima are a countable dense set in $(0, \infty)$.

Theorem 2.7. *Almost surely Brownian motion is nowhere differentiable. Furthermore, almost surely for all* t *either* $D^*B(t) = +\infty$ *or* $D_*B(t) = -\infty$.

For local maxima we have $D^*B(t) \le 0$, and for local minima, $D_*B(t) \ge 0$, so it is important to have the either-or in the statement.

The result is due to Paley, Weiner and Zygmund in 1933.

We give a proof due to Dvoretzky, Erdős and Kakutani from 1961.

Proof. Suppose that there is a $t_0 \in [0,1]$ such that $-\infty < D_*B(t_0) \leq D^*B(t_0) < \infty$. Then for some finite constant M we would have

$$\sup_{h\in[0,1]} \frac{|B(t_0+h)-B(t_0)|}{h} \leq M. \tag{2.4}$$

If t_0 is contained in the binary interval $[(k-1)/2^n, k/2^n]$ for $n > 2$, then for all $1 \leq j \leq n$ the triangle inequality gives

$$|B\left((k+j)/2^n\right) - B\left((k+j-1)/2^n\right)| \leq M(2j+1)/2^n. \tag{2.5}$$

Let $\Omega_{n,k}$ be the event that (2.5) holds for $j = 1$, 2, and 3. Then by the scaling property

$$\mathbb{P}(\Omega_{n,k}) \le \mathbb{P}\left(|B(1)| \le 7M/\sqrt{2^n}\right)^3,$$

which is at most $(7M2^{-n/2})^3$, since the normal density is less than 1/2. Hence

$$\mathbb{P}\left(\bigcup_{k=1}^{2^n} \Omega_{n,k}\right) \le 2^n(7M2^{-n/2})^3 = (7M)^3 2^{-n/2},$$

Therefore by the Borel–Cantelli Lemma,

$$\mathbb{P}\left((2.4) \text{ holds}\right) \le \mathbb{P}\left(\bigcup_{k=1}^{2^n} \Omega_{n,k} \text{ holds for infinitely many } n\right) = 0. \quad \square$$

3. Hausdorff dimension

The **Hausdorff dimension** of K is defined to be

$$\dim(K) = \inf\{\alpha : \mathcal{H}^\alpha_\infty(K) = 0\}.$$

More generally we define

$$\mathcal{H}^\alpha_\epsilon(K) = \inf\Big\{\sum_i |U_i|^\alpha : K \subset \bigcup_i U_i, |U_i| < \epsilon\Big\},$$

where each U_i is now required to have diameter less than ϵ. The **α-dimensional Hausdorff measure** of K is defined as

$$\mathcal{H}^\alpha(K) = \lim_{\epsilon\to 0} \mathcal{H}^\alpha_\epsilon(K).$$

This is an outer measure; an **outer measure** on a nonempty set X is a function μ^* from the family of subsets of X to $[0, \infty]$ that satisfies

- $\mu^*(\emptyset) = 0$,
- $\mu^*(A) \leq \mu^*(B)$ if $A \subset B$,
- $\mu^*(\bigcup_{j=1}^{\infty} A_j) \leq \sum_{j=1}^{\infty} \mu^*(A_j)$.

When $\alpha = d \in \mathbb{N}$, then $\mathcal{H}^\alpha$ is a constant multiple of $\mathcal{L}_d$, d-dimensional Lebesgue measure.

The construction of Hausdorff measure can be made a little more general by considering a positive, increasing function φ on $[0, \infty)$ with $\varphi(0) = 0$. This is called a **gauge function** and we may associate to it the Hausdorff content

$$\mathcal{H}^{\varphi}_{\infty}(K) = \inf\{\sum_i \varphi(|U_i|) : K \subset \bigcup_i U_i\};$$

then $\mathcal{H}^{\varphi}_{\epsilon}(K)$, and $\mathcal{H}^{\varphi}(K) = \lim_{\epsilon \to 0} \mathcal{H}^{\varphi}_{\epsilon}(K)$ are defined as before.

Lemma 3.1. *If* $\mathcal{H}^\alpha(K) < \infty$ *then* $\mathcal{H}^\beta(K) = 0$ *for any* $\beta > \alpha$*.*

Proof. It follows from the definition of $\mathcal{H}^\alpha_\epsilon$ that

$$\mathcal{H}^\beta_\epsilon(K) \leq \epsilon^{\beta-\alpha}\mathcal{H}^\alpha_\epsilon(K),$$

which gives the desired result as $\epsilon \to 0$. $\square$

Thus if we think of $\mathcal{H}^\alpha(K)$ as a function of α, the graph of $\mathcal{H}^\alpha(K)$ versus α shows that there is a critical value of α where $\mathcal{H}^\alpha(K)$ jumps from ∞ to 0. This critical value is equal to the Hausdorff dimension of the set.

Lemma 3.2 (Mass Distribution Principle)**.** *If E supports a strictly positive Borel measure μ that satisfies*

$$\mu(B(x,r)) \leq Cr^{\alpha},$$

for some constant $0 < C < \infty$ and for every ball $B(x,r)$, then $\mathcal{H}^{\alpha}(E) \geq \mathcal{H}^{\alpha}_{\infty}(E) \geq \mu(E)/C$. In particular, $\dim(E) \geq \alpha$.

Proof. Let $\{U_i\}$ be a cover of E. For $\{r_i\}$, where $r_i = |U_i|$, we look at the following cover: choose x_i in each U_i, and take open balls $B(x_i, r_i)$. Then

$$\mu(U_i) \leq \mu(B(x_i, r_i)) \leq C{r_i}^{\alpha} = C|U_i|^{\alpha}.$$

whence

$$\sum_i |U_i|^{\alpha} \geq \sum_i \frac{\mu(U_i)}{C} \geq \frac{\mu(E)}{C}.$$

Thus $\mathcal{H}^{\alpha}(E) \geq \mathcal{H}^{\alpha}_{\infty}(E) \geq \mu(E)/C$. □

Lemma 3.3. *If f is α-Hölder on $X \to Y$ and $A \subset X$ then*

$$\dim(f(A)) \leq \dim(A)/\alpha \leq \min(\dim(A)/\alpha, \dim(Y)).$$

Proof. If $\{U_j\}$ is a cover of A, then $\{f(U_j\}$ covers $f(A)$ and

$$\mathrm{diam}(f(U_j)) \leq C\mathrm{diam}(U_j)^\alpha.$$

so

$$\sum \mathrm{diam}(f(U_j))^{s/\alpha} \lesssim \sum \mathrm{diam}((U_j))^s \to 0$$

for $s \geq \dim(A)$. $\square$

Since Brownian motion is Hölder of every order $\beta < 1/2$ we have:

Corollary 3.4.

$$\dim(G_B) \leq \underline{\dim}_{\mathcal{M}}(G_B) \leq \overline{\dim}_{\mathcal{M}}(G_B) \leq 3/2 \quad a.s.$$

Corollary 3.5. *For* $A \subset [0, \infty)$, *we have* $\dim(B(A)) \leq (2 \dim(A)) \wedge 1$ $a.s.$

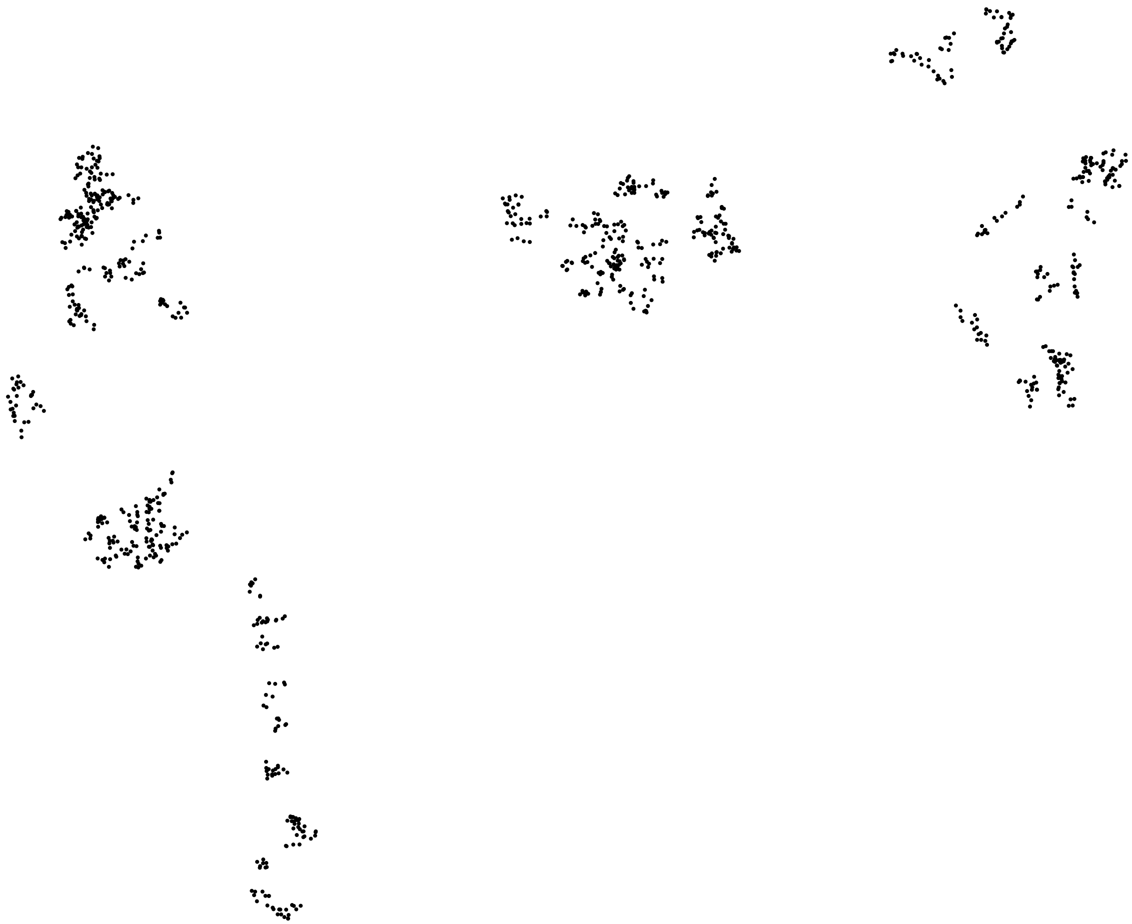
Theorem 3.6 (McKean). *For $A \subset [0, \infty)$, the image of A under d-dimensional Brownian motion has Hausdorff dimension $(2 \dim A) \wedge d$ almost surely.*

Theorem 3.7 (Uniform Dimension Doubling, Kaufman). *Let B be Brownian motion in dimension at least 2. Almost surely, for every $A \subset [0, \infty)$, we have* $\dim B(A) = 2 \dim(A)$.

Notice the difference between these results. In McKean's Theorem, the null probability set depends on A, while Kaufman's Theorem has a much stronger claim: it states dimension doubling uniformly for all sets.

Kaufman's theorem fails for $d = 1$. The zero set of 1-dimensional Brownian motion has dimension half, while its image is the single point 0.

We will prove McKean's theorem for $A = [0, 1]$.



The image of the middle thirds Cantor set under a 2-dimensional Brownian path. By dimension doubling, this set has Hausdorff dimension $\log 4/\log 3$.

Theorem 3.8 (Frostman's Energy Method)**.** *Given a metric space* (X, ρ), *if* μ *is a finite Borel measure supported on* $A \subset X$ *and*

$$\mathcal{E}_\alpha(\mu) \stackrel{\text{def}}{=} \iint \frac{d\mu(x)d\mu(y)}{\rho(x,y)^\alpha} < \infty,$$

then $\mathcal{H}^\alpha_\infty(A) = \infty$, *and hence* $\dim(A) \geq \alpha$.

Proof. For sufficiently large M, the set

$$K_1 = \{x \in K : \int_K \frac{d\mu(y)}{|x-y|^\alpha} \leq M\}$$

has positive μ measure. One can write this integral as a sum and use summation by parts to get,

$$\begin{aligned}
\int_K \frac{d\mu(y)}{|x-y|^\alpha} &\geq \sum_{n=-r}^{\infty} \mu(\{2^{-n-1} \leq |y-x| \leq 2^{-n}\})2^{n\alpha} \\
&= \sum_{n=-r}^{\infty} \Big(\mu(B(x,2^{-n})) - \mu(B(x,2^{-n-1}))\Big)2^{n\alpha} \\
&= C_1 + \sum_{n=-r}^{\infty} \Big((2^{n\alpha} - 2^{(n-1)\alpha})\mu(B(x,2^{-n}))\Big) \\
&= C_1 + C_2 \sum_{n=-r}^{\infty} 2^{n\alpha}\mu(B(x,2^{-n})).
\end{aligned}$$

If $x \in K_1$, then the integral, and thus the sum, is finite, so

$$\lim_{n\to\infty} \frac{\mu(B(x, 2^{-n}))}{2^{-n\alpha}} = 0.$$

The Mass Distribution Principle implies $\mathcal{H}^\alpha(K_1) = \infty$, hence $\mathcal{H}^\alpha(K) = \infty$. $\square$

Theorem 3.9. *Brownian trace $B_d([0,1])$ in $\mathbb{R}^d$ has dimension 2.*

Proof. From Corollary 2.2 we have that B_d is β Hölder for every $\beta < 1/2$ a.s. Therefore $\dim B_d[0,1] \leq 2$ almost surely.

For the other inequality, we will use Frostman's Energy Method. A natural measure on $B_d[0,1]$ is the occupation measure $\mu_B \stackrel{\text{def}}{=} \mathcal{L}B^{-1}$, which means that $\mu_B(A) = \mathcal{L}B^{-1}(A)$, for all measurable subsets A of $\mathbb{R}^d$, or, equivalently,

$$\int_{\mathbb{R}^d} f(x)\, d\mu_B(x) = \int_0^1 f(B_t)\, dt$$

for all measurable functions f.

Note that by definition,

$$\mathbb{E}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\frac{d\mu_B(x)d\mu_B(y)}{|x-y|^\alpha}=\mathbb{E}\int_0^1\int_0^1\frac{dsdt}{|B(t)-B(s)|^\alpha} \tag{3.1}$$

We want to show that for any $0<\alpha<2$, this is finite.

Let us evaluate the right hand side expectation:

$$\mathbb{E}|B(t)-B(s)|^{-\alpha}=\mathbb{E}((|t-s|^{1/2}|Z|)^{-\alpha})=|t-s|^{-\alpha/2}\int_{\mathbb{R}^d}\frac{c_d}{|z|^\alpha}e^{-|z|^2/2}\,dz.$$

Here Z denotes the d-dimensional standard Gaussian random variable.

The integral can be evaluated using polar coordinates, but all we need is that, for $d \geq 2$, it is a finite constant c depending on d and α only. Finiteness can be checked by considering $|z| < 1$ and $|z| \geq 1$ separately (details left to the reader).

Substituting c into (3.1) and using Fubini's Theorem we get

$$\mathbb{E}\mathcal{E}_\alpha(\mu_B) = c \int_0^1 \int_0^1 \frac{dsdt}{|t-s|^{\alpha/2}} \leq 2c \int_0^1 \frac{du}{u^{\alpha/2}} < \infty. \tag{3.2}$$

Therefore $\mathcal{E}_\alpha(\mu_B) < \infty$ almost surely and we are done by Frostman's method.

$\square$

Theorem 3.10. *The graph of Brownian motion in* $\mathbb{R}$ *has dimension* $3/2$.

Proof. We have shown in Corollary 3.4 that

$$\dim G_B \leq 3/2.$$

For the other inequality, let $\alpha < 3/2$ and let A be a subset of the graph. Define a measure on the graph using projection to the time axis:

$$\mu(A) \stackrel{\text{def}}{=} \mathcal{L}(\{0 \leq t \leq 1 : (t, B(t)) \in A\}).$$

Changing variables, the α energy of μ can be written as

$$\iint \frac{d\mu(x)d\mu(y)}{|x-y|^\alpha} = \int_0^1 \int_0^1 \frac{dsdt}{(|t-s|^2 + |B(t)-B(s)|^2)^{\alpha/2}}.$$

Bounding the integrand, taking expectations, and applying Fubini's Theorem we get that

$$\mathbb{E}\mathcal{E}_\alpha(\mu) \leq 2\int_0^1 \mathbb{E}\left((t^2 + B(t)^2)^{-\alpha/2}\right) dt. \tag{3.3}$$

Let $n(z)$ denote the standard normal density. By scaling, the expected value above can be written as

$$2\int_0^{+\infty} (t^2 + tz^2)^{-\alpha/2} n(z)\, dz. \tag{3.4}$$

Cut the integral into $z \le \sqrt{t}$ and $z > \sqrt{t}$. Then (3.4) is bounded above by twice

$$\int_0^{\sqrt{t}} (t^2)^{-\alpha/2}\, dz + \int_{\sqrt{t}}^{\infty} (tz^2)^{-\alpha/2} n(z)\, dz = t^{\frac{1}{2}-\alpha} + t^{-\alpha/2} \int_{\sqrt{t}}^{\infty} z^{-\alpha} n(z)\, dz.$$

Furthermore, we separate the last integral at 1. We get

$$\int_{\sqrt{t}}^{\infty} z^{-\alpha} n(z)\, dz \le c_\alpha + \int_{\sqrt{t}}^{1} z^{-\alpha}\, dz.$$

The latter integral is of order $t^{(1-\alpha)/2}$. Substituting these results into (3.3), we see that the expected energy is finite when $\alpha < 3/2$. Therefore $\mathcal{E}_\alpha(\mu_B) < \infty$ almost surely. The claim now follows from Frostman's Energy Method. $\square$

4. Stopping times and the Markov property

A filtration on a probability space $(\Omega, \mathcal{F}, P)$ is a family $\{\mathcal{F}(t) : t > 0\}$ of σ-algebras such that $\mathcal{F}(s) \subset \mathcal{F}(t) \subset \mathcal{F}$ for all $s < t$.

A probability space together with a filtration is called a filtered probability space.

A stochastic process $\{X(t) : t > 0\}$ defined on a filtered probability space with filtration $\{\mathcal{F}(t) : t > 0\}$ is called adapted if $X(t)$ is $\mathcal{F}(t)$-measurable for any $t > 0$.

Suppose we have a Brownian motion $\{B(t) : t > 0\}$ defined on some probability space, then we can define a filtration $\{\mathcal{F}_0(t) : t > 0\}$ by letting $\mathcal{F}0(t) = \sigma(B(s) : 0 < s < t)$ be the σ-algebra generated by the random variables $B(s)$, for $0 < s < t$. With this definition, the Brownian motion is obviously adapted to the filtration.

Intuitively, this σ-algebra contains all the information available from observing the process up to time t.

For each $t \geq 0$ let $\mathcal{F}_0(t) = \sigma\{B(s) : s \leq t\}$ be the smallest σ-field making every $B(s)$, $s \leq t$, measurable

Set $\mathcal{F}_+(t) = \bigcap_{u>t} \mathcal{F}_0(u)$. This allows us to look "infinitesimally" into the future. This is a right-continuous filtration; these are sometimes more convenient for technical reasons.

If E is closed, then $\inf\{t \in B(t) \in E\}$ is in $\mathcal{F}_0(t)$.

If U is open, then $\inf\{t \in B(t) \in U\}$ is in $\mathcal{F}_+(t)$, but need not be in $\mathcal{F}_0(t)$.

Proposition 4.1 (Markov property)**.** *For every $t \geq 0$ the process*

$$\{B(t+s) - B(t)\}_{s \geq 0}$$

is standard Brownian motion independent of $\mathcal{F}_0(t)$ and $\mathcal{F}_+(t)$.

It is evident from independence of increments that $\{B(t+s)-B(t)\}_{s\geq 0}$ is standard Brownian motion independent of $\mathcal{F}_0(t)$. That this process is independent of $\mathcal{F}_+(t)$ follows from continuity.

Corollary 4.2. *The process* $\{B(t), t > 0\}$ *is independent of the* σ*-algebra* $\mathcal{F}_+(0)$

Theorem 4.3. *(Blumenthal's 0-1 law) Let* $A \in \mathcal{F}_+(0)$. *Then* $\mathbb{P}(A) \in \{0, 1\}$.

Proof. Using the corollary above, for any $A \in \sigma(B(t) : t > 0)$ is independent of $\mathcal{F}_+(0)$. But $\mathcal{F}_+(0) \subset \sigma(B(t) : t > 0)$, so A is independent of itself, i.e., $\mathbb{P}(A) = \mathbb{P}(A \cap A) = \mathbb{P}(A)^2$ and hence has probability zero or one. $\square$

We can use this to give another proof of something we proved earlier:

Corollary 4.4. *Let* $\tau = \inf\{t > 0 : B(t) > 0\}$ *and* $\sigma = \inf\{t > 0 : B(t) < 0\}$*. Then* $\mathbb{P}(\tau = 0) = \mathbb{P}(\sigma = 0) = 1$*.*

Proof. The event

$$\{\tau = 0\} = \bigcap_n \{\exists\ 0 < \epsilon < 1/n \text{ with } B(\epsilon) > 0\}.$$

is clearly in $\mathcal{F}_+(0)$. Thus it suffices to show this event has positive probability. But for $t > 0$, then $\mathbb{P}(\tau < t) > \mathbb{P}(B(t) > 0) = 1/2$. Hence $\mathbb{P}(\tau = 0) = 1$.

The argument works replacing $B(t) > 0$ by $B(t) < 0$ to show $\mathbb{P}(\sigma = 0) = 1$. $\square$

A random variable τ is a **stopping time** for a Brownian filtration $\{\mathcal{F}(t)\}_{t\geq 0}$ if $\{\tau \leq t\} \in \mathcal{F}(t)$ for all t. For any random time τ we define the pre-τ σ-field

$$\mathcal{F}(\tau) := \{A : \forall t,\ A \cap \{\tau \leq t\} \in \mathcal{F}(t)\}.$$

Roughly speaking, we stop a Brownian motion at time t based behavior up to time t.

A simple example of a stopping time is the time when a Brownian motion first enters a closed set. More generally, if A is a Borel set then the hitting time τ_A is a stopping time.

Theorem 4.5. *Suppose that τ is a stopping time for the Brownian filtration $\{\mathcal{F}(t)\}_{t\geq 0}$. Then $\{B(\tau+s)-B(\tau)\}_{s\geq 0}$ is Brownian motion independent of $\mathcal{F}(\tau)$.*

Sketch of Proof. Suppose first that τ is an integer valued stopping time with respect to a Brownian filtration $\{\mathcal{F}(t)\}_{t\geq 0}$. For each integer j the event $\{\tau = j\}$ is in $\mathcal{F}(j)$ and the process $\{B(t+j)-B(j)\}_{t\geq 0}$ is independent of $\mathcal{F}(j)$, so the result follows from the Markov property in this special case. It also holds if the values of τ are integer multiples of some $\varepsilon > 0$, and approximating τ by such discrete stopping times gives the conclusion in the general case. $\square$

The following is an elementary fact we need below:

Lemma 4.6. *Let X, Y, Z be random variables with X, Y independent and X, Z independent. If $Y \stackrel{\mathrm{d}}{=} Z$ then $(X, Y) \stackrel{\mathrm{d}}{=} (X, Z)$.*

Theorem 4.7 (Reflection Principle)**.** *If τ is a stopping time then*

$$B^*(t) := B(t)\mathbf{1}_{(t\le\tau)} + (2B(\tau) - B(t))\mathbf{1}_{(t>\tau)}$$

(Brownian motion reflected at time τ) is also standard Brownian motion.

Proof. The strong Markov property states that $\{B(\tau+t)-B(\tau)\}_{t\ge0}$ is Brownian motion independent of $\mathcal{F}(\tau)$, and by symmetry this also holds for

$$\{-(B(\tau+t) - B(\tau))\}_{t\ge0}.$$

We see from Lemma 4.6 that

$$\begin{aligned}(\{B(t)\}_{0\le t\le\tau}, \{B(t+\tau) - B(\tau)\}_{t\ge0}) \\ \overset{\mathrm{d}}{=} (\{B(t)\}_{0\le t\le\tau}, \{(B(\tau) - B(t+\tau))\}_{t\ge0}),\end{aligned}$$

and the reflection principle follows immediately. $\square$

Set $M(t) = \max_{0 \le s \le t} B(s)$.

Theorem 4.8. *If* $a > 0$, *then* $\mathbb{P}[M(t) > a] = 2\mathbb{P}[B(t) > a] = \mathbb{P}[|B(t)| > a]$.

Proof. Set $\tau_a = \min\{t \ge 0 : B(t) = a\}$ and let $\{B^*(t)\}$ be Brownian motion reflected at τ_a. Then $\{M(t) > a\}$ is the disjoint union of the events $\{B(t) > a\}$ and $\{M(t) > a, B(t) \le a\}$, and since $\{M(t) > a, B(t) \le a\} = \{B^*(t) \ge a\}$ the desired conclusion follows immediately. □

Lemma 4.9 (Wald's Lemma for Brownian Motion)**.** *Let τ be a stopping time for Brownian motion in $\mathbb{R}^d$.*

(i) $\mathbb{E}[|B_\tau|^2] = d\,\mathbb{E}[\tau]$ *(possibly both ∞),*

(ii) *if $\mathbb{E}[\tau] < \infty$, then $\mathbb{E}[B_\tau] = 0$.*

Proof. (i) We break the proof of (i) into three steps:

(a) τ is integer valued and bounded,

(b) τ is integer valued and unbounded and

(c) general τ.

(a) Let τ be bounded with integer values. Write

$$|B_\tau|^2 = \sum_{k=1}^{\infty} (|B_k|^2 - |B_{k-1}|^2) \mathbf{1}_{\tau \geq k}. \tag{4.1}$$

If $Z_k = B_k - B_{k-1}$ and $\mathcal{F}_t$ is the σ-field determined by Brownian motion in $[0, t]$, then

$$\mathbb{E}[|B_k|^2 - |B_{k-1}|^2 | \mathcal{F}_{k-1}] = \mathbb{E}[|Z_k|^2 + 2B_{k-1} \cdot Z_k | \mathcal{F}_{k-1}] = d.$$

The expectation of $|Z_k|^2$ is d, since Z_k is Brownian motion run for time 1. Given the conditioning, B_{k-1} is constant and Z_k has mean zero, so their product has zero expectation. Since the event $\{k \leq \tau\}$ is $\mathcal{F}_{k-1}$ measurable, we deduce that $\mathbb{E}[(|B_k|^2 - |B_{k-1}|^2) \mathbf{1}_{k \leq \tau}] = d\, \mathbb{P}(k \leq \tau)$. Therefore by (4.1),

$$\mathbb{E}(|B_\tau|^2) = d \sum_{k=1}^{\infty} \mathbb{P}(k \leq \tau) = d\mathbb{E}(\tau).$$

(b) Next, suppose that τ is integer valued but unbounded. Applying (4.1) to $\tau \wedge n$ and letting $n \to \infty$ yields $\mathbb{E}|B_\tau|^2 \leq d\mathbb{E}\tau$ by Fatou's Lemma.

On the other hand, the strong Markov property yields independence of $B_{\tau\wedge n}$ and $B_\tau - B_{\tau\wedge n}$, which implies that $\mathbb{E}(|B_\tau|^2) \geq d\mathbb{E}(\tau \wedge n)$. Letting $n \to \infty$ proves (4.1) in this case. By scaling, (4.1) also holds if τ takes values that are multiples of a fixed ϵ.

(c) Now suppose just that $\mathbb{E}\tau < \infty$ and write $\tau_\ell := 2^{-\ell}\lceil \tau 2^\ell \rceil$. Then (4.1) holds for the stopping times τ_ℓ, which decrease to τ as $\ell \to \infty$. Since we have $0 \le \tau_\ell - \tau \le 2^{-\ell}$ and $\mathbb{E}[|B(\tau_\ell) - B(\tau)|^2] \le 2^{-\ell}$ by the strong Markov property, (4.1) follows.

(ii) We also break the proof of (ii) into the integer-valued and general cases.

(a) Suppose τ is integer valued and has finite expectation. Then

$$B_\tau = \sum_{k=1}^{\infty} (B_k - B_{k-1}) \mathbf{1}_{\tau \geq k}.$$

The terms in the sum are orthogonal in L^2 (with respect to the Wiener measure), and the second moment of the kth term is $d\mathbb{P}(\tau \geq k)$.

The second moment of the sum is the sum of the second moments:

$$\mathbb{E}[B_\tau^2] = d \sum_{k=1}^{\infty} \mathbb{P}(\tau \geq k) = d \sum_{j=1}^{\infty} j\, \mathbb{P}(\tau = j) = d\mathbb{E}[\tau] < \infty.$$

Thus

$$\mathbb{E}[B_\tau] = \sum_{k=1}^{\infty} \mathbb{E}[(B_k - B_{k-1})\mathbf{1}_{\tau \geq k}],$$

since taking the expectation of an L^2 function is just taking the inner product with the constant 1 function, and this is continuous with respect to convergence in L^2. Finally, every term on the right-hand side is 0 by independence of increments.

(b) We apply Lebesgue dominated convergence to deduce the general case.

Suppose τ is any stopping time with finite expectation and define $\tau_\ell := 2^{-\ell}\lceil 2^\ell \tau\rceil$. Note that $B(\tau_\ell) \to B(\tau)$ almost surely; thus we just want to show this sequence is dominated in L^1. Define

$$Y = \max\{|B(\tau + s) - B(\tau_1)| : 0 \le s \le 1\}$$

and note that

$$Y \le 2\max\{|B(\tau + s) - B(\tau)| : 0 \le s \le 1\},$$

by the triangle inequality.

The right-hand side is in L^2 (hence L^1) since we earlier showed the maximum process has the same distribution as $|B_t|$. Moreover $|B(\tau_\ell)| \le |B(\tau_1)| + Y$, for every ℓ, so the sequence $\{B(\tau_\ell)\}_{\ell\ge 1}$ is dominated by a single L^1 function. Thus $\mathbb{E}[B(\tau)] = \lim_\ell \mathbb{E}[B(\tau_\ell)]$ and the limit is zero by part (a). $\square$

5. Area of 2-dimensional Brownian motion

Lemma 5.1. *If $A_1, A_2 \subset \mathbb{R}^2$ are Borel sets with positive area, then*

$$\mathcal{L}_2(\{x \in \mathbb{R}^2 : \mathcal{L}_2(A_1 \cap (A_2 + x)) > 0\}) > 0.$$

Proof. We may assume A_1 and A_2 are bounded. By Fubini's Theorem,

$$\begin{aligned}
\int_{\mathbb{R}^2} \mathbf{1}_{A_1} * \mathbf{1}_{-A_2}(x)\, dx &= \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} \mathbf{1}_{A_1}(w) \mathbf{1}_{A_2}(w - x)\, dw\, dx \\
&= \int_{\mathbb{R}^2} \mathbf{1}_{A_1}(w) \left(\int_{\mathbb{R}^2} \mathbf{1}_{A_2}(w - x)\, dx \right) dw \\
&= \mathcal{L}_2(A_1) \mathcal{L}_2(A_2) \\
&> 0.
\end{aligned}$$

Thus $\mathbf{1}_{A_1} * \mathbf{1}_{-A_2}(x) > 0$ on a set of positive area. But $\mathbf{1}_{A_1} * \mathbf{1}_{-A_2}(x)$ is exactly the area of $A_1 \cap (A_2 + x)$, so this proves the Lemma. $\square$

Theorem 5.2. *Almost surely* $\mathcal{L}_2(B[0,1]) = 0$.

Proof. Let X denote the area of $B[0,1]$, and M be its expected value. First we check that $M < \infty$. If $a \geq 1$ then

$$\mathbb{P}[X > a] \leq 2\mathbb{P}[|W(t)| > \sqrt{a}/2 \text{ for some } t \in [0,1]\,] \leq 8e^{-a/8}$$

where W is standard one-dimensional Brownian motion. Thus

$$M = \int_0^\infty \mathbb{P}[X > a]\, da \leq 8 \int_0^\infty e^{-a/8}\, da + 1 < \infty.$$

Thus the area is finite, almost surely.

Note that $B(3t)$ and $\sqrt{3}B(t)$ have the same distribution, and hence

$$\mathbb{E}\mathcal{L}_2(B[0,3]) = 3\mathbb{E}\mathcal{L}_2(B[0,1]) = 3M.$$

Note that we have

$$\mathcal{L}_2(B[0,3]) \leq \sum_{j=0}^{2} \mathcal{L}_2(B[j,j+1])$$

with equality if and only if for $0 \leq i < j \leq 2$ we have

$$\mathcal{L}_2(B[i,i+1] \cap B[j,j+1]) = 0.$$

On the other hand, for $j = 0, 1, 2$, we have $\mathbb{E}\mathcal{L}_2(B[j, j+1]) = M$ and

$$3M = \mathbb{E}\mathcal{L}_2(B[0,3]) \leq \sum_{j=0}^{2} \mathbb{E}\mathcal{L}_2(B[j, j+1]) = 3M,$$

whence the intersection of any two of the $B[j, j+1]$ has measure zero almost surely. In particular, $\mathcal{L}_2(B[0,1] \cap B[2,3]) = 0$ almost surely.

For $x \in \mathbb{R}^2$, let $R(x)$ denote the area of $B[0,1] \cap (x + B[2,3] - B(2) + B(1))$.

If we condition on the values of $B[0,1]$, $B[2,3] - B(2)$, then in order to evaluate the expected value of $\mathcal{L}_2(B[0,1] \cap B[2,3])$ we should integrate $R(x)$ where x has the distribution of $B(2) - B(1)$. Thus

$$0 = \mathbb{E}[\mathcal{L}_2(B[0,1] \cap B[2,3])] = (2\pi)^{-1} \int_{\mathbb{R}^2} e^{-|x|^2/2} \mathbb{E}[R(x)] \, dx,$$

where we average with respect to the Gaussian distribution of $B(2) - B(1)$.

Thus $R(x) = 0$ a.s. for $\mathcal{L}_2$-almost all x, or, by Fubini's Theorem, the area of the set where $R(x)$ is positive is a.s. zero. From the lemma we get that a.s.

$$\mathcal{L}_2(B[0,1]) = 0 \quad \text{or} \quad \mathcal{L}_2(B[2,3]) = 0.$$

The observation that $\mathcal{L}_2(B[0,1])$ and $\mathcal{L}_2(B[2,3])$ are identically distributed and independent completes the proof that $\mathcal{L}_2(B[0,1]) = 0$ almost surely. $\square$

This also follows from the fact that Brownian motion has probability zero of hitting a given point (other than its starting point), a fact we will prove using potential theory later.

6. The Law of the Iterated Logarithm

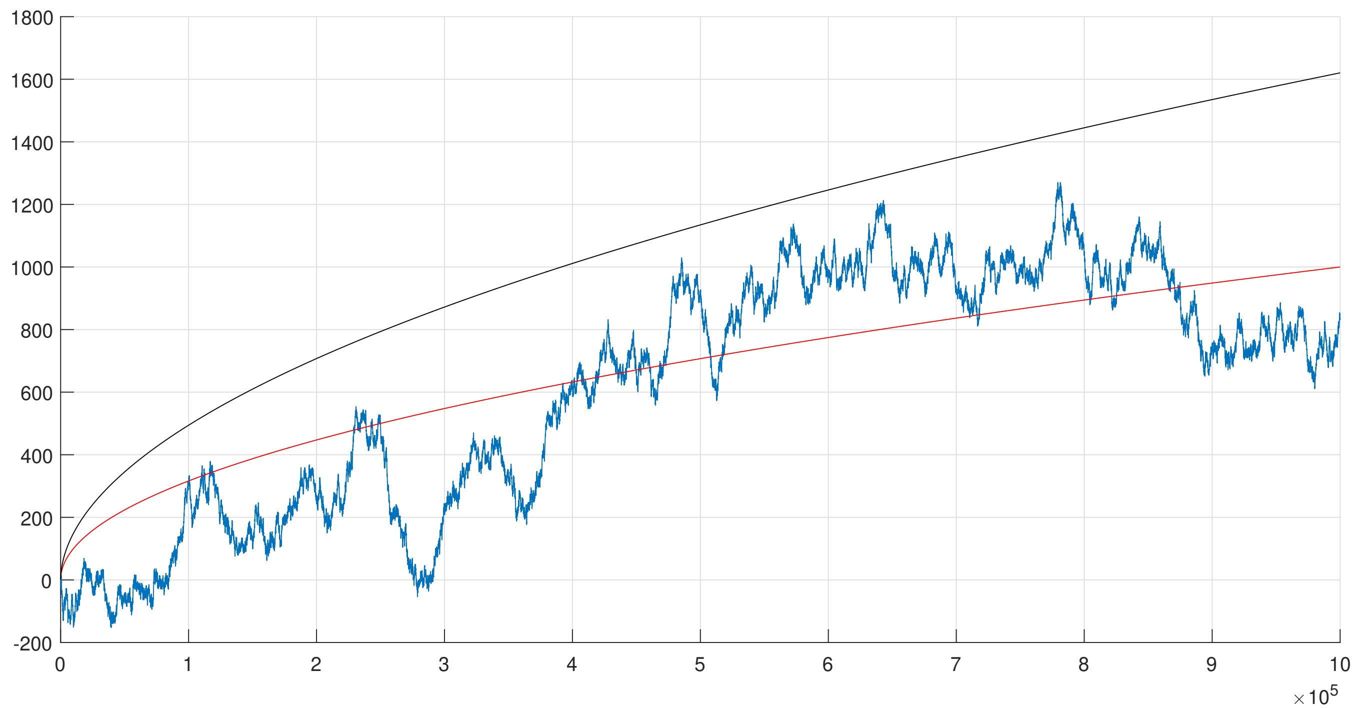
Theorem 6.1 (The Law of the Iterated Logarithm). *For* $\psi(t) = \sqrt{2t \log \log t}$

$$\limsup_{t \to \infty} \frac{B(t)}{\psi(t)} = 1 \quad a.s.$$

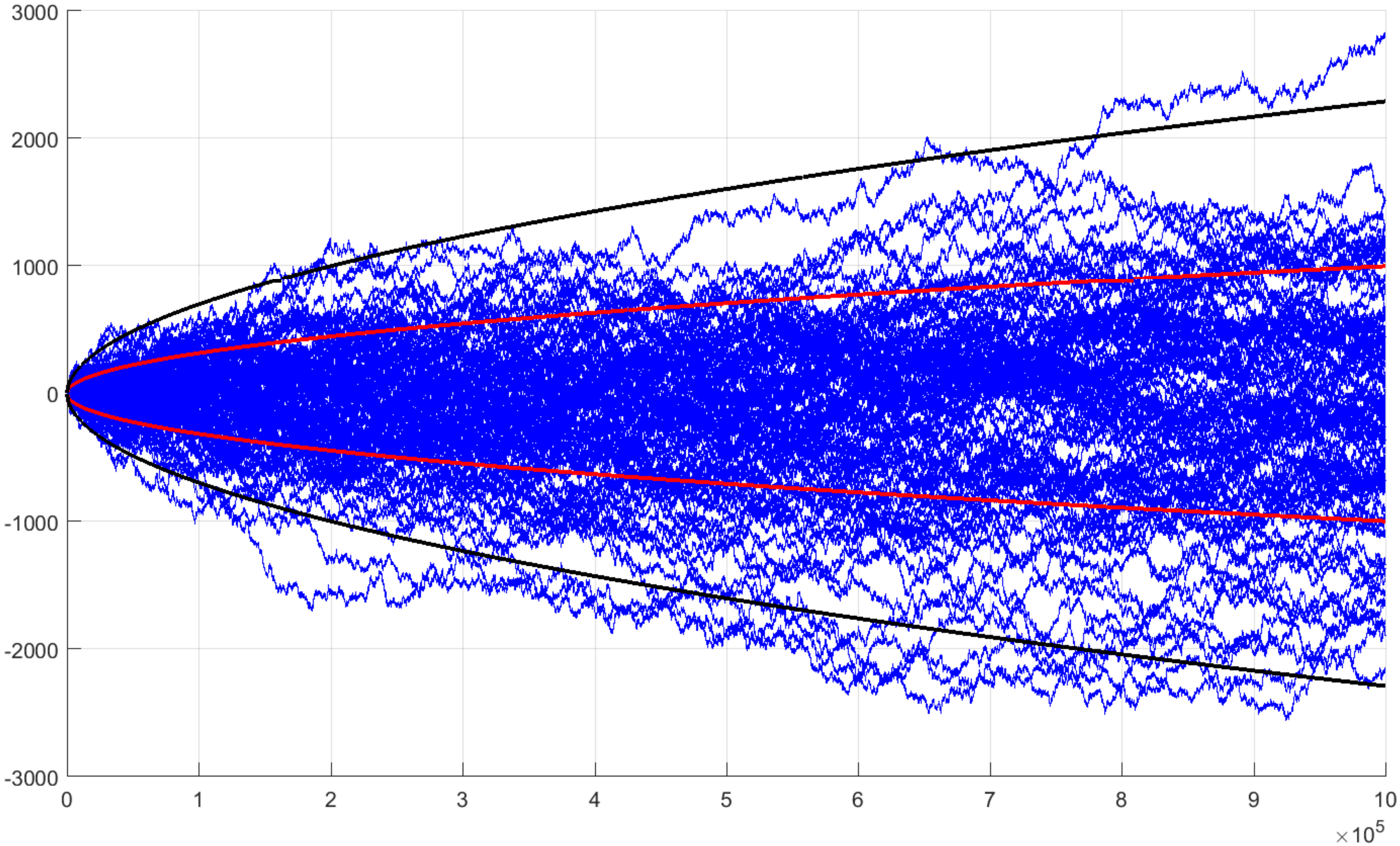
By symmetry it follows that

$$\liminf_{t \to \infty} \frac{B(t)}{\psi(t)} = -1 \quad \text{a.s.}$$

Khinchin (1924) proved the Law of the Iterated Logarithm for simple random walks, Kolmogorov (1929) for other walks, and Lévy for Brownian motion. The proof for general random walks is much simpler through Brownian motion than directly.



1800
1600
1400
1200
1000
800
600
400
200
0
-200
0
1
2
3
4
5
6
7
8
9
10
$\times 10^5$



3000
2000
1000
0
-1000
-2000
-3000
0
1
2
3
4
5
6
7
8
9
10
$\times 10^5$

Proof. The main idea is to first consider a geometric sequence of times, and later "fill in" between these times.

We start by proving the upper bound. Fix $\epsilon > 0$ and $q > 1$. Let

$$A_n = \left\{ \max_{0 \le t \le q^n} B(t) \ge (1+\epsilon)\psi(q^n) \right\}.$$

By Theorem 4.8 the maximum of Brownian motion up to a fixed time t has the same distribution as $|B(t)|$. Therefore

$$\begin{aligned}
\mathbb{P}(A_n) &= \mathbb{P}\left[|B(q^n)| \ge (1+\epsilon)\psi(q^n)\right] \\
&= \mathbb{P}\left[\frac{|B(q^n)|}{\sqrt{q^n}} \ge \frac{(1+\epsilon)\psi(q^n)}{\sqrt{q^n}}\right]
\end{aligned}$$

We use the estimate for a standard normal random variable (Lemma 1.3)

$$\mathbb{P}(Z > x) \leq e^{-x^2/2}$$

for $x > 1$ to conclude that for large n (recall $\psi(t) = \sqrt{2t \log \log t}$):

$$\mathbb{P}(A_n) \leq 2 \exp\left(-(1+\epsilon)^2 \log \log q^n\right) = \frac{2}{(n \log q)^{(1+\epsilon)^2}},$$

which is summable in n.

Since $\sum_n \mathbb{P}(A_n) < \infty$, by the Borel–Cantelli Lemma, only finitely many of these events occur. So for all large enough n, $|B(q^n)| \geq (1+\epsilon)\psi(q^n)$.

This is the LIL upper bound for times q^n. Next we consider $t \in [q^{n-1}, q^n]$.

For large t choose n so that $q^{n-1} \le t < q^n$. We have

$$\frac{B(t)}{\psi(t)} = \frac{B(t)}{\psi(q^n)}\frac{\psi(q^n)}{\psi(t)} \le \frac{B(t)}{\psi(q^n)}\frac{\psi(q^n)}{\psi(q^{n-1})} \le (1+\epsilon)q.$$

Thus

$$\limsup_{t\to\infty} \frac{B(t)}{\psi(t)} \le (1+\epsilon)q \quad \text{a.s.}$$

Since this holds for any $\epsilon > 0$ and $q > 1$ we have proved that

$$\limsup_{t\to\infty} \frac{B(t)}{\psi(t)} \le 1.$$

This is the upper bound in the LIL.

For the lower bound, fix $q > 1$. In order to use the Borel–Cantelli lemma in the other direction, we need to create a sequence of *independent* events.

[**Borel–Cantelli, part 2:**]
If $\{A_i\}$ are pairwise independent, and $\sum_{i=0}^{\infty} \mathbb{P}(A_i) = \infty$, then $\mathbb{P}(A_i \text{ i.o.}) = 1$. where “i.o.” abbreviates “infinitely often”.

Let

$$D_n = \left\{ B(q^n) - B(q^{n-1}) \geq \psi(q^n - q^{n-1}) \right\}.$$

We will now use Lemma 1.3 for large x:

$$\mathbb{P}(Z > x) \geq \frac{ce^{-x^2/2}}{x}.$$

Using this estimate with $x = \psi(q^n - q^{n-1})/\sqrt{q^n - q^{n-1}}$ we get

$$\begin{aligned}
\mathbb{P}(D_n) &= \mathbb{P}\left(Z \geq \frac{\psi(q^n - q^{n-1})}{\sqrt{q^n - q^{n-1}}} \right) \geq c\frac{\exp(-\log\log(q^n - q^{n-1}))}{\sqrt{2\log\log(q^n - q^{n-1})}} \\
&\geq \frac{c\exp(-\log(n\log q))}{\sqrt{2\log(n\log q)}} > \frac{c'}{n\log n}.
\end{aligned}$$

Thus $\sum_n \mathbb{P}(D_n) = \infty$. Hence D_n happens infinitely often with probability 1.

Thus for infinitely many n

$$B(q^n) \geq B(q^{n-1}) + \psi(q^n - q^{n-1}) \geq -2\psi(q^{n-1}) + \psi(q^n - q^{n-1}),$$

where the second inequality follows from applying the previously proven upper bound to $-B(q^{n-1})$.

We claim that that for infinitely many n

$$\frac{B(q^n)}{\psi(q^n)} \geq \frac{-2\psi(q^{n-1}) + \psi(q^n - q^{n-1})}{\psi(q^n)} \geq \frac{-2}{\sqrt{q}} + \frac{q^n - q^{n-1}}{q^n}. \tag{6.1}$$

To obtain the second inequality first note that

$$\frac{\psi(q^{n-1})}{\psi(q^n)} = \frac{\psi(q^{n-1})}{\sqrt{q^{n-1}}} \frac{\sqrt{q^n}}{\psi(q^n)} \frac{1}{\sqrt{q}} \leq \frac{1}{\sqrt{q}}$$

since $\psi(t)/\sqrt{t}$ is increasing in t for large t.

For the second term we just use the fact that $\psi(t)/t$ is decreasing in t. Since $q^n - q^{n-1} < q^n$ this implies

$$\frac{\psi(q^n - q^{n-1})}{q^n - q^{n-1}} \geq \frac{\psi(q^n)}{q^n}$$

$$\frac{\psi(q^n - q^{n-1})}{\psi(q^n)} \leq \frac{q^n - q^{n-1}}{q^n}.$$

Now (6.1) implies that

$$\limsup_{t \to \infty} \frac{B(t)}{\psi(t)} \geq -\frac{2}{\sqrt{q}} + 1 - \frac{1}{q} \quad \text{a.s.}. \tag{6.2}$$

Letting $q \uparrow \infty$ concludes the proof of the lower bound. $\square$

Corollary 6.2. *If* $\{\lambda_n\}$ *is a sequence of random times (not necessarily stopping times) satisfying* $\lambda_n \to \infty$ *and* $\lambda_{n+1}/\lambda_n \to 1$ *almost surely, then*

$$\limsup_{n\to\infty} \frac{B(\lambda_n)}{\psi(\lambda_n)} = 1 \quad a.s.$$

Furthermore, if $\lambda_n/n \to a > 0$ *almost surely, then*

$$\limsup_{n\to\infty} \frac{B(\lambda_n)}{\psi(an)} = 1 \quad a.s.$$

Proof. The upper bound follows from the upper bound for continuous time.

To prove the lower bound, we might run into the problem that λ_n and q^n may not be close for large n; we have to exclude the possibility that λ_n is a sequence of times where the value of Brownian motion is too small.

To get around this problem recall in previous proof we set

$$D_k = \{B(q^k) - B(q^{k-1}) \geq \psi(q^k - q^{k-1})\}$$

. Now define

$$D_k^* = D_k \cap \left\{ \min_{q^k \leq t \leq q^{k+1}} B(t) - B(q^k) \geq -\sqrt{q^k} \right\} \stackrel{\text{def}}{=} D_k \cap \Omega_k.$$

Note that D_k and Ω_k are independent events since the depend on disjoint time intervals.

By Brownian scaling, $\mathbb{P}(\Omega_k) = c_q > 0$ does not depend on k.

By independence, $\mathbb{P}(D_k^*) = c_q \mathbb{P}(D_k)$, so the sum of these probabilities is infinite.

The events $\{D_{2k}^*\}$ are independent (disjoint time intervals), so by the Borel–Cantelli lemma, for infinitely many (even) k,

$$\begin{aligned} \min_{q^k \le t \le q^{k+1}} B(t) &\geq B(q^k) - \sqrt{q^k} \\ &\geq \psi(q^k)\left(1 - \frac{1}{q} - \frac{2}{\sqrt{q}}\right) - \sqrt{q^k}. \end{aligned}$$

Here we have used (6.2).

Now define $n(k) = \min\{n \ : \ \lambda_n > q^k\}$. Since the ratios λ_{n+1}/λ_n tend to 1, it follows that $q^k \le \lambda_{n(k)} < q^{k+1}$ for all large k. Thus for infinitely many k

$$\frac{B(\lambda_{n(k)})}{\psi(\lambda_{n(k)})} \ge \frac{\psi(q^k)}{\psi(\lambda_{n(k)})}\left[1 - \frac{1}{q} - \frac{2}{\sqrt{q}}\right] - \frac{\sqrt{q^k}}{\psi(\lambda_{n(k)})}.$$

Note that as $k \nearrow \infty$,

$$q^k/\lambda_{n(k))} \to 1, \quad \psi(q^k)/\psi(\lambda_{n(k))}) \to 1, \quad \sqrt{q^k}/\psi(q^k) \to 0.$$

Thus

$$\limsup_{n\to\infty} \frac{B(\lambda_n)}{\psi(\lambda_n)} \ge 1 - \frac{1}{q} - \frac{2}{\sqrt{q}}.$$

Since the left-hand side does not depend on q, we are done.

For the last part, note that if $\lambda_n/n \to a > 0$ then $\psi(\lambda_n)/\psi(an) \to 1$. $\square$

Corollary 6.3. *If $\{S_n\}$ is a simple random walk on $\mathbb{Z}$, then almost surely*
$$\limsup_{n\to\infty} \frac{S_n}{\psi(n)} = 1.$$

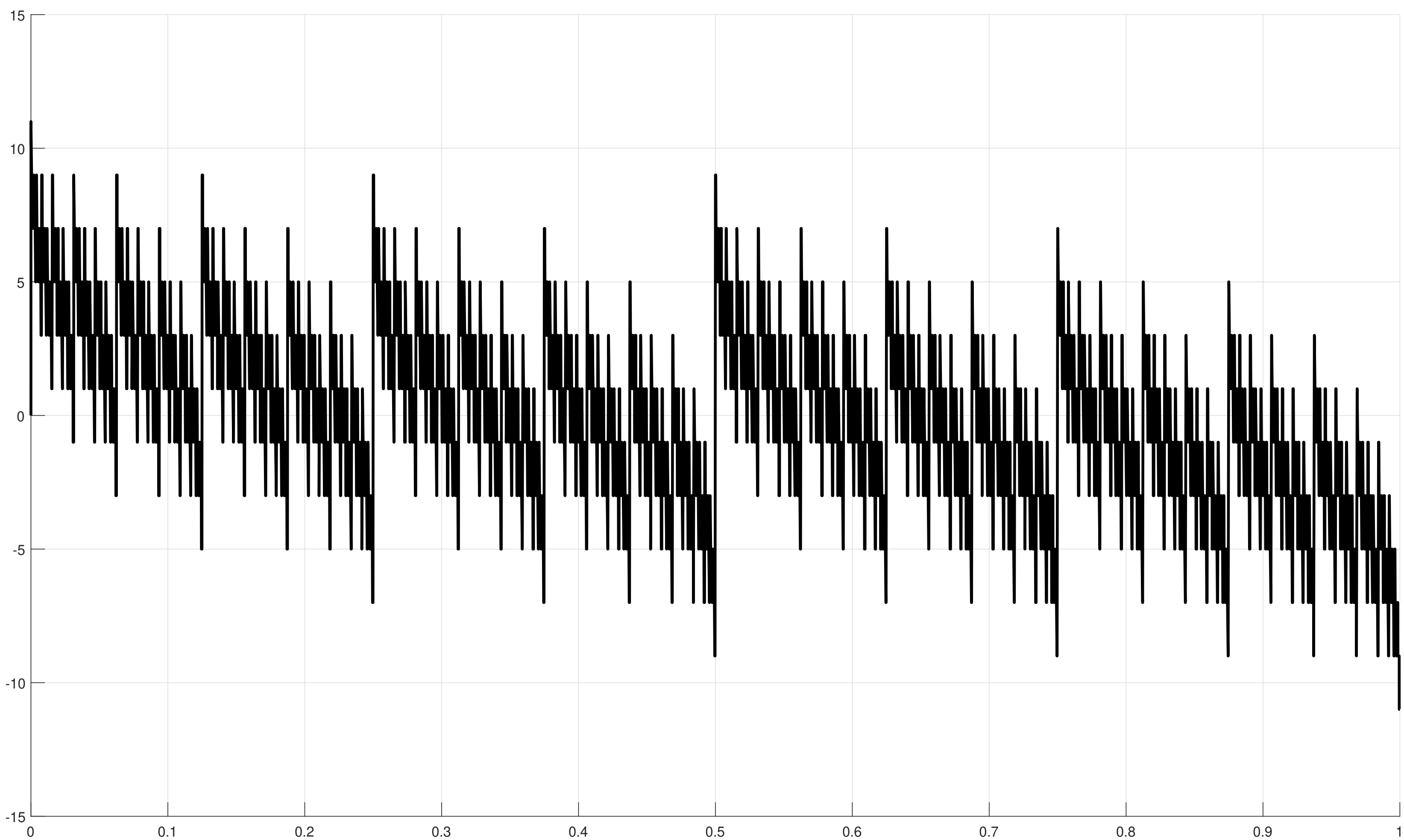
Proof. Set
$$\lambda_0 = 0, \qquad \lambda_n = \min\{t > \lambda_{n-1} : |B(t) - B(\lambda_{n-1})| = 1\}.$$
The waiting times $\{\lambda_n - \lambda_{n-1}\}$ are i.i.d. random variables with mean 1; see Wald's lemma.

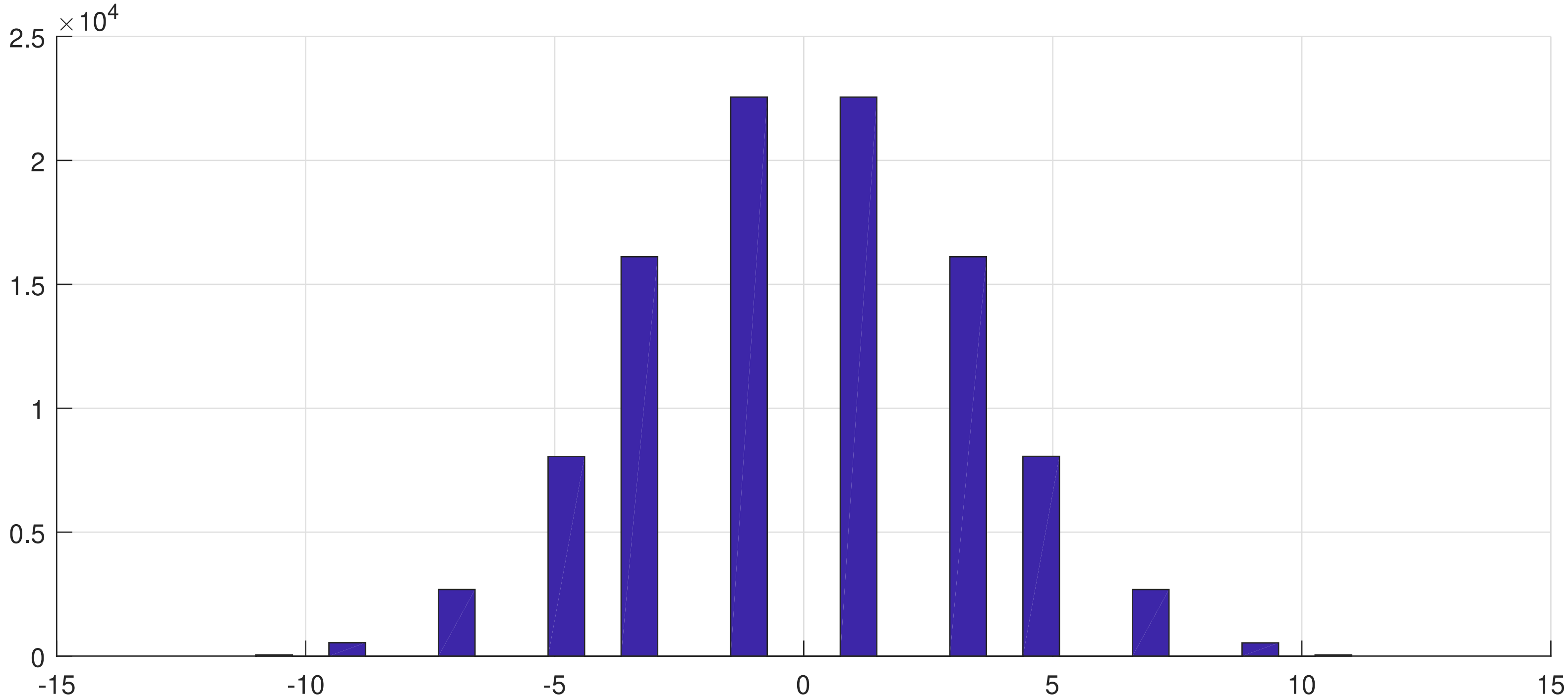
By the Markov property, $\mathbb{P}[\lambda_1 > k] \le \mathbb{P}[|B(1)| < 2]^k$ for every positive integer k, so λ_1 has finite variance. By the Law of Large Numbers λ_n/n will converge to 1, and the corollary follows. $\square$

Theorem 6.4 (Strong Law of Large Numbers)**.** *Let* $(X, d\nu)$ *be a probability space and* $\{f_n\}$, $n = 1, 2\ldots$ *a sequence of orthogonal functions in* $L^2(X, d\nu)$*. Suppose* $E(f_n^2) = \int |f_n|^2 \, d\nu \leq 1$*, for all* n*. Then*

$$\frac{1}{n} S_n = \frac{1}{n} \sum_{k=1}^{n} f_k \to 0,$$

a.e. (with respect to ν*) as* $n \to \infty$*.*


15
10
5
0
-5
-10
-15
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1



$\times 10^4$
2.5
2
1.5
1
0.5
0
-15
-10
-5
0
5
10
15

By 1915 Hausdorff had proved that if $\{f_n\}$ are independent and satisfy $\int f_n d\nu = 0$ and $\int f_n^2\, d\nu = 1$, then

$$\lim_{N\to\infty} \frac{1}{N^{\frac{1}{2}+\epsilon}} \sum_{n=0}^{N} f_n(x) = 0 \text{ for a.e. } x$$

and for every $\epsilon > 0$.

After that Hardy–Littlewood, and independently Khinchin, proved

$$\lim_{N\to\infty} \frac{1}{\sqrt{N \log N}} \sum_{n=0}^{N} f_n(x) = 0 \text{ for a.e. } x.$$

The "final" result, found by Khinchin for a special case in 1928 and proved in general by Hartman–Wintner says

$$\limsup_{N\to\infty} \frac{1}{\sqrt{2N \log\log N}} \sum_{n=0}^{N} f_n(x) = 1 \text{ for a.e. } x.$$

Proof of SLLN. We begin with the simple observation that if $\{g_n\}$ is a sequence of functions on a probability space $(X, d\nu)$ such that

$$\sum_n \int |g_n|^2 \, d\nu < \infty,$$

then $\sum_n |g_n|^2 < \infty$ ν-a.e. and hence $g_n \to 0$ ν-a.e.

Using this, it is easy to verify the Strong Law of Large Numbers (LLN) for $n \to \infty$ along the sequence of squares.

The functions $\{f_n\}$ are orthogonal, so

$$\int \left(\frac{1}{n} S_n\right)^2 d\nu = \frac{1}{n^2} \int |S_n|^2 d\nu = \frac{1}{n^2} \sum_{k=1}^{n} \int |f_k|^2 \, d\nu \leq \frac{1}{n}.$$

Thus if we set $g_n = \frac{1}{n^2} S_{n^2}$, we have

$$\int g_n^2 \, d\nu \leq \frac{1}{n^2}.$$

This is summable, so the previous observation implies $g_n = n^{-2} S_{n^2} \to 0 \;\; \nu$-a.e.

For the limit over all positive integers, suppose that $m^2 \le n < (m+1)^2$. Then

$$
\begin{aligned}
\int |\frac{1}{m^2} S_n - \frac{1}{m^2} S_{m^2}|^2 \, d\nu &= \frac{1}{m^4} \int | \sum_{k=m^2+1}^{n} f_k|^2 \, d\nu \\
&= \frac{1}{m^4} \int \sum_{k=m^2+1}^{n} |f_k|^2 \, d\nu \\
&\le \frac{2}{m^3},
\end{aligned}
$$

since the sum has at most $2m$ terms, each of size at most 1.

Set $m(n) = \lfloor \sqrt{n} \rfloor$ and
$$h_n = \frac{S_n}{m(n)^2} - \frac{S_{m(n)^2}}{m(n)^2}.$$
Now each integer m equals $m(n)$ for at most $2m+1$ different choices of n. Therefore,

$$\sum_{n=1}^{\infty} \int |h_n|^2 \, d\mu \leq \sum_{n=1}^{\infty} \frac{2}{m(n)^3} \leq \sum_m (2m+1) \frac{2}{m^3} < \infty,$$

so by the initial observation, $h_n \to 0$ a.e. with respect to ν. This yields that
$$\frac{1}{m(n)^2} S_n \to 0 \text{ a.e.}$$
But $m(n)^2 \leq n$, so $\frac{1}{n} S_n \to 0$ a.e., as claimed. $\square$

7. Solving the Dirichlet problem with Brownian motion

Definition 1. *Let $D \subset \mathbb{R}^d$ be a domain. We say that D satisfies the* **Poincaré cone condition** *if for each point $x \in \partial D$ there exists a cone $C_x(\alpha, h)$ of height $h(x)$ and angle $\alpha(x)$ such that $C_x(\alpha, h) \subset D^c$ and $C_x(\alpha, h)$ is based at x.*

Proposition 7.1 (Dirichlet Problem)**.** *Suppose $D \subset \mathbb{R}^d$ is a bounded domain with boundary ∂D, such that D satisfies the Poincaré cone condition, and f is a continuous function on ∂D. Then there exists a unique function u that is harmonic on D, continuous on $\overline{D}$ and satisfies $u(x) = f(x)$ for all $x \in \partial D$.*

Proof. The uniqueness claim follows from the maximum principle for harmonic functions.

To prove existence, let W be a Brownian motion in $\mathbb{R}^d$ and define

$$u(x) = \mathbb{E}_x f(W_{\tau_{\partial D}}), \quad \text{where} \quad \tau_A = \inf\{t \geq 0 : W_t \in A\}$$

for any Borel set $A \subset \mathbb{R}^d$.

For a ball $B(x, r) \subset D$, the strong Markov property implies that

$$u(x) = \mathbb{E}_x[\mathbb{E}_x[f(W_{\tau_{\partial D}})|\mathcal{F}_{\tau_{S(x,r)}}]] = \mathbb{E}_x[u(W_{\tau_{S(x,r)}})] = \int_{S(x,r)} u(y) d\, \mu_r,$$

where μ_r is the uniform distribution on the sphere $S(x, r)$. Therefore, u has the mean value property and so it is harmonic on D.

It remains to be shown that the Poincaré cone condition implies

$$\lim_{x \to z, x \in D} u(x) = f(z) \text{ for all } z \in \partial D.$$

Fix $z \in \partial D$, then there is a cone with height $h > 0$ and angle $\alpha > 0$ in D^c based at z. Let

$$\phi = \sup_{x \in B(0,\frac{1}{2})} \mathbb{P}_x[\tau_{S(0,1)} < \tau_{C_0(\alpha,1)}].$$

Then $\phi < 1$.

Note that if $x \in B(0, 2^{-k})$ then by the strong Markov property:

$$\mathbb{P}_x[\tau_{S(0,1)} < \tau_{C_0(\alpha,1)}] \leq \prod_{i=0}^{k-1} \sup_{x \in B(0,2^{-k+i})} \mathbb{P}_x[\tau_{S(0,2^{-k+i+1})} < \tau_{C_0(\alpha,2^{-k+i+1})}] = \phi^k.$$

Therefore, for any positive integer k, we have

$$\mathbb{P}_x[\tau_{S(z,h)} < \tau_{C_z(\alpha,h)}] \leq \phi^k$$

for all x with $|x - z| < 2^{-k}h$.

Given $\epsilon > 0$, there is a $0 < \delta \leq h$ such that $|f(y) - f(z)| < \epsilon$ for all $y \in \partial D$ with $|y - z| < \delta$. For all $x \in \bar{D}$ with $|z - x| < 2^{-k}\delta$,

$$|u(x) - u(z)| = |\mathbb{E}_x f(W_{\tau_{\partial D}}) - f(z)| \leq \mathbb{E}_x |f(W_{\tau_{\partial D}}) - f(z)|. \tag{7.1}$$

If the Brownian motion hits the cone $C_z(\alpha, \delta)$, which is outside the domain D, before it hits the sphere $S(z, \delta)$, then $|z - W\tau_{\partial D}| < \delta$, and $f(W_{\tau_{\partial D}})$ is close to $f(z)$. The complement has small probability. More precisely, (7.1) is bounded above by

$$2\|f\|_\infty \mathbb{P}_x\{\tau_{S(z,\delta)} < \tau_{C_z(\alpha,\delta)}\} + \epsilon \mathbb{P}_x\{\tau_{\partial D} < \tau_{S(z,\delta)}\} \leq 2\|f\|_\infty \phi^k + \epsilon.$$

Hence u is continuous on $\bar{D}$. $\square$

8. 2-dimensional Brownian motion is recurrent

Given $x \in \mathbb{R}^2, 1 \leq |x| \leq R$, we know that

$$\mathbb{P}_x[\tau_{S(0,R)} < \tau_{S(0,1)}] = a + b \log |x|.$$

The left-hand side is clearly a function of $|x|$, and it is a harmonic function of x for $1 < |x| < R$ by averaging over a small sphere surrounding x.

Setting $|x| = 1$ implies $a = 0$, and $|x| = R$ implies $b = \frac{1}{\log R}$. It follows that

$$\mathbb{P}_x[\tau_{S(0,R)} < \tau_{S(0,1)}] = \frac{\log |x|}{\log R}.$$

By scaling, for $0 < r < R$ and $r \leq |x| \leq R$,

$$\mathbb{P}_x[\tau_{S(0,R)} < \tau_{S(0,r)}] = \frac{\log \frac{|x|}{r}}{\log \frac{R}{r}}. \tag{8.1}$$

Definition 2. *A set A is* **polar** *for a Markov process X if for all x we have*

$$\mathbb{P}_x[X_t \in A \textit{ for some } t > 0] = 0.$$

The image of (δ, ∞) under Brownian motion W is the random set

$$W(\delta, \infty) \stackrel{\text{def}}{=} \bigcup_{\delta < t < \infty} \{W_t\}.$$

Proposition 8.1. *Points are polar for a planar Brownian motion W, that is, for all $z \in \mathbb{R}^2$ we have $\mathbb{P}_0\{z \in W(0,\infty)\} = 0$.*

Proof. Take $z \neq 0$ and $0 < \epsilon < |z| < R$,

$$\mathbb{P}_0\{\tau_{S(z,R)} < \tau_{S(z,\epsilon)}\} = \frac{\log \frac{|z|}{\epsilon}}{\log \frac{R}{\epsilon}}.$$

Let $\epsilon \to 0+$,

$$\mathbb{P}_0\{\tau_{S(z,R)} < \tau_{\{z\}}\} = \lim_{\epsilon \to 0+} \mathbb{P}_0\{\tau_{S(z,R)} < \tau_{S(z,\epsilon)}\} = 1,$$

and then

$$\mathbb{P}_0\{\tau_{S(z,R)} < \tau_{\{z\}} \text{ for all integers } R > |z|\} = 1.$$

It follows that

$$\mathbb{P}_0\{z \in W(0,\infty)\} = \mathbb{P}_0\{\tau_{\{z\}} < \infty\} = 0.$$

Let $f(z) = \mathbb{P}_z(0 \in W(0,\infty))$. Given $\delta > 0$, by the Markov property

$$\mathbb{P}_0\{0 \in W(\delta,\infty)\} = \mathbb{E}_0[f(W_\delta)] = 0.$$

Finally, $f(0) = \mathbb{P}(\bigcup_{n=1}^{\infty}\{0 \in W(\frac{1}{n},\infty)\}) = 0$. Hence any fixed single point is a polar set for a planar Brownian motion. $\square$

Corollary 8.2. *Almost surely, a Brownian path has zero area.*

Proof. The expected area $\mathbb{E}_0[\mathcal{L}_2(W(0,\infty))]$ of planar Brownian motion is

$$\mathbb{E}_0[\int_{\mathbb{R}^2} I_{\{z\in W(0,\infty)\}}\,dz] = \int_{\mathbb{R}^2} \mathbb{P}_0\{z \in W(0,\infty)\}\,dz = 0,$$

where the first equality is by Fubini's Theorem, the second from the previous theorem. So almost surely, the image of a planar Brownian motion is a set with zero Lebesgue measure. $\square$

Proposition 8.3. *Planar Brownian motion W is neighborhood recurrent. In other words,*

$$\mathbb{P}_0\{W(0,\infty) \textit{ is dense in } \mathbb{R}^2\} = 1.$$

Proof. Note that $\limsup_{t\to\infty} |W_t| = \infty$, so for all $z \in \mathbb{R}^2$ and $\epsilon > 0$,

$$\mathbb{P}_0\{\tau_{B(z,\epsilon)} = \infty\} = \lim_{R\to\infty} \mathbb{P}_0\{\tau_{S(z,R)} < \tau_{B(z,\epsilon)}\} = 0.$$

Summing over all rational z and ϵ completes the proof. □

Evan's theorem says that $X \subset \mathbb{R}^2$ is a G_δ with zero capacity iff it supports a measure μ whose potential U_μ tends to infinity everywhere on X.

G.C. Evans. Potentials and positively infinite singularities of harmonic functions. Monatshefte für Math. u. Phys., 43: 419–424, 1936.

Using this, we can mimic the proof that a point is never hit by Brownian motion to show that a set of zero capacity is never hit: evaluating the expected value of the potential function along Brownian paths gives a finite value, so the probability of visit the set where the potential is infinite is zero.

A recent generalization of Even's theorem is given in
On Evans' and Choquet's theorems for polar sets by Hansen and Netuka.

Lemma 8.4. *Define*

$$a = \mathbb{P}_x(\textit{Brownian motion } W \textit{ hits } S(0,r) \textit{ before } S(0,R)),$$

where $r < |x| < R$. *Then*

$$a = \frac{(r/|x|)^{d-2} - (r/R)^{d-2}}{1 - (r/R)^{d-2}}. \tag{8.2}$$

Proof. The given function in harmonic in the annulus, equals 1 on the inner boundary and equals 0 on the outer boundary. □

In $\mathbb{R}^2$ a Brownian motion started on $\{|z| = 2^n\}$ is equally likely to firs hit $\{|z| = 2^{n-1}\}$ or $\{|z| = 2^{n+1}\}$.

In $\mathbb{R}^3$ a Brownian motion started on $\{|z| = 2^n\}$ hits $\{|z| = 2^{n-1}\}$ with probability $1/3$ and hits $\{|z| = 2^{n+1}\}$ with probability $2/3$.

The corresponding random walk on powers of two has positive mean, so by the strong law of large numbers, Brownian motion in $\mathbb{R}^3$ tends to ∞.

Proposition 8.5. *Brownian motion W in dimension $d \geq 3$ is transient, i.e.,* $\lim_{t\to\infty} |W(t)| = \infty$.

Kakutani is credited with summarizing this by saying "A drunk man will find his way home, but a drunk bird may get lost forever."

9. 2-dimensional Brownian motion is conformally invariant

In $\mathbb{R}^3$ the conformal image of a Brownian motion is not Brownian motion.

In $\mathbb{R}^3$ Brownian motion tends to ∞ almost surely.

Reflection through a sphere is conformal, but maps Brownian motion to a process that converges to the center of the sphere. This image can't be Brownian motion.

Alternatively, reflect $\{\frac{1}{2} < |z| < 2\}$ through the unit sphere. Harmonic measures w.r.t $z = (1, 0, 0)$ of the two boundaries are not preserved since they are not equal.

If u is harmonic on U and $f : V \to U$ is conformal, then $u \circ f$ is harmonic on V. In fact, conformal and anti-conformal maps are the only homeomorphisms with this property.

Since the hitting distribution of Brownian motion solves the Dirichlet problem on both domains, it is easy to verify that, assuming f extends continuously to ∂V, it maps the Brownian hitting distribution on ∂V (known as harmonic measure) to the harmonic measure on ∂U.

Does f take individual Brownian paths in V to Brownian paths in U? This is not quite correct: if $f(z) = 2z$, then $f(B(t))$ leaves a disk of radius 2 in the same expected time that $B(t)$ leaves a disk of radius 1, so $f(B(t))$ is "too fast" to be Brownian motion. However, it is Brownian motion up to a time change.

What does this mean? Suppose $f : V \to U$ is conformal and suppose $0 \in V$. For a Brownian path started at 0 let τ be the first hitting time on ∂V. For $0 \le t < \tau$, define

$$\varphi(t) = \int_0^t |f'(B(t))|^2 \, dt \,. \tag{9.1}$$

Why does the integral makes sense? For $0 \le t < \tau$, we know that $B([0,t])$ is a compact subset of V and that $|f'|$ is a continuous function that is bounded above and below on this compact set. Thus the integrand above is a bounded continuous function. Therefore $\varphi(t)$ is continuous and strictly increasing, and so $\varphi^{-1}(t)$ is well defined.

We will need the following well known result.

Lemma 9.1. *(Kolomogorov's maximal inequality) Let X_i be independent with mean zero and finite variance. Write $S_n = \sum_{k=1}^n X_i$. Then*

$$\mathbb{P}[\max_{1\le k\le n} |S_k| \ge h] \le \frac{\operatorname{Var} S_n}{h^2}.$$

Proof. Let A_k denote the event that k is minimal such that $|S_k| \ge h$. Then since independence implies orthogonality,

$$\mathbb{E}S_n^2 \mathbf{1}_{A_k} \ge \mathbb{E}S_k^2 \mathbf{1}_{A_k} \ge h^2 \mathbb{P}(A_k).$$

Now sum over k,

$$\operatorname{Var} S_n = \mathbb{E}S_n^2 = \sum_{k=1}^n \mathbb{E}S_n^2 \mathbf{1}_{A_k} \ge h^2 \sum_{k=1}^n \mathbb{P}(A_k) = h^2 \mathbb{P}[\max_{1\le k\le n} |S_k| \ge h].$$

$\square$

Theorem 9.2. *Suppose that $V, U \subset \mathbb{C}$ are open sets with $0 \in V$, the map $f : V \to U$ is conformal and φ is defined by (9.1). If $B(\cdot)$ is Brownian motion in V and τ is the exit time from V, then $\{X(t) : 0 \leq t \leq \phi(\tau)\}$ defined by $X(t) := f(B(\varphi^{-1}(t)))$ is a Brownian motion in U started at $f(0)$ and stopped at ∂U.*

Proof. Let $Y(t)$ be Brownian motion in U started at $f(0)$. The idea of the proof is to show that both $X(t)$ and $Y(t)$ are limits of discrete random walks that depend on a parameter ϵ, and that as $\epsilon \to 0$, the two random walks get closer and closer, and hence have the same limiting distribution.

Fix a small $\epsilon > 0$, and starting at $f(0)$, sample $Y(t)$ every time it moves distance ϵ from the previous sample point. We stop when a sample point lands within 2ϵ of ∂U. Since $Y(t)$ is almost surely continuous, it is almost surely the limit of the linear interpolation of these sampled values.

Because Brownian motion is rotationally invariant, the increment between samples is uniformly distributed on a circle of radius ϵ. Thus $Y(t)$ is the limit of the following discrete process: starting at $z_0 = f(0)$, choose z_1 uniformly on $|z - z_0| = \epsilon$. In general, z_{n+1} is chosen uniformly on $|z - z_n| = \epsilon$.

Now sample $X(t)$ starting at z_0, each time it first moves distance ϵ from the previous sample. We claim that, as above, z_{n+1} is uniformly distributed on an ϵ-circle around z_n. Note that if $D = D(z_n, \epsilon) \subset U$, then the probability that $X(t)$ first hits ∂D in a set $E \subset \partial D$ is the same as the probability that $B(t)$ started at $w_n = f^{-1}(z_n)$ first hits $\partial f^{-1}(D)$ in $F = f^{-1}(E)$.

This probability is the solution at w_n of the Dirichlet problem on $f^{-1}(D)$ with boundary data $\mathbf{1}_F$.

Since f is conformal, this value is the same as the solution at z_n of the Dirichlet problem on D with boundary data $\mathbf{1}_E$, which is just the normalized angle measure of E.

Thus the hitting distribution of $X(t)$ on ∂D starting from z_n is the uniform distribution, just as it is for usual Brownian motion, only the time needed for $f(B(t))$ to hit ∂D may be different.

How different? The time $T_n - T_{n-1}$ between the samples z_{n-1} and z_n for the Brownian motion Y are i.i.d. random variables with expectation $\epsilon^2/2$ and variance $O(\epsilon^4)$, (exercise using Wald's lemma).

So taking $t_n := n\epsilon^2/2$, the time T_n for $Y(t)$ to reach z_n satisfies $\mathbb{E}[|T_n - t_n|^2] = O(n\epsilon^4) = O(t_n\epsilon^2)$. Moreover, by Kolomogorov's maximal inequality we have

$$\mathbb{P}[\max_{1\le k\le n} |T_k - t_k| \ge h] \le O(n\epsilon^4/h^2)\,.$$

This remains true even if we condition on the sequence $\{z_k\}$.

Next we do the same calculation for the process $X(t)$. Note that we may pick the points $\{z_n\}$ to be the same for the two processes X and Y without altering their distributions – this amounts to a *coupling* of X and Y.

The exit time for $X(t)$ from $D = D(z_n, \epsilon)$ is same as the exit time for $B(\varphi^{-1}(t))$ from $f^{-1}(D)$ starting at the point $p = f^{-1}(z_n)$. Since f is conformal, f is close to linear on a neighborhood of p with estimates that only depend on the distance of p from ∂V.

Thus for any $\delta > 0$, we can choose ϵ so small (uniformly for all p in any compact subset of V) that

$$D(p, \frac{\epsilon}{(1+\delta)|f'(p)|}) \subset f^{-1}(D) \subset D(p, \frac{\epsilon(1+\delta)}{|f'(p)|}).$$

Therefore the expected exit time for the Brownian motion $B(\cdot)$ from $f^{-1}(D)$ starting at p is bounded above and below by the expected exit times for these two disks, which are

$$\frac{1}{2}\left(\frac{\epsilon}{(1+\delta)|f'(p)|}\right)^2 \qquad \text{and} \qquad \frac{1}{2}\left(\frac{\epsilon(1+\delta)}{|f'(p)|}\right)^2.$$

As long as $B(s)$ is inside $f^{-1}(D)$, and ϵ is small enough, we have

$$\frac{|f'(p)|^2}{(1+\delta)^2} \leq |f'(B(s))|^2 \leq |f'(p)|^2(1+\delta)^2.$$

Therefore $\varphi(s)$ has derivative between these two bounds during this time and so φ^{-1} has derivative bounded between the reciprocals, i.e.,

$$\frac{1}{(1+\delta)^2|f'(p)|^2} \leq \frac{d}{ds}\varphi^{-1} \leq \frac{(1+\delta)^2}{|f'(p)|^2}.$$

Thus the expected exit time of $B \circ \varphi^{-1}$ from $f^{-1}(D)$ is between

$$\frac{\epsilon^2}{2(1+\delta)^4} \qquad \text{and} \qquad \frac{\epsilon^2(1+\delta)^4}{2}.$$

The bounds are uniform as long as ϵ is small enough and z_n is in a compact subset of U.

Let S_n denote the time it takes $X(\cdot)$ to reach z_n. The random variables $S_n - S_{n-1}$ are not i.i.d., but they are independent given the sequence $\{z_k\}$, and have variances $O(\epsilon^4)$, so

$$\mathbb{P}[\max_{1\leq k\leq n} |S_k - s_k| \geq h|\, \{z_k\}_{k=1}^n] = O(n\epsilon^4/h^2)\,,$$

where $s_n = \mathbb{E}(S_n|\, \{z_k\}_{k=1}^n)$.

We have already proved that

$$(1+\delta)^{-4} \le s_n/t_n \le (1+\delta)^4 \,,$$

so it follows that for $n \le 2C\epsilon^{-2}$

$$\mathbb{P}[\max_{1\le k\le n} |S_k - T_k| \ge 5C\delta + 2h \,|\, \{z_k\}_{k=1}^n] \to 0 \text{ as } \epsilon \to 0.$$

Using the uniform continuity of Brownian motion on $[0, C]$, given any $\eta > 0$ we can choose δ and h sufficiently small so that

$$\mathbb{P}[\max_{t\le C} |X(t) - Y(t)| \ge \eta \,|\, \{z_k\}_{k=1}^n] \to 0$$

as $\epsilon \to 0$. Since η can be taken arbitrarily small, this coupling implies that $X(\cdot)$ and $Y(\cdot)$ indeed have the same distribution until the first exit from U. $\square$

The following elegant result is due to Markowsky.

Lemma 9.3. *Suppose that $f(z) = \sum_{n=0}^{\infty} a_n z^n$ is conformal in $\mathbb{D}$. Then the expected time for Brownian motion to leave $\Omega = f(\mathbb{D})$ starting at $f(0)$ is $\frac{1}{2}\sum_{n=1}^{\infty} |a_n|^2$.*

We give two proofs, the first relies on Wald's Lemma and the second on an exercise involving Green's theorem.

Proof 1. We may assume that $f(0) = 0$ since translating the domain and starting point does not change either the expected exit time or the infinite sum (it doesn't include a_0). We use the identity

$$2\mathbb{E}[\tau] = \mathbb{E}[|B_\tau|^2],$$

where τ is a stopping time for a 2-dimensional Brownian motion B (Lemma 4.9). W apply it in the case when B starts at $p = f(0)$ and is stopped when it hits $\partial\Omega$.

Then the expectation on the right side above is

$$\int_{\partial\Omega} |z|^2 \, d\omega_p(z),$$

where ω_p is harmonic measure on $\partial\Omega$ with respect to p, i.e., the hitting distribution of Brownian motion started at p. By the conformal invariance of Brownian motion, we get

$$\mathbb{E}[|B_\tau|^2] = \frac{1}{2\pi} \int_{\partial\mathbb{D}} |f(z)|^2 \, d\theta = \sum_{n=1}^{\infty} |a_n|^2.$$

$\square$

Proof 2. By definition, the expected exit time from Ω for Brownian motion started at $w = f(0)$ is (writing $z = x + iy$)

$$\iint_{\Omega} G_{\Omega}(z, w)dxdy.$$

By the conformal invariance of Green's functions, this is the same as

$$\iint_{\mathbb{D}} G_{\mathbb{D}}(z, 0)|f'(z)|^2 dxdy.$$

The Green's function for the disk is $G_{\mathbb{D}}(z, 0) = \frac{1}{\pi} \log |z|^{-1}$ (we leave this as an exercise), so this formula becomes

$$\frac{1}{\pi} \iint_{\mathbb{D}} |f'(z)|^2 \log \frac{1}{|z|} dxdy.$$

This can be evaluated using the identities (writing $z = e^{i\theta}$),

$$\int_0^{2\pi} |\sum c_n z^n|^2 \, d\theta = \int_0^{2\pi} \left(\sum c_n z^n\right)\left(\overline{\sum c_n z^n}\right) d\theta = 2\pi \sum |c_n|^2,$$

and

$$\int_0^x t^m \log t \, dt = x^{m+1} \left(\frac{\log x}{m+1} - \frac{1}{(m+1)^2} \right), \qquad m \neq -1,$$

as follows:

$$\begin{aligned}
\frac{1}{\pi} \iint_{\mathbb{D}} \log \frac{1}{|z|} |f'(z)|^2 dx dy &= \frac{1}{\pi} \int_0^{2\pi} \int_0^1 \log \frac{1}{r} |f'(re^{i\theta})|^2 r \, dr d\theta \\
&= 2 \sum_{n=1}^{\infty} n^2 |a_n|^2 \int_0^1 r^{2n-1} \log \frac{1}{r} \, dr \\
&= 2 \sum_{n=1}^{\infty} n^2 |a_n|^2 \left[-r^{2n} \left(\frac{\log r}{2n} - \frac{1}{(2n)^2} \right) \right]_0^1 \\
&= 2 \sum_{n=1}^{\infty} n^2 |a_n|^2 \frac{1}{4n^2} \\
&= \frac{1}{2} \sum_{n=1}^{\infty} |a_n|^2. \quad \square
\end{aligned} \tag{9.2}$$

Corollary 9.4. *Among all simply connected domains with area π and containing 0, Brownian motion started at 0 has the largest expected exit time for the unit disk.*

Proof. If $f : \mathbb{D} \to \Omega$ is conformal, then

$$\begin{aligned}
\pi = \text{area}(\Omega) &= \iint_{\mathbb{D}} |f'(z)|^2 dxdy \\
&= \int_0^{2\pi} \int_0^1 |\sum_{n=1}^{\infty} n a_n r^{n-1} e^{i(n-1)\theta}|^2 r \, dr d\theta \\
&= 2\pi \int_0^1 \sum_{n=1}^{\infty} n^2 |a_n|^2 r^{2n-1} \, dr \\
&= \pi \sum_{n=1}^{\infty} n|a_n|^2 \geq \pi \sum_{n=1}^{\infty} |a_n|^2 .
\end{aligned}$$

By Lemma 9.3 the expected exit time is $\leq \frac{1}{2}$ with equality if and only if $|a_1| = 1$, $a_n = 0$ for $n \geq 2$, so the disk is optimal. ∎

The expected time for a 1-dimensional Brownian motion started at zero to leave $[-1, 1]$ is 1 (this was calculated earlier using Wald's lemma and is the same as the time for a 2-dimensional path to leave the infinite strip $S = \{x + iy : |y| < 1\}$.

This strip is the image of the unit disk under the conformal map

$$f(z) = \frac{2}{\pi} \log \frac{1+z}{1-z}$$

since the linear fractional map $(1 + z)/(1 - z)$ maps the disk to the right half-plane and the logarithm carries the half-plane to the strip $\{|y| \leq \pi/2\}$.

Since

$$f(z) = \frac{2}{\pi}[\log(1+z) - \log(1-z)] = \frac{4}{\pi}\left(z + \frac{1}{3}z^3 + \frac{1}{5}z^5 + \cdots\right),$$

the expected time a Brownian motion spends in S is

$$1 = \frac{1}{2}\left(\frac{4}{\pi}\right)^2\left(1 + \frac{1}{9} + \frac{1}{25} + \cdots\right),$$

so

$$\frac{\pi^2}{8} = \left(1 + \frac{1}{9} + \frac{1}{25} + \cdots\right).$$

From this we can deduce

$$
\begin{aligned}
\zeta(2) &= 1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \cdots \\
&= (1 + \frac{1}{9} + \frac{1}{25} + \cdots) + (\frac{1}{4} + \frac{1}{16} + \frac{1}{36} + \cdots) \\
&= \frac{\pi^2}{8} + \frac{1}{4}\zeta(2)
\end{aligned}
$$

which implies $\zeta(2) = \pi^2/6$.