Path Properties of Brownian Motion

Lecture notes for Statistics 206b
UC Berkeley, Spring 1998

Yuval Peres

Notes edited by Bálint Virág
Version of October 8, 1998

Contents

CHAPTER 1

Brownian Motion

The following are from week 1, (W 1/21, F 1/23, 1998).
Notes taken by Frédéric Latrémolière, `fred@stat.berkeley.edu`.

1. Motivation – Intersection of Brownian paths

Consider a number of Brownian motion paths started at different points. Say that they intersect if there is a point which lies on all of the paths. Do the paths intersect? The answer to this question depends on the dimension:

- In $\mathbf{R}^2$, any finite number of paths intersect with positive probability (this is a theorem of Dvoretsky, Erdős, Kakutani in the 1950's),
- In $\mathbf{R}^3$, 2 but not 3 paths intersect with positive probability (this is a theorem of Dvoretsky, Erdős, Kakutani and Taylor),
- In $\mathbf{R}^d$ for $d \geq 4$, no pair of paths intersect with positive probability.

The principle we will use to establish these results is intersection equivalence between Brownian motion and certain random Cantor-type sets. Here we will introduce the concept for $\mathbb{R}^3$ only. Partition the cube $[0,1]^3$ in eight congruent sub-cubes, and keep each of the sub-cubes with probability $\frac{1}{2}$. For each cube that remained at the end of this stage, partition it into eight sub-cubes, keeping each of them with probability $\frac{1}{2}$, and so on. Let $Q\left(3, \frac{1}{2}\right)$ denote the limiting set— that is, the intersection of the cubes remaining at all steps. This set is not empty with positive probability, since, if we consider that the remaining sub-cubes of a given cube as its "children" in a branching process, then the expected number of offsprings is four, so this process has positive probability not to die out.

One can prove that, there exists two positive constants C_1, C_2 such that, if Λ is a closed subset of $[0,1]^3$, and $\{B_t\}$ is a Brownian motion started at a point uniformly chosen in $[0,1]^3$, then:

$$C_1 \mathbf{P}\left(Q\left(3, \frac{1}{2}\right) \cap \Lambda \neq \emptyset\right) \leq \mathbf{P}\left(\exists t \geq 0 \;\; B_t \in \Lambda\right) \leq C_2 \mathbf{P}\left(Q\left(3, \frac{1}{2}\right) \cap \Lambda \neq \emptyset\right)$$

The motivation is that, though the intersection of two independent Brownian paths is a complicated object, the intersection of two sets of the form $Q\left(3, \frac{1}{2}\right)$ is a set of the same kind—namely, $Q\left(3, \frac{1}{4}\right)$. The previously described branching process dies out as soon as we intersect more than two of these Cantor-type sets—hence the result about intersection of paths in $\mathbb{R}^3$.

2. Gaussian random variables

Brownian motion is at the meeting point of the most important categories of stochastic processes: it is a martingale, a strong Markov process, a process with independent and stationary increments, and a Gaussian process. We will construct Brownian motion as a specific Gaussian process. We start with the definitions of Gaussian random variables:

DEFINITION 2.1. A real-valued random variable X on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ has a **standard Gaussian** (or **standard normal**) distribution if

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

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 $Y : \Omega \to \mathbf{R}^p$ 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:

$$Y = AX + b \tag{2.1}$$

We are now ready to define the Gaussian processes.

DEFINITION 2.2. A stochastic process $(X_t)_{t \in I}$ is said to be a **Gaussian process** if for all k and $t_1, \dots, t_k \in I$ the vector $(X_{t_1}, \dots, X_{t_k})^t$ is Gaussian.

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 in (2.1) has

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

Recall that an $n \times n$ matrix A is said to be **orthogonal** if $AA^t = I_n$. The following lemma shows that the distribution of a Gaussian vector is determined by its mean and covariance.

LEMMA 2.3.

(i) *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.*

(ii) *If Y and Z are Gaussian vector in $\mathbb{R}^n$ such that $\mathbb{E}Y = \mathbb{E}Z$ and $\mathrm{Cov}(Y) = \mathrm{Cov}(Z)$, then Y and Z have the same distribution.*

PROOF.

(i) As the coordinates of X are independent standard Gaussian, X has density given by:

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

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

(ii) It is sufficient to consider the case when $\mathbb{E}Y = \mathbb{E}Z = 0$. Then, using Definition 2.1), there exist standard Gaussian vectors X_1, X_2 and matrices A, B so that

$$Y = AX_1 \quad \text{and} \quad Z = BX_2.$$

By adding some columns of zeroes to A or B if necessary, we can assume that X_1, X_2 are both k-vectors for some k and A, B are both $n \times k$ matrices.

Let $\mathcal{A}$, $\mathcal{C}$ denote the vector spaces generated by the row vectors of A and C, respectively. To simplify notations, assume without loss of generality that the first ℓ row vectors of A form a basis for the space $\mathcal{A}$. For any matrix M let M_i denote the ith row vector of M, and define the linear map Θ from $\mathcal{A}$ to $\mathcal{C}$ by

$$\Theta A_i = C_i \quad \text{for } i = 1 \dots \ell.$$

We want to check that Θ is an isomorphism. Assume that there is a vector $v_1 A_1 + \ldots + v_\ell A_\ell$ whose image is 0. Then the k-vector $v = (v_1, v_2, \dots, v_\ell, 0, \dots, 0)$ satisfies $v^t C = 0$, and so $\|v^t A\|^2 = v^t A A^t v = v^t C C^t v = 0$, giving $v^t A = 0$. This shows that Θ is one-to-one, in particular $\dim \mathcal{A} \le \dim \mathcal{C}$. By symmetry $\mathcal{A}$ and $\mathcal{C}$ must have the same dimension, so Θ is an isomorphism.

As the coefficient (i, j) of the matrix AA^t is the scalar product of A_i and A_j, the identity $AA^t = CC^t$ implies that Θ is an orthogonal transformation from $\mathcal{A}$ to $\mathcal{C}$. We can extend it to map the orthocomplement of $\mathcal{A}$ to the orthocomplement of $\mathcal{C}$ orthogonally, getting an orthogonal map $\Theta : \mathbb{R}^k \to \mathbb{R}^k$. Then

$$Y = AX_1, \qquad Z = CX_2 = A\Theta X_2,$$

and (ii) follows from (i). ∎

Thus, the two first moments of a Gaussian vector are sufficient to characterize its distribution, hence the introduction of the notation $\mathcal{N}(\mu, \Sigma)$ to designate the normal distribution with expectation μ and covariance matrix Σ. A useful corollary of this lemma is:

COROLLARY 2.4. *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. ∎

As a conclusion of this section, we state the following tail estimate of the standard Gaussian distribution:

LEMMA 2.5. *Let Z be distributed as $\mathcal{N}(0, 1)$. Then for all $x \ge 0$:*

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

PROOF. The right inequality is obtained by the estimate:

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

since, in the integral, $u \geq x$. The left inequality is proved as follows: Let us define

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

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

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

so the right inequality implies $f'(x) \geq 0$ for all $x \geq 0$. This implies $f(x) \leq 0$, proving the lemma. ■

3. Lévy's construction of Brownian motion

3.1. Definition. Standard Brownian motion on an interval I containing 0 is defined by the following properties:

DEFINITION 3.1. A 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) $\forall k$ natural $\forall t_1 < \ldots < t_k \quad B_{t_k} - B_{t_{k-1}}, \ldots, B_{t_2} - B_{t_1}$ are independent,
(iii) $\forall t, h \in [0,1] \quad t + h \in [0,\infty) \Rightarrow B_{t+h} - B_t$ has $\mathcal{N}(0,h)$ distribution.
(iv) Almost surely, $t \mapsto B_t$ is continuous on $[0,1]$.

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

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

Indeed, assume that $t \geq s$. Then $\mathrm{Cov}(B_t, B_s) = \mathrm{Cov}(B_t - B_s, B_s) + \mathrm{Cov}(B_s, B_s)$ by bilinearity of the covariance. The first term vanishes by the independence of increments, and the second term equals s by properties (iii) and (i).

Kolmogorov's extension theorem implies the existence of any countable time set stochastic process $\{X_t\}$ if we know its finite-dimensional distributions and they are consistent. Thus, standard Brownian motion could be easily constructed on any countable time set. However knowing finite dimensional distributions is not sufficient to get **continuous paths**, as the following example shows.

EXAMPLE 3.2. Suppose that standard Brownian motion $\{B_t\}$ on $[0,1]$ has been constructed, and consider an independent random variable U uniformly distributed on $[0,1]$. Define:

$$\tilde{B}_t = \begin{cases} B_t & \text{if } t \neq U \\ 0 & \text{otherwise} \end{cases}$$

The finite-dimensional distributions of $\{\tilde{B}_t\}$ are the same as the ones of $\{B_t\}$. However, the process $\{\tilde{B}_t\}$ has almost surely discontinuous paths.

In measure theory, one often identifies functions with their equivalence class for almost-everywhere equality. As the above example shows, it is important not to make this identification in the study of continuous-time stochastic processes. Here we want to define a probability measure on the set of the continuous functions.

3.2. Construction. The following construction, due to the French probabilist Paul Lévy, consist 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 an induction, 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.

We will use the following basic lemma. The proof can be found, for instance, in Durrett (1995).

LEMMA 3.3 (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 $\sum_{i=0}^{\infty} \mathbf{P}(A_i) < \infty$, then $\mathbf{P}(A_i \ i.o.) = 0$.*

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

THEOREM 3.4 (Wiener 1923). *Standard Brownian motion on $[0,\infty)$ exists.*

A proof similar to the one below can be found in Bass (1995).

PROOF. (Lévy 1948)

For $n \geq -1$, 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 $B_1 = Z_{0,1}$. 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 $\mathcal{N}(0, r-s)$ 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 $n = -1$. Suppose that they hold for $n-1$. Define, for all $d \in D_n \setminus D_{n-1}$ a random variable B_d by

$$B_d = \frac{B_{d^-} + B_{d^+}}{2} + \frac{Z_d}{2^{(n-1)/2}} \tag{3.1}$$

where $d^+ = d + 2^{-n}$, and $d^{-1} = d - 2^{-n}$, and both are in D_{n-1}. Since $\frac{1}{2}[B_{d^+} - B_{d^-}]$ 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 2.4. 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, for $n \geq 0$, let us introduce the function:

$$F_n(x) = \begin{cases} 2^{-(n+1)/2} Z_d & \text{for } x \in D_n \setminus D_{n-1}, \\ 0 & \text{for } x \in D_{n-1}, \\ \text{linear} & \text{between consecutive points in } D_n. \end{cases} \tag{3.2}$$

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

$$B_d = \sum_{i=0}^{n} F_i(d) = \sum_{i=0}^{\infty} F_i(d). \tag{3.3}$$

This can be seen by induction. Suppose that it holds for $n-1$. Let $d \in D_n - D_{n-1}$. Since for $0 \le i \le n-1$ F_i is linear on $[d^-, d^+]$, we get

$$\sum_{i=0}^{n-1} F_i(d) = \sum_{i=1}^{n-1} \frac{F_i(d^-) + F_i(d^+)}{2} = \frac{B_{d^-} + B_{d^+}}{2}. \tag{3.4}$$

Since $F_n(d) = 1/2^{(n-1)/2} Z_d$, comparing (3.1) and (3.4) gives (3.3).

On the other hand, we have, by definition of Z_d and by Lemma 2.5:

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

so the series $\sum_{n=0}^{\infty} \sum_{d \in D_n} \mathbb{P}(|Z_d| \ge c\sqrt{n})$ converges as soon as $c > \sqrt{2\log 2}$. Fix such a c. By the Borel-Cantelli Lemma 3.3 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:

$$\|F_n\|_\infty < c\sqrt{n} 2^{-n/2}. \tag{3.5}$$

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 limit 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 and 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 $[0,1]$. To conclude, if $\{B_t^n\}_t$ for $n \ge 0$ are independent Brownian motions on $[0,1]$, then

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

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

The following are from week 2, (W 1/28, F 1/30, 1998).
Notes taken by Ben Morris, `morris@stat.berkeley.edu`.

4. Basic properties of Brownian motion

Let $\{B(t)\}_{t\ge 0}$ be a standard Brownian motion, and let $a \ne 0$. The following **scaling relation** is a simple consequence of the definitions.

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

Also, define the **time inversion** of $\{B_t\}$ as

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

We claim that W is a standard Brownian motion. Indeed,

$$\mathrm{Cov}(W(t), W(s)) = ts\,\mathrm{Cov}(B(\frac{1}{t}, \frac{1}{s})) = ts\ (\frac{1}{t} \wedge \frac{1}{s}) = t \wedge s,$$

so W and B have the same finite dimensional distributions, and they have the same distributions as processes on the rational numbers. Since the paths of $W(t)$ are continuous except maybe at 0, we have

$$\lim_{t \downarrow 0} W(t) = \lim_{t \downarrow 0, t \in Q} W(t) = 0 \quad \text{a.s.}$$

so the paths of $W(t)$ are continuous on $[0, \infty)$ a.s. As a corollary, we get

COROLLARY 4.1. *[Law of Large Numbers for Brownian motion]*

$$\lim_{t \to \infty} \frac{B(t)}{t} = 0 \ \text{ a.s.}$$

PROOF. $\lim_{t \to \infty} \frac{B(t)}{t} = \lim_{t \to \infty} W(\frac{1}{t}) = 0$ a.s. ■

EXERCISE 4.2. Prove this result directly. Use the usual Law of Large Numbers to show that $\lim_{n \to \infty} \frac{B(n)}{n} = 0$. Then show that $B(t)$ does not oscillate too much between n and $n+1$.

REMARK. The symmetry inherent in the time inversion property becomes more apparent if one considers the Ornstein-Uhlenbeck diffusion, which is given by

$$X(t) = e^{-t} B(e^{2t})$$

This is a stationary Markov chain with $X(t) \sim N(0,1)$ for all t. It is a diffusion with a drift toward the origin proportional to the distance from the origin. Unlike Brownian motion, the Ornstein-Uhlenbeck diffusion is time reversible. The time inversion formula gives $\{X(t)\}_{t \geq 0} \stackrel{\mathrm{d}}{=} \{X(-t)\}_{t \geq 0}$. For t near $-\infty$, $X(t)$ relates to the Brownian motion near 0, and for t near ∞, $X(t)$ relates to the Brownian motion near ∞.

One of the advantages of Lévy's construction of Brownian motion is that it easily yields a modulus of continuity result. Following Lévy, 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 (3.2). Its derivative exists almost everywhere, and by definition and (3.5)

$$\|F_n'\|_\infty \leq \frac{\|F_n\|_\infty}{2^{-n}} \leq C_1(\omega) + \sqrt{n}\, 2^{n/2} \tag{4.1}$$

The random constant $C_1(\omega)$ is introduced to deal with the finitely many exceptions to (3.5). Now for $t, t+h \in [0,1]$, we have

$$|B(t+h) - B(t)| \leq \sum_n |F_n(t+h) - F_n(t)| \leq \sum_{n \leq \ell} h\|F_n'\|_\infty + \sum_{n > \ell} 2\|F_n\|_\infty \tag{4.2}$$

By (3.5) and (4.1) if $\ell > N$ for a random N, then the above is bounded by

$$h(C_1(\omega) + \sum_{n\le\ell} c\sqrt{n}\, 2^{n/2}) + 2\sum_{n>\ell} c\sqrt{n}\, 2^{-n/2}$$
$$\le C_2(\omega) h\sqrt{\ell}\, 2^{\ell/2} + C_3(\omega)\sqrt{\ell}\, 2^{-\ell/2}$$

The inequality holds because each series is bounded by a constant times its dominant term. Choosing $\ell = \log_2 h$, and choosing $C(\omega)$ to take care of the cases when $\ell \le N$ we get

$$|B(t+h) - B(t)| \le C(\omega)\sqrt{h \log\frac{1}{h}}. \tag{4.3}$$

The result is a form of Lévy's modulus of continuity. This is not enough to make $\{B_t\}$ a differentiable function since $\sqrt{h} \gg h$ for small h. But we still have

COROLLARY 4.3. *For every $\alpha < \frac{1}{2}$, Brownian paths are α - Hölder a.s.*

EXERCISE 4.4. Show that a Brownian motion is a.s. not $\frac{1}{2}$ - Hölder.

REMARK. There does exist a $t = t(\omega)$ such that $|B(t+h) - B(t)| \le C(\omega)h^{\frac{1}{2}}$ for every h, a.s. However, such t have measure 0. This is the slowest movement that is locally possible.

SOLUTION. Let $A_{k,n}$ be the event that $B((k+1)2^{-n}) - B(k2^{-n}) > c\sqrt{n}2^{-n/2}$. Then

$$\mathbb{P}(A_{k,n}) = \mathbb{P}(B(1) > c\sqrt{n}) \ge \frac{c\sqrt{n}}{c^2 n + 1} e^{-c^2 n/2}.$$

If $0 < c < \sqrt{2\log(2)}$ then $c^2/2 < \log(2)$ and $2^n \mathbb{P}(A_{k,n}) \to \infty$. Therefore,

$$\mathbb{P}(\bigcap_{k=1}^{2^n} A_{k,n}^c) = [1 - \mathbb{P}(A_{k,n})]^{2^n} \le e^{-2^n \mathbb{P}(A_{k,n})} \to 0.$$

The last inequality comes from the fact that $1 - x \le e^{-x}$ for all x. By considering $h = 2^{-n}$, one can see that

$$P\left(\forall h \;\; B(t+h) - B(t) \le c\sqrt{h \log_2 h^{-1}}\right) = 0 \qquad \text{if } c < \sqrt{2\log 2}.$$

With more careful work, it can be shown that this also holds for all $c < \sqrt{2}$ (and the $\sqrt{2}$ is sharp). Hence the paths are a.s. not $\frac{1}{2}$-Hölder. This shows that the $\log_2(\frac{1}{h})$ factor in (4.3) cannot be ignored. We will see later that the Hausdorff dimension of the graph of a Brownian motion is $\frac{3}{2}$ a.s.

Having proven that Brownian paths are somewhat "regular", let us see why they are "bizarre". One reason is that the paths of Brownian motion have no intervals of monotonicity. 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.

We will now show that for any time t_0, Brownian motion is not differentiable at t_0. For this, we need a simple proposition.

PROPOSITION 4.5. *A.s.*

$$\limsup_{n\to\infty} \frac{B(n)}{\sqrt{n}} = +\infty, \qquad \liminf_{n\to\infty} \frac{B(n)}{\sqrt{n}} = -\infty. \tag{4.4}$$

REMARK. Comparing this with Corollary 4.1, it is natural to ask what sequence $B(n)$ should be divided by to get a lim sup which is greater than 0 but less than ∞. An answer is given by the Law of the iterated logarithm in a later section.

The proof of (4.4) relies on the following standard fact, whose proof can be found, for example, in Durrett (1995). 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, that is 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. Here *finite permutation* means that $\sigma_n = n$ for all sufficiently large n.

PROPOSITION 4.6 (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.*

PROOF OF PROPOSITION 4.5.

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

By the scaling property, the expression in the lim sup equals $\mathbb{P}(B(1) > c)$, which is positive. Letting $X_n = B(n) - B(n-1)$ the Hewitt-Savage 0-1 law gives that $B(n) > c\sqrt{n}$ infinitely often. Taking the intersection over all natural c gives the first part of (4.4), and the second is proved similarly. ■

The two claims of Proposition 4.5 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

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

Then

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

which is infinite by Proposition 4.5. Similarly, $D_*W(0) = -\infty$, showing that W is not differentiable at 0.

COROLLARY 4.7. *Fix $t_0 \geq 0$. Brownian motion W a.s. 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. $t \to W(t_0 + t) - W(t_0)$ is a standard Brownian motion. ■

Does this imply that each t_0 is an accumulation point from the right for the level set $\{s : W(s) = W(t_0)\}$ a.s.? Certainly not; consider, for example the last 0 of $\{B_t\}$ before time 1. However, Z_B a.s. has no isolated points, as we will see later. Also, the set of exceptional

t_0-s must have measure 0. This is true in general. Suppose A is a measurable event (set of paths) such that

$$\forall t_0 \; \mathbb{P}(t \to W(t_0+t) - W(t_0) \text{ satisfies } A) = 1.$$

Let Θ_t be the operator that shifts paths by t. Then $\mathbb{P}(\bigcap_{t_0 \in \mathbf{Q}} \Theta_{t_0}(A)) = 1$. In fact, the Lebesgue measure of points t_0 so that W does not satisfy $\Theta_{t_0}(A)$ is 0 a.s. To see this, apply Fubini to the double integral

$$\int \int_0^\infty \mathbf{1}(W \notin \Theta_{t_0}(A)) \; dt_0 \; d\mathbb{P}(W)$$

EXERCISE 4.8. Show that $\forall t_0 \; \mathbb{P}(t_0$ is a local maximum for B$) = 0$, but a.s. local maxima are a countable dense set in $(0,\infty)$.

Nowhere differentiability of Brownian motion therefore requires a more careful argument than non-differentiability at a fixed point.

THEOREM 4.9 (Paley, Wiener, Zygmund 1933). *A.s. Brownian motion is nowhere differentiable. Furthermore, almost surely for all t either $D^*B(t) = +\infty$ or $D_*B(t) = -\infty$.*

REMARK. 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.

PROOF. (Dvoretsky, Erdős, Kakutani 1961) Suppose that there is a $t_0 \; \in [0,1]$ such that $-\infty < D_*B(t_0) \le D^*B(t_0) < \infty$. Then for some M we would have

$$\sup_{h \in [0,1]} \frac{|B(t_0+h) - B(t_0)|}{h} \le M. \tag{4.5}$$

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

$$|B\left((k+j)/2^n\right) - B\left((k+j-1)/2^n\right)| \le M(2j+1)/2^n. \tag{4.6}$$

Let $\Omega_{n,k}$ be the event that (4.6) 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

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

whose the sum over all n is finite. By the Borel-Cantelli lemma:

$$\mathbb{P}\left((4.5) \text{ is satisfed}\right) \le \mathbb{P}\left(\bigcup_{k=1}^{2^n} \Omega_{n,k} \text{ for infinitely many } n\right) = 0.$$

That is, for sufficiently large n, there are no three good increments in a row so (4.5) is not satisfied. ∎

EXERCISE 4.10. Let $\alpha > 1/2$. Show that a.s. for all $t > 0, \exists\; h > 0$ such that $|B(t+h) - B(t)| > h^\alpha$.

SOLUTION. Suppose that there is a $t_0 \in [0,1]$ such that

$$\sup_{h\in[0,1]} \frac{B(t_0+h)-B(t_0)}{h^\alpha} \leq 1, \quad \text{and}$$
$$\inf_{h\in[0,1]} \frac{B(t_0+h)-B(t_0)}{h^\alpha} \geq -1$$

If $t_0 \in (k-1)2^n, \frac{k}{2^n})$ for $n > 2$, then the triangle inequality gives

$$|B(\frac{k+j}{2^n}) - B(\frac{k+j-1}{2^n})| \leq 2(\frac{j+1}{2^n})^\alpha.$$

Fix $l \geq 1/(\alpha - \frac{1}{2})$ and let $\Omega_{n,k}$ be the event

$$\left(|B(\tfrac{k+j}{2^n}) - B(\tfrac{k+j-1}{2^n})| \leq 2\left[\tfrac{(j+1)}{2^n}\right]^\alpha \text{ for } \text{j} = 1,2 \ldots \text{l}\right)$$

Then

$$\mathbb{P}(\Omega_{n,k}) \leq [\mathbb{P}(|B(1)| \leq 2^{n/2}\cdot 2\cdot(\frac{l+1}{2^n})^\alpha]^l \leq [2^{n/2}\cdot 2\cdot(\frac{l+1}{2^n})^\alpha]^l$$

since the normal density is less than 1/2. Hence

$$\mathbb{P}(\bigcup_{k=1}^{2^n}\Omega_{n,k}) \leq 2^n\cdot[2^{n/2}\cdot 2(\frac{l+1}{2^n})^\alpha]^l = C[2^{(1-l(\alpha-1/2))}]^n,$$

which sums. Thus

$$\mathbb{P}(\limsup_{n\to\infty}\bigcup_{k=1}^{2^n}\Omega_{n,k}) = 0.$$

EXERCISE 4.11. (Hard.) A.s. if $B(t_0) = \max_{0\leq t\leq 1} B(t)$ then $D^*B(t_0) = -\infty$.

The following are from week 3, (W 2/4, F 2/5, 1998).
Notes taken by Wei Li, `wli@stat.berkeley.edu`.

5. Hausdorff dimension and Minkowski dimension

DEFINITION 5.1 (Hausdorff 1918). Let A be a subset of a metric space. For $\alpha > 0$, and a possibly infinite $\epsilon > 0$ define a measure of size for A by

$$\mathcal{H}^\alpha_{(\epsilon)}(A) = \inf\{\sum_j |A_j|^\alpha : A \subset \bigcup_j A_j, |A_j| < \epsilon\},$$

where $|\cdot|$ applied to a set means its diameter. The quantity $\mathcal{H}_{(\infty)}(A)$ is called the **Hausdorff content** of A. Define the α-dimensional **Hausdorff measure** of A as

$$\mathcal{H}^\alpha(A) = \lim_{\epsilon\downarrow 0}\mathcal{H}^\alpha_{(\epsilon)}(A) = \sup_{\epsilon>0}\mathcal{H}^\alpha_{(\epsilon)}(A).$$

$\mathcal{H}^\alpha(\cdot)$ is a Borel measure. Sub-additivity is obvious since $\mathcal{H}^\alpha_{(\epsilon)}(A\cup D) \leq \mathcal{H}^\alpha_{(\epsilon)}(A) + \mathcal{H}^\alpha_{(\epsilon)}(D)$. Countable additivity can be shown with more effort. The graph of $\mathcal{H}^\alpha(A)$ versus α shows that there is a critical value of α where $\mathcal{H}^\alpha(A)$ jumps from ∞ to 0. This critical value is called the Hausdorff dimension of A.

DEFINITION 5.2. The **Hausdorff dimension** of A is

$$\dim_{\mathcal{H}}(A) = \inf\{\alpha : H^\alpha(A) = 0\}.$$

Note that given $\beta > \alpha > 0$ and $\mathcal{H}^{\alpha}_{(\epsilon)}(A) < \infty$, we have

$$\mathcal{H}^{\beta}_{(\epsilon)}(A) \leq \mathcal{H}^{\alpha}_{(\epsilon)}(A) \leq \epsilon^{(\beta-\alpha)}\mathcal{H}^{\alpha}(A).$$

The first inequality is obvious from the definition. The second one is due to the fact that $|A_j|^{\beta} \leq \epsilon^{(\beta-\alpha)}|A_j|^{\alpha}$ for a covering set A_j of diameter $\leq \epsilon$. Since ϵ is arbitrary here, we see that if $\mathcal{H}^{\alpha}(A) < \infty$, then $\mathcal{H}^{\beta}(A) = 0$. In general, we can also define the Hausdorff dimension of A as

$$\sup\{\alpha : \mathcal{H}^{\alpha}(A) = \infty\}.$$

Now let us look at another kind of dimension defined as follows.

Definition 5.3. The **upper and lower Minkowski dimension** (also called"Box" dimension) of A is defined as

$$\begin{aligned} \overline{\dim}_{\mathcal{M}}(A) &= \overline{\lim}_{\epsilon\downarrow 0}\frac{\log N_{\epsilon}(A)}{\log(1/\epsilon)}, \\ \underline{\dim}_{\mathcal{M}}(A) &= \underline{\lim}_{\epsilon\downarrow 0}\frac{\log N_{\epsilon}(A)}{\log(1/\epsilon)}, \end{aligned}$$

where $N_{\epsilon}(A)$ is the minimum number of balls of radius ϵ needed to cover A. If $\overline{\dim}_{\mathcal{M}}(A) = \underline{\dim}_{\mathcal{M}}(A)$, we call the common value the **Minkowski dimension** of A, denoted by $\dim_{\mathcal{M}}$.

Exercise 5.4. Show that $\dim_{\mathcal{M}}([0,1]^d) = d$.

Proposition 5.5. *For any set A, we have* $\dim_{\mathcal{H}}(A) \leq \underline{\dim}_{\mathcal{M}}(A)$.

Proof. By definition, $\mathcal{H}^{\alpha}_{(2\epsilon)} \leq N_{\epsilon}(A)(2\epsilon)^{\alpha}$. If $\alpha > \underline{\dim}_{\mathcal{M}}(A)$, then we also have $\underline{\lim} N_{\epsilon}(A)(2\epsilon)^{\alpha} = 0$. It then follows that $\dim_{\mathcal{H}}(A) \leq \underline{\dim}_{\mathcal{M}}(A)$. ∎

We now look at some examples to illustrate the above results. Let A be all the rational numbers in the unit interval. $\dim_{\mathcal{M}}(A) = 1$ since $\dim_{\mathcal{M}}(A) = \dim_{\mathcal{M}}(\overline{A})$ and $\dim_{\mathcal{M}}(A) = \dim_{\mathcal{M}}(\overline{A}) = \dim_{\mathcal{M}}([0,1])$. On the other hand, a countable set A always has $\dim_{\mathcal{H}}(A) = 0$. This is because for any $\alpha, \epsilon > 0$, we can cover the countable points in A by balls of radius $\epsilon/2$, $\epsilon/4$, $\epsilon/8$, and so on. Then

$$\mathcal{H}^{\alpha}_{(\epsilon)}(A) \leq \sum_j |A_j|^{\alpha} \leq \sum_j (\epsilon 2^{-j})^{\alpha} < \infty.$$

Another example is the harmonic sequence $A = \{0\}\bigcup\{1/n\}_{n\geq 1}$. It is a countable set, so it has Hausdorff dimension 0. It can be shown that $\dim_{\mathcal{M}}(A) = 1/2$. (Those points in A that are less than $\sqrt{\epsilon}$ can be covered by $1/\sqrt{\epsilon}$ balls of radius ϵ. The rest can be covered by $2/\sqrt{\epsilon}$ balls of radius ϵ, one on each point.)

We have shown in Lemma 4 that Brownian motion is β-Hölder for any $\beta < 1/2$ a.s. This will allow us to infer an upper bound on the Hausdorff dimension of its image and graph. Define the **graph** G_f of a function as the set of points $(t, f(t))$ as t ranges over the domain of f.

Proposition 5.6. *Let $f : [0,1] \to \mathbb{R}$ be a β-Hölder continuous function. Then*

$$\overline{\dim}_{\mathcal{M}}(G_f) \leq 2 - \beta.$$

PROOF. Since f is β-Hölder, there exists a constant C_1 such that, if $|t-s| \leq \epsilon$, then $|f(t)-f(s)| \leq C_1\epsilon^\beta = C_1\epsilon \cdot (1/\epsilon^{1-\beta})$. Hence, the minimum number of balls of radius ϵ to cover G_f satisfies $N_\epsilon(G_f) \leq C_2(1/\epsilon^{2-\beta})$ for some other constant C_2. ∎

COROLLARY 5.7.

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

A function $f : [0,1] \to \mathbb{R}$ is called "reverse" β-Hölder if there exists a constant $C > 0$ such that for any $[t,s]$, there is a subinterval $[t_1, s_1] \subset [t,s]$, such that $|f(t_1)-f(s_1)| \geq C|t_1-s_1|^\beta$.

PROPOSITION 5.8. *Let $f : [0,1] \to \mathbb{R}$ be β-Hölder and "reverse" β-Hölder. Then $\dim_{\mathcal{H}}(G_f) = 2-\beta$.*

REMARK. It follows from the above proposition that such a function f has the property $\dim_{\mathcal{H}}(G_f) > 1$. (Przytycki-Urbanalki, 1989.)

EXERCISE 5.9. Prove Proposition 5.8.

EXAMPLE 5.10. The Weierstrass nowhere differentiable function

$$W(t) = \sum_{n=0}^{\infty} a^n \cos(b^n t),$$

$ab > 1$, $0 < a < 1$ is β-Hölder and "reverse" β-Hölder for some $0 < \beta < 1$. For example, if $a = 1/2$, $b = 4$, then $\beta = 1/2$.

LEMMA 5.11. *Suppose X is a complete metric space. Let $f : X \to Y$ be β-Hölder. For any $A \subset X$, we have $\dim_{\mathcal{H}} f(A) \leq \dim_{\mathcal{H}}(A)/\beta$. Similar statements for $\overline{\dim}_{\mathcal{M}}$ and $\underline{\dim}_{\mathcal{M}}$ are also true.*

PROOF. If $\dim_{\mathcal{H}}(A) < \alpha < \infty$, then there exists a cover $\{A_j\}$ such that $A \subset \bigcup_j A_j$ and $\sum_j |A_j|^\alpha < \epsilon$. Then $\{f(A_j)\}$ is a cover for $f(A)$, and $|f(A_j)| \leq C|A_j|^\beta$, where C is the constant from the β-Hölder condition. Thus,

$$\sum_j |f(A_j)|^{\alpha/\beta} \leq C^{\alpha/\beta} \sum_j |A_j|^\alpha < C^{\alpha/\beta}\epsilon \to 0$$

as $\epsilon \to 0$, and hence $\dim_{\mathcal{H}} f(A) \leq \alpha/\beta$. ∎

COROLLARY 5.12. *For $A \subset [0,\infty)$, we have $\dim_{\mathcal{H}} B(A) \leq 2\dim_{\mathcal{H}}(A) \wedge 1$ a.s.*

6. Hausdorff dimension of Brownian motion and Brownian path

The nowhere differentiability of Brownian motion established in the previous section suggests that its graph has dimension higher than one. Recall that the graph G_f of a function f is the set of points $(x, f(x))$ where x ranges over the set where f is defined. Taylor (1953) showed that the graph of Brownian motion has Hausdorff dimension 3/2. The same holds in higher dimensions.

Define the d-dimensional standard Brownian motion whose coordinates are independent one dimensional standard Brownian motions. This measure is invariant under orthogonal transformations of $\mathbb{R}^d$, since Gaussian random variables are, see Lemma 2.3. For $d \geq 2$ it is interesting to look at the image set of Brownian motion. We will see that planar Brownian motion is neighborhood recurrent, that is it visits every neighborhood in the plane infinitely often. In this sense, the image of planar Brownian motion comparable to the plane itself; another sense in which this happens is that of Hausdorff dimension: the image of planar and higher dimensional Brownian motion has Hausdorff dimension two. Summing up, we will prove

THEOREM 6.1 (Taylor 1953). *Let B be $d-$ dimensional Brownian motion defined on the time set $[0,1]$. Then*

$$\dim_{\mathcal{H}} G_B = 3/2 \quad a.s.$$

Moreover, if $d \geq 2$, then

$$\dim_{\mathcal{H}} B[0,1] = 2 \quad a.s.$$

Higher dimensional Brownian motion therefore doubles the dimension of the time line. Naturally, the question arises whether this holds for subsets of the time line as well. In certain sense, this even holds for $d=1$: note the "$\wedge d$" in the following theorem.

THEOREM 6.2 (McKean 1955). *For every subset A of $[0,\infty)$, the image of A under d dimensional Brownian motion has Hausdorff dimension $2\dim_{\mathcal{H}} A \wedge d$ a.s.*

THEOREM 6.3 (Uniform Dimension Doubling (Kaufman 1969)).
Let B be Brownian motion in dimension at least 2. A.s, for any $A \subset [0,\infty)$, we have $\dim_{\mathcal{H}} B(A) = 2\dim_{\mathcal{H}}(A)$.

Notice the difference between the last two results. Kaufman's theorem has a much stronger claim: it states dimension doubling uniformly for all sets. For this theorem, $d \geq 2$ is a necessary condition: we will see later that the zero set of one dimensional Brownian motion has dimension half, while its image is the single point 0. We will prove Kaufman's theorem in a later section. For Theorem (6.1) we need the following fact.

THEOREM 6.4 (Mass Distribution Principle). *If $A \subset X$ supports a positive Borel measure μ such that $\mu(D) \leq C|D|^\alpha$ for any Borel set D, then $\mathcal{H}^\alpha_\infty(A) \geq \frac{\mu(A)}{C}$. This means that $\mathcal{H}^\alpha(A) \geq \frac{\mu(A)}{C}$, and hence $\dim_{\mathcal{H}}(A) \geq \alpha$.*

PROOF. If $A \subset \bigcup_j A_j$, then $\sum_j |A_j|^\alpha \geq C^{-1}\sum_j \mu(A_j) \geq C^{-1}\mu(A)$. ∎

EXAMPLE 6.5. Consider the middle third Cantor set $K_{1/3}$. At the nth level during the construction of $K_{1/3}$, we have 2^n intervals of length 3^{-n}. Thus, we have $\mathcal{H}^\alpha K_{1/3} \leq \lim 2^n 3^{-n\alpha}$. For the last expression to have value 1, α should be equal to $\log 2/\log 3$. Let μ be the Cantor-Lebesgue measure, which assigns mass 2^{-n} for any level n interval. Let D be a subset of $K_{1/3}$, and let n be the integer such that $3^{-n} \leq |D| \leq 3^{1-n}$. Then every such D is covered by at most four shorter terracing intervals of the form $[k3^{-n},(k+1)3^{-n}]$ for some k. We then have

$$\mu(D) \leq 4\cdot 2^{-n} \leq 4|D|^{-\alpha},$$

and so

$$\mathcal{H}^\alpha(K_{1/3}) \geq \mathcal{H}^\alpha_\infty(K_{1/3}) \geq 1/4.$$

Next, we introduce the energy method due to Frostman.

THEOREM 6.6 (Frostman 1935). *Given a metric space X, if μ is a finite Borel measure supported on $A \subset X$ and*

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

then $\mathcal{H}^\alpha_\infty(A) > 0$, and hence $\dim_{\mathcal{H}}(A) \geq \alpha$.

It can be shown that under the above conditions $\mathcal{H}^\alpha(A) = \infty$. The converse of this theorem is also true. That is, for any $\alpha < \dim_{\mathcal{H}}(A)$, there exists a measure μ on A that satisfies $\mathcal{E}_\alpha(\mu) < \infty$. A proof will be given later.

PROOF. Given a measure μ, define the function

$$\phi_\alpha(\mu, x) \stackrel{\text{def}}{=} \int \frac{d\mu(y)}{|x-y|^\alpha},$$

so that $\mathcal{E}_\alpha(\mu) = \int \phi_\alpha(\mu, x) d\mu(x)$. Let $A[M]$ denote the subset of A where $\phi_\alpha(\mu, \cdot)$ is at most M. There exists a number M such that $A[M]$ has positive μ-measure, since $\mathcal{E}_\alpha(\mu)$ is finite. Let ν denote the measure μ restricted to the set $A[M]$. Then for any $x \in A[M]$, we have $\phi_\alpha(\nu, x) \leq \phi_\alpha(\mu, x) \leq M$. Now let D be a bounded subset of the unit interval, and let $x \in D$. Let m be the largest integer such that D is contained in the open ball of radius 2^{-m} about x. Then

$$M \geq \int \frac{d\nu(x)}{|x-y|^\alpha} \geq \int_D \frac{d\nu(x)}{|x-y|^\alpha} \geq 2^{m\alpha}\nu(D).$$

The last inequality comes from the fact that $|x - y| \leq 2^{-m}$ for each y in D. Thus, we have $\nu(D) \leq M2^{-m\alpha} \leq M(2|D|)^\alpha$. By the Mass Distribution Principle, we conclude that $\mathcal{H}^\alpha_\infty(A) > \mathcal{H}^\alpha_\infty(A[M]) > 0$. ∎

Now we are ready to prove the second part of Taylor's theorem.

PROOF OF THEOREM 6.1, PART 2. Form Corollary 5.12 we have the inequality

$$\dim_{\mathcal{H}} B_d[0,1] \leq 2.$$

For the other inequality, we will use Frostman's energy method. Let $\alpha < 2$. We use 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,

$$\int f(x) d\mu_B(x) = \int_0^1 f(B_t) dt$$

for all measurable functions f. We want to show that

$$\mathbb{E}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \frac{d\mu_B(x)d\mu_B(y)}{|x-y|^\alpha} = \mathbb{E}\int_0^1\int_0^1 \frac{dsdt}{|B(t)-B(s)|^\alpha} < \infty. \tag{6.1}$$

Let us evaluate the 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 it is a finite constant c depending on d and α only. Substituting this expression into (6.1) and using Fubini's theorem we get

$$\mathcal{E}_\alpha(\mu_B) = c\int_0^1\int_0^1 \frac{dsdt}{|t-s|^{\alpha/2}} \leq 2c\int_0^1 \frac{du}{u^{\alpha/2}} < \infty. \quad \blacksquare \tag{6.2}$$

REMARK. Lévy showed earlier in 1940 that, when $d=2$, we have $\mathcal{H}^2(B[0,1]) = 0$ a.s. The statement is actually also true for all $d \geq 2$.

Now let us turn to the graph G_B of Brownian motion. We will show a proof of the first half of Taylor's theorem for one dimensional Brownian motion.

PROOF OF THEOREM 6.1, PART 1. We have shown in Corollary 5.7 that

$$\dim_{\mathcal{H}} G_B \leq 3/2.$$

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

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

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 we get that

$$\mathbb{E}\mathcal{E}_\alpha(\mu) \leq 2\int_0^1 \left(\mathbb{E}(t^2 + B(t)^2)^{-\alpha/2}\right) dt. \tag{6.3}$$

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

$$2\int_0^{+\infty} (t^2 + tz^2)^{-\alpha/2} n(z) dz. \tag{6.4}$$

Comparing the size of the summands in the integration suggests separating $z \leq \sqrt{t}$ from $z > \sqrt{t}$. Then we can bound (6.4) 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

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

The later integral is of order $t^{(1-\alpha)/2}$. Substituting these results into (6.3), we see that the expected energy is finite since $\alpha < 3/2$. The claim now follows form Frostman's Energy Method. $\blacksquare$

The following are from week 4, (W 2/11, F 2/13, 1998).
Notes taken by Charles Holton, `cholton@math.berkeley.edu`.

7. On nowhere differentiability

Lévy (1954) asks whether it is true that

$$\mathbb{P}\left[\forall t\, D^*B(t) \in \{\pm\infty\}\right] = 1.$$

We sketch a proof that almost surely there is an uncountable set of times t at which the upper right derivative $D^*B(t)$ is zero. Stronger and more general results can be found in Barlow and Perkins (1984). Put

$$I = \left[B(1), \sup_{0\le s\le 1} B(s)\right]$$

and define a function $g : I \to [0,1]$ by setting

$$g(x) = \sup\{s \in [0,1] : B(s) = x\}.$$

The interval I is almost surely non-degenerate, and g is strictly decreasing, right continuous and satisfies $B(g(x)) = x$ whenever B is continuous. Furthermore, the set of discontinuities of g is dense in I provided B has no interval of monotonicity. We restrict our attention to the event of probability 1 on which these assertions hold. Let

$$V_n = \{x \in I : g(x) - g(x+h) > nh \text{ for some } h \in (0, n^{-1})\}.$$

One readily verifies that V_n is open; it is also dense in I as every point of discontinuity of g is a limit from the left of points of V_n. By the Baire category theorem, $V := \bigcap_n V_n$ is uncountable and dense in I. Now if $x \in V$ then there is a sequence $x_n \downarrow x$ such that $g(x) - g(x_n) > n(x_n - x)$. Setting $t = g(x)$ and $t_n = g(x_n)$ we have $t_n \uparrow t$ and $t - t_n > n(B(t_n) - B(t))$, from which it follows that $D^*B(t) \ge 0$. On the other hand $D^*B(t) \le 0$ since $B(s) \le B(t)$ for all $s \in (t,1)$, by definition of $t = g(x)$.

EXERCISE 7.1. Let $f \in C([0,1])$. Prove that $B(t) + f(t)$ is nowhere differentiable almost surely.

Is the "typical" function in $C([0,1])$ nowhere differentiable? It is an easy application of the Baire category theorem to show that nowhere differentiability is a generic property for $C([0,1])$. This result leaves something to be desired, perhaps, as topological and measure theoretic notions of a "large" set need not coincide. For example, the set of points in $[0,1]$ whose binary expansion has zeros with asymptotic frequency $1/2$ is a meager set, yet it has Lebesgue measure 1. We consider a related idea proposed by Christensen (1972) and by Hunt, Sauer and Yorke (1992). Let X be a separable Banach space. Say that $A \subset X$ is **prevalent** if there exists a Borel probability measure μ on X such that $\mu(x + A) = 1$ for every $x \in X$. A set is called **negligible** if its complement is prevalent.

PROPOSITION 7.2. *If $A_1, A_2, \ldots$ are negligible subsets of X then $\bigcup_{i\ge 1} A_i$ is also negligible.*

PROOF. For each $i \ge 1$ let μ_{A_i} be a Borel probability measure satisfying $\mu_{A_i}(x + A_i) = 0$ for all $x \in X$. Using separability we can find for each i a ball D_i of radius 2^{-i} centered at $x_i \in X$ with $\mu_{A_i}(D_i) > 0$. Define probability measures μ_i, $i \ge 1$, by setting $\mu_i(E) = \mu_{A_i}(E+$

$x_i|D_i)$ for each Borel set E, so that $\mu_i(x + A_i) = 0$ for all x and for all i. Let $(Y_i; i \geq 0)$ be a sequence of independent random variables with $\mathrm{dist}(Y_i) = \mu_i$. Then $S = \sum_i Y_i$ converges almost surely. Writing μ for the distribution of S and putting $\nu_j = \mathrm{dist}(S - Y_j)$, we have $\mu = \mu_j * \nu_j$, and hence $\mu(x + A_j) = \mu_j * \nu_j(x + A_j) = 0$ for all x and for all j. Thus $\mu(x + \cup_{i\geq 1} A_i) = 0$ for all x. ■

PROPOSITION 7.3. *A subset A of $\mathbb{R}^d$ is negligible iff $\mathcal{L}_d(A) = 0$.*

PROOF. ($\Rightarrow$) Assume A is negligible. Let μ_A be a (Borel) probability measure such that $\mu_A(x + A) = 0$ for all $x \in X$. Since $\mathcal{L}_d * \mu_A = \mathcal{L}_d$ (indeed $\mathcal{L}_d * \mu = \mathcal{L}_d$ for any Borel probability measure μ on $\mathbb{R}^d$) we have $0 = \mathcal{L}_d * \mu_A(x + A) = \mathcal{L}_d(x + A)$ for all $x \in X$.

($\Leftarrow$) If $\mathcal{L}_d(A) = 0$ then the restriction of $\mathcal{L}_d$ to the unit cube is a probability measure which vanishes on every translate of A. ■

REMARK. It follows from Exercise 7.1 that the set of nowhere differentiable functions is prevalent in $C([0, 1])$.

8. Strong Markov property and the reflection principle

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, and set $\mathcal{F}_+(t) = \cap_{u>t} \mathcal{F}_0(u)$ (the right-continuous filtration). It is known (see, for example, Durrett (1996), theorem 7.2.4) that $\mathcal{F}_0(t)$ and $\mathcal{F}_+(t)$ have the same completion. A filtration $\{\mathcal{F}(t)\}_{t\geq 0}$ is a **Brownian filtration** if for all $t \geq 0$ the process $\{B(t + s)\}_{s\geq 0}$ is independent of $\mathcal{F}(t)$ and $\mathcal{F}(t) \supset \mathcal{F}_0(t)$. 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

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

PROPOSITION 8.1. *(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; see, e.g., Durrett (1996, 7.2.1) for details.

The main result of this section is the **strong Markov property** for Brownian motion, established independently by Hunt (1956) and Dynkin (1957):

THEOREM 8.2. *If 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. See, e.g., Durrett (1996, 7.3.7) for more details. ■

One important consequence of the strong Markov property is the following:

THEOREM 8.3 (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. We shall use an elementary fact:

LEMMA 8.4. *Let X, Y, Z be random variables with X, Y independent and X, Z independent. If $Y \overset{d}{\equiv} Z$ then $(X, Y) \overset{d}{\equiv} (X, Z)$.*

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 is also true of $\{-(B(\tau+t) - B(\tau))\}_{t\ge0}$. We see from the lemma that

$$(\{B(t)\}_{0\le t\le\tau}, \{B(t+\tau) - B(\tau)\}_{t\ge0}) \overset{d}{\equiv} (\{B(t)\}_{0\le t\le\tau}, \{-(B(\tau) - B(t+\tau))\}_{t\ge0}),$$

and the reflection principle follows immediately. ■

REMARK. Consider $\tau = \inf\{t : B(t) = \max_{0\le s\le1} B(s)\}$. Almost surely $\{B(\tau+t) - B(\tau)\}_{t\ge0}$ is non-positive on some interval about $t = 0$ and hence is *not* Brownian motion. The strong Markov property does not apply here because τ is not a stopping time for any Brownian filtration. We will later see that Brownian motion almost surely has no point of increase. Since τ is a point of increase of the reflected process $\{B^*(t)\}$, it follows that the distributions of Brownian motion and of $\{B^*(t)\}$ are singular.

EXERCISE 8.5. Prove that if A is a closed set then τ_A is a stopping time.

SOLUTION. $\{\tau_A \le t\} = \bigcap_{n\ge1} \bigcup_{s\in[0,t]\cap\mathbb{Q}} \{\mathrm{dist}(B(s), A) \le \frac{1}{n}\} \in \overline{\mathcal{F}_0(t)}$.

More generally, if A is a Borel set then the hitting time τ_A is a stopping time (see Bass (1995)).

Set $M(t) = \max_{0\le s\le t} B(s)$. Our next result says $M(t) \overset{d}{\equiv} |B(t)|$.

THEOREM 8.6. *If $a > 0$ then $\mathbb{P}[M(t) > a] = 2\mathbb{P}[B(t) > a]$.*

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 from the strong Markov property. ■

9. Local extrema of Brownian motion

PROPOSITION 9.1. *Almost surely, every local maximum of Brownian motion is a strict local maximum.*

For the proof we shall need

LEMMA 9.2. *Given two disjoint closed time intervals, the maxima of Brownian motion on them is different almost surely.*

PROOF. For $i = 1, 2$ let $[a_i, b_i]$, m_i, denote the lower, respectively the higher interval, and the corresponding maximum of Brownian motion. Note that $B(a_2) - B(b_1)$ is independent of the pair $m_1 - B(b_1)$ and $m_2 - B(a_2)$. Conditioning on the values of the later three random variables, The event $m_1 = m_2$ can be written as

$$B(a_2) - B(b_1) = m_1 - B(b_1) - (m_2 - B(a_2)).$$

The left hand side being a continuous random variable, and the right hand side a constant, we see that this event has probability 0. ∎

The statement of the lemma holds jointly for all pairs of intervals with rational endpoints. The proposition follows, since if Brownian motion had a non-strict local maximum, then there were two disjoint rational intervals where Brownian motion has the same maximum.

COROLLARY 9.3. *The set M of times where Brownian motion assumes its local maximum is countable and dense almost surely.*

PROOF. Consider the function from the set of non-degenerate closed intervals with rational endpoints to $\mathbb{R}$ given by

$$[a, b] \mapsto \inf\left\{t \geq a : B(t) = \max_{a \leq s \leq b} B(s)\right\}.$$

The image of this map contains the set M almost surely by the lemma. This shows that M is countable almost surely. We already know that B has no interval of increase or decrease almost surely. It follows that B almost surely has a local maximum in every interval with rational endpoints, implying the corollary. ∎

10. Area of planar Brownian motion paths

We have seen that the image of Brownian motion is always 2 dimensional, so one might ask what its 2 dimensional Hausdorff measure is. It turns out to be 0 in all dimensions; we will prove it for the planar case. We will need the following lemma.

LEMMA 10.1. *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. One proof of this fact relies on (outer) regularity of Lebesgue measure. The proof below is more streamlined.

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

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

Thus $\mathbf{1}_{A_1} * \mathbf{1}_{-A_2}(x) > 0$ on a set of positive area. But $\mathbf{1}_{A_1} * \mathbf{1}_{-A_2}(x) = 0$ unless $A_1 \cap (A_2 + x)$ has positive area, so this proves the lemma. ∎

Throughout this section B denote planar Brownian motion. We are now ready to prove Lévy's theorem on the area of its image.

THEOREM 10.2 (Lévy). *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

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

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

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

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

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

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.

Let $R(x)$ denote the area of $B[0,1] \cap (x + B[2,3] - B(2) + B(1))$. Since $B(2) - B(1)$ is independent of the triple $B[0,1]$, $B(1)$, and $B[2,3] - B(2)$, it must be independent of $R(x)$, so we have

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

where we are averaging 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, 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. ∎

The following are from week 5, (W 2/18, F 2/20, 1998).
Notes taken by Diogo Aguiar Gomes, `dgomes@math.berkeley.edu`.

11. Zeros of the Brownian motion

In this section, we start the study of the properties of the zero set Z_B of one dimensional Brownian motion. We will prove that this set is an uncountable closed set with no isolated points. This is, perhaps, surprising since, almost surely, a Brownian motion has isolated zeros from the left (for instance, the first zero after $1/2$) or from the right (the last zero before $1/2$). However, according to the next theorem, with probability one, it does not have any isolated zero.

THEOREM 11.1. *Let B be a one dimensional Brownian motion and Z_B be its zero set, i.e.,*

$$Z_B = \{t \in [0, +\infty) : B(t) = 0\}.$$

Then, a.s., Z_B is an uncountable closed set with no isolated points.

PROOF. Clearly, with probability one, Z_B is closed because B is continuous a.s.. To prove that no point of Z_B is isolated we consider the following construction: for each rational $q \in [0, \infty)$ consider the first zero after q, i.e., $\tau_q = \inf\{t > q : B(t) = 0\}$. Note that $\tau_q < \infty$ a.s. and, since Z_B is closed, the inf is a.s. a minimum. By the strong Markov property we have that for each q, a.s. τ_q is not an isolated zero from the right. But, since there are only countably many rationals, we conclude that a.s., for all q rational, τ_q is not an isolated zero from the right. Our next task is to prove that the remaining points of Z_B are not isolated from the left. So we claim that any $0 < t \in Z_B$ which is different from τ_q for all rational q is not an isolated point from the left. To see this take a sequence $q_n \uparrow t$. Define $t_n = \tau_{q_n}$. Clearly $q_n < t_n < t$ and so $t_n \uparrow t$. Thus t is not isolated from the left.

Finally, recall (see, for instance, Hewitt-Stromberg, 1965) that a closed set with no isolated points is uncountable and this finishes the proof. ∎

11.1. General Markov Processes. In this section, we define general Markov processes. Then we prove that Brownian motion, reflected Brownian motion and a process that involves the maximum of Brownian motion are Markov processes.

DEFINITION 11.2. A function $p(t, x, A)$, $p : \mathbb{R} \times \mathbb{R}^d \times \mathcal{B} \to \mathbb{R}$, where $\mathcal{B}$ is the Borel σ-algebra in $\mathbb{R}^d$, is a Markov transition kernel provided

1. $p(\cdot, \cdot, A)$ is measurable as a function of (t, x), for each $A \in \mathcal{B}$,
2. $p(t, x, \cdot)$ is a Borel probability measure for all $t \in \mathbb{R}$ and $x \in \mathbb{R}^d$,
3. $\forall A \in \mathcal{B},\ x \in \mathbb{R}^d$ and $t, s > 0$,

$$p(t + s, x, A) = \int_{\mathbb{R}^d} p(t, y, A) p(s, x, dy).$$

DEFINITION 11.3. A process $\{X(t)\}$ is a Markov process with kernel $p(t, x, A)$ if for all $t > s$ and Borel set $A \in \mathcal{B}$ we have

$$\mathbb{P}(X(t) \in A | \mathcal{F}_s) = p(t - s, X(s), A),$$

where $\mathcal{F}_s = \sigma(X(u),\ u \le s)$.

The next two examples are trivial consequences of the Markov Property for Brownian motion.

EXAMPLE 11.4. A d-dimensional Brownian motion is a Markov process and its transition kernel $p(t, x, \cdot)$ has $N(x, t)$ distribution in each component.

Suppose Y has $N(x, t)$ distribution. Define $|N(x, t)|$ to be the distribution of $|Y|$.

EXAMPLE 11.5. The reflected one-dimensional Brownian motion $|B(t)|$ is a Markov process. Moreover, its kernel $p(t, x, \cdot)$ has $|N(x, t)|$ distribution.

THEOREM 11.6 (Lévy, 1948). *Let $M(t)$ be the maximum process of a one dimensional Brownian motion $B(t)$, i.e. $M(t) = \max_{0 \le s \le t} B(s)$. Then, the process $Y(t) = M(t) - B(t)$ is Markov and its transition kernel $p(t, x, \cdot)$ has $|N(x, t)|$ distribution.*

PROOF. For $t > 0$, consider the two processes $\hat{B}(t) = B(s+t) - B(s)$ and $\hat{M}(t) = \max_{0 \le u \le t} \hat{B}(u)$. Define $\mathcal{F}_B(s) = \sigma(B(t), 0 \le t \le s)$. To prove the theorem it suffices to check that conditional on $\mathcal{F}_B(s)$ and $Y(s) = y$, we have $Y(s+t) \stackrel{\mathrm{d}}{=} |y + \hat{B}(t)|$.

Note that since $Y(s)$ is $\mathcal{F}_B(s)$ measurable this is the same as conditioning on $\mathcal{F}_B(s)$. To prove the claim note that $M(s+t) = M(s) \vee (B(s) + \hat{M}(t))$, and so we have

$$Y(s+t) = M(s) \vee (B(s) + \hat{M}(t)) - (B(s) + \hat{B}(t)).$$

Using the fact that $a \vee b - c = (a - c) \vee (b - c)$, we have

$$Y(s+t) = Y(s) \vee \hat{M}(t) - \hat{B}(t).$$

To finish, it suffices to check, for every $y \ge 0$, that $y \vee \hat{M}(t) - \hat{B}(t) \stackrel{\mathrm{d}}{=} |y + \hat{B}(t)|$. For any $a \ge 0$ write

$$\mathbb{P}(y \vee \hat{M}(t) - \hat{B}(t) > a) = I + II,$$

where $I = \mathbb{P}(y - \hat{B}(t) > a)$ and

$$II = \mathbb{P}(y - \hat{B}(t) \le a \text{ and } \hat{M}(t) - \hat{B}(t) > a).$$

Since $\hat{B} \stackrel{\mathrm{d}}{=} -\hat{B}$ we have

$$I = \mathbb{P}(y + \hat{B}(t) > a).$$

To study the second term is useful to define the "time reversed" Brownian motion

$$W(u) = \hat{B}(t-u) - \hat{B}(t),$$

for $0 \le u \le t$. Note that W is also a Brownian motion for $0 \le u \le t$ since it is continuous and its finite dimensional distributions are Gaussian with the right covariances.

Let $M_W(t) = \max_{0 \le u \le t} W(u)$. Then $M_W(t) = \hat{M}(t) - \hat{B}(t)$, since $W(t) = -\hat{B}(t)$. Thus

$$II = \mathbb{P}(y + W(t) \le a \text{ and } M_W(t) > a).$$

If we use the reflection principle by reflecting $W(u)$ at the first time it hits a we get another Brownian motion $W^*(u)$. In terms of this Brownian motion we have $II = \mathbb{P}(W^*(t) \ge a+y)$. Since $W^* \stackrel{\mathrm{d}}{=} -\hat{B}$, it follows $II = \mathbb{P}(y + \hat{B}(t) \le -a)$. The Brownian motion $\hat{B}(t)$ has continuous distribution, and so, by adding I and II, we get

$$\mathbb{P}(y \vee \hat{M}(t) - \hat{B}(t) > a) = \mathbb{P}(|y + \hat{B}(t)| > a).$$

This proves the claim and, consequently, the theorem. ∎

PROPOSITION 11.7. *Two Markov processes in $\mathbb{R}^d$ with continuous paths, with the same initial distribution and transition kernel are identical in law.*

OUTLINE OF PROOF. The finite dimensional distributions are the same. From this we deduce that the restriction of both processes to rational times agree in distribution. Finally we can use continuity of paths to prove that they agree, as processes, in distribution (see Freedman 1971 for more details). ∎

Since the process $Y(t)$ is continuous and has the same distribution as $|B(t)|$ (they have the same Markov transition kernel and same initial distribution) this proposition implies $\{Y(t)\} \stackrel{\mathrm{d}}{=} \{|B(t)|\}$.

11.2. Hausdorff dimension of Z_B. We already know that Z_B is an uncountable set with no isolated points. In this section, we will prove that, with probability one, the Hausdorff dimension of Z_B is $1/2$. It turns out that is relatively easy to bound from bellow the dimension of the zero set of $Y(t)$ (also known as set of record values of B). Then, by the results in the last section, this dimension must be the same of Z_B since these two (random) sets have the same distribution.

DEFINITION 11.8. A time t is a record time for B if $Y(t) = M(t) - B(t) = 0$, i.e., if t is a global maximum from the left.

The next lemma gives a lower bound on the Hausdorff dimension of the set of record times.

LEMMA 11.9. *With probability 1,* $\dim\{t \in [0,1] : Y(t) = 0\} \geq 1/2$.

PROOF. Since $M(t)$ is an increasing function, we can regard it as a distribution function of a measure μ, with $\mu(a,b] = M(b) - M(a)$. This measure is supported on the set of record times. We know that, with probability one, the Brownian motion is Hölder continuous with any exponent $\alpha < 1/2$. Thus

$$M(b) - M(a) \leq \max_{0 \leq h \leq b-a} B(a+h) - B(a) \leq C_\alpha (b-a)^\alpha,$$

where $\alpha < 1/2$ and C_α is some random constant that doesn't depend on a or b. By the Mass Distribution Principle, we get that, a.s., $\dim\{t \in [0,1] : Y(t) = 0\} \geq \alpha$. By choosing a sequence $\alpha_n \uparrow 1/2$ we finish the proof. ■

Recall that the upper Minkowski dimension of a set is an upper bound for the Hausdorff dimension. To estimate the Minkowski dimension of Z_B we will need to know

$$\mathbb{P}(\exists t \in (a, a+\epsilon) : B(t) = 0). \tag{11.1}$$

This probability can be computed explicitly and we will leave this as an exercise.

EXERCISE 11.10. Compute (11.1).

SOLUTION. Conditional on $B(a) = x > 0$ we have

$$\mathbb{P}(\exists t \in (a, a+\epsilon) : B(t) = 0 | B(a) = x) = \mathbb{P}(\min_{a \leq t \leq a+\epsilon} B(t) < 0 | B(a) = x).$$

But the right hand side is equal to

$$\mathbb{P}(\max_{0<t<\epsilon} B(t) > x) = 2\mathbb{P}(B(\epsilon) > x),$$

using reflection principle.
By considering also the case where x is negative we get

$$\mathbb{P}(\exists t \in (a, a+\epsilon) : B(t) = 0) = 4\int_0^\infty \int_x^\infty \frac{e^{-\frac{y^2}{2\epsilon} - \frac{x^2}{2a}}}{2\pi\sqrt{a\epsilon}} dy dx.$$

Computing this last integral explicitly, we get

$$\mathbb{P}(\exists t \in (a, a+\epsilon) : B(t) = 0) = \frac{2}{\pi} \arctan \sqrt{\frac{\epsilon}{a}}.$$

However, for our purposes, an estimate will suffice and we will disregard the solution of the exercise and prove a (much simpler) lemma.

LEMMA 11.11. *For any $a, \epsilon > 0$ we have*

$$\mathbb{P}(\exists t \in (a, a+\epsilon) : B(t) = 0) \leq C\sqrt{\frac{\epsilon}{a+\epsilon}},$$

for some appropriate positive constant c.

PROOF. Consider the event A given by $|B(a+\epsilon)| \leq \sqrt{\epsilon}$. By the scaling property of the Brownian motion, we can give the upper bound

$$\mathbb{P}A = \mathbb{P}\left(|B(1)| \leq \sqrt{\frac{\epsilon}{a+\epsilon}}\right) \leq 2\sqrt{\frac{\epsilon}{a+\epsilon}}. \tag{11.2}$$

However, knowing that Brownian motion has a zero in $(a, a+\epsilon)$ makes the event A very likely. Indeed, we certainly have

$$\mathbb{P}A \geq \mathbb{P}(A \text{ and } 0 \in B[a, a+\epsilon]),$$

and the strong Markov property implies that

$$\mathbb{P}A \geq \tilde{c}\mathbb{P}(0 \in B[a, a+\epsilon]), \tag{11.3}$$

where

$$\tilde{c} = \min_{a \leq t \leq a+\epsilon} \mathbb{P}(A | B(t) = 0).$$

Because the minimum is achieved when $t = a$ have

$$\tilde{c} = \mathbb{P}(|B(1)| \leq 1) > 0,$$

by using the scaling property of the Brownian motion.

From inequalities (11.2) and (11.3), we conclude

$$\mathbb{P}(0 \in B[a, a+\epsilon]) \leq \frac{2}{\tilde{c}}\sqrt{\frac{\epsilon}{a+\epsilon}}. \quad \blacksquare$$

For any, possibly random, closed set $A \subset [0,1]$ define a function

$$N_m(A) = \sum_{k=1}^{2^m} 1_{\left\{A \cap [\frac{k-1}{2^m}, \frac{k}{2^m}] \neq \emptyset\right\}}.$$

This function counts the number of intervals of the form $[\frac{k-1}{2^m}, \frac{k}{2^m}]$ intersected by the set A and so is a natural object if we want to compute Minkowski dimension. In the special case where $A = Z_B$ we have

$$N_m(Z_B) = \sum_{k=1}^{2^m} 1_{\left\{0 \in B[\frac{k-1}{2^m}, \frac{k}{2^m}]\right\}}.$$

The next lemma shows that estimates on the expected value of $N_m(A)$ will give us bounds on the Minkowski dimension (and hence in the Hausdorff dimension).

LEMMA 11.12. *Suppose A is a closed random subset of $[0,1]$ such that*

$$EN_m(A) \leq c2^{m\alpha},$$

for some $c, \alpha > 0$. Then $\underline{dim}_M(A), \overline{dim}_M(A) \leq \alpha$.

PROOF. By Fatou's lemma we have

$$c \geq \liminf \frac{EN_m(A)}{2^{m\alpha}} \geq E \liminf \frac{N_m(A)}{2^{m\alpha}},$$

and so, a.s.,

$$\liminf \frac{N_m(A)}{2^{m\alpha}} < \infty.$$

This proves $\underline{dim}_M(A) \leq \alpha$. To estimate $\overline{dim}_M(A)$ consider

$$E \sum_{m=1}^{\infty} \frac{N_m(A)}{2^{m(\alpha+\epsilon)}},$$

for $\epsilon > 0$. Then, by monotone convergence theorem,

$$E \sum_{m=1}^{\infty} \frac{N_m(A)}{2^{m(\alpha+\epsilon)}} = \sum_{m=1}^{\infty} \frac{EN_m(A)}{2^{m(\alpha+\epsilon)}} < \infty.$$

This estimate implies that

$$\sum_{m=1}^{\infty} \frac{N_m(A)}{2^{m(\alpha+\epsilon)}} < \infty \quad \text{a.s.},$$

and so, with probability one,

$$\limsup_{m\to\infty} \frac{N_m(A)}{2^{m(\alpha+\epsilon)}} = 0.$$

From the last equation follows

$$\overline{dim}_M(A) \leq \alpha + \epsilon, \quad \text{a.s..}$$

Let $\epsilon \to 0$ through some countable sequence to get

$$\overline{dim}_M(A) \leq \alpha, \quad \text{a.s..}$$

And this completes the proof of the lemma. ∎

To get an upper bound on the Hausdorff dimension of Z_B note that

$$EN_m(Z_B) \leq C \sum_{k=1}^{2^m} \frac{1}{\sqrt{k}} \leq \tilde{C} 2^{m/2},$$

since $\mathbb{P}\left(\exists t \in \left[\frac{k-1}{2^m}, \frac{k}{2^m}\right] : B(t) = 0\right) \leq \frac{C}{\sqrt{k}}$. Thus, by the last lemma, $\overline{dim}_M(Z_B) \leq 1/2$ a.s. This implies immediately $\dim_H(Z_B) \leq 1/2$ a.s. Combining this estimate with Lemma 11.9 we have

THEOREM 11.13. *With probability one we have*

$$\dim_H(Z_B) = \frac{1}{2}.$$

From this proof we can also infer that $\mathcal{H}^{1/2}(Z_B) < \infty$ a.s. Later in the course we will prove that $\mathcal{H}^{1/2}(Z_B) = 0$. However, it is possible to define a more sophisticated Hausdorff measure for which, with probability one, $0 < \mathcal{H}(Z_B) < \infty$.

The following are from week 6, (W 3/4, F 3/6, 1998).
Notes taken by Jason Schweinsberg, `jason@stat.berkeley.edu`.

12. Harris' Inequality and its consequences

We begin this section by proving Harris' inequality.

LEMMA 12.1 (Harris' inequality). *Suppose that $\mu_1, \ldots, \mu_d$ are Borel probability measures on $\mathbb{R}$ and $\mu = \mu_1 \times \mu_2 \times \ldots \times \mu_d$. Let $f, g : \mathbb{R}^d \to \mathbb{R}$ be measurable functions that are nondecreasing in each coordinate. Then,*

$$\int_{\mathbb{R}^d} f(x)g(x)\,d\mu \geq \left(\int_{\mathbb{R}^d} f(x)\,d\mu\right)\left(\int_{\mathbb{R}^d} g(x)\,d\mu\right), \tag{12.1}$$

provided the above integrals are well-defined.

PROOF. One can argue, using the Monotone Convergence Theorem, that it suffices to prove the result when f and g are bounded. We assume f and g are bounded and proceed by induction. Suppose $d = 1$. Note that

$$(f(x) - f(y))(g(x) - g(y)) \geq 0$$

for all $x, y \in \mathbb{R}$. Therefore,

$$\begin{aligned} 0 &\leq \int_{\mathbb{R}}\int_{\mathbb{R}} (f(x) - f(y))(g(x) - g(y))\,d\mu(x)d\mu(y) \\ &= 2\int_{\mathbb{R}} f(x)g(x)\,d\mu(x) - 2\left(\int_{\mathbb{R}} f(x)\,d\mu(x)\right)\left(\int_{\mathbb{R}} g(y)\,d\mu(y)\right), \end{aligned}$$

and (12.1) follows easily. Now, suppose (12.1) holds for $d-1$. Define

$$f_1(x_1) = \int_{\mathbb{R}^{d-1}} f(x_1, \ldots, x_d)\,d\mu(x_2)\ldots d\mu_d(x_d),$$

and define g_1 similarly. Note that $f_1(x_1)$ and $g_1(x_1)$ are increasing functions of x_1. Since f and g are bounded, we may apply Fubini's Theorem to write the left hand side of (12.1) as

$$\int_{\mathbb{R}}\left(\int_{\mathbb{R}^{d-1}} f(x_1, \ldots, x_d)g(x_1, \ldots, x_d)\,d\mu_2(x_2)\ldots d\mu_d(x_d)\right)\,d\mu_1(x_1). \tag{12.2}$$

The integral in the parenthesis is at least $f_1(x_1)g_1(x_1)$ by the induction hypothesis. Thus, by using the result for the $d = 1$ case we can bound (12.2) below by

$$\left(\int_{\mathbb{R}} f_1(x_1)\,d\mu_1(x_1)\right) \times \left(\int_{\mathbb{R}} g_1(x_1)\,d\mu_1(x_1)\right)$$

which equals the right hand side of (12.1), completing the proof.

REMARK. We say an event $A \subset \mathbb{R}^d$ is an **increasing event** if when $\tilde{x}_i \geq x_i$ and $(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots x_d) \in A$, we have $(x_1, \ldots, x_{i-1}, \tilde{x}_i, x_{i+1}, \ldots x_d) \in A$. If A and B are increasing events, then it is easy to see by applying Harris' Inequality to the indicator functions 1_A and 1_B that $\mathbb{P}(A \cap B) \geq \mathbb{P}(A)\mathbb{P}(B)$.

REMARK. Let $X_1, \dots, X_n$ be an i.i.d. sample, where each X_i has distribution μ. Given any $(x_1, \dots, x_n) \in \mathbb{R}^n$, define the relabeling $x_{(1)} \geq x_{(2)} \geq \dots \geq x_{(n)}$. Fix i and j, and define $f(x_1, \dots, x_n) = x_{(i)}$ and $g(x_1, \dots, x_n) = x_{(j)}$. Then f and g are measurable and nondecreasing in each component. Therefore, if $X_{(i)}$ and $X_{(j)}$ denote the ith and jth order statistics of $X_1, \dots, X_n$, then it follows from Harris' Inequality that $\mathbb{E}[X_{(i)}X_{(j)}] \geq \mathbb{E}[X_{(i)}]\mathbb{E}[X_{(j)}]$, provided these expectations are well-defined. See Lehmann (1966) and Bickel (1967) for further discussion.

For the rest of this section, let $X_1, X_2, \dots$ be i.i.d. random variables, and let $S_k = \sum_{i=1}^k X_i$ be their partial sums. Denote

$$p_n = \mathbb{P}(S_i \geq 0 \text{ for all } \ 1 \leq i \leq n)\,. \tag{12.3}$$

Observe that the event that $\{S_n$ is largest among $S_0, S_1, \dots S_n\}$ is precisely the event that the reversed random walk $X_n + \dots + X_{n-k+1}$ is nonnegative for all $k = 1, \dots, n$; thus this event also has probability p_n. The following theorem gives the order of magnitude of p_n.

THEOREM 12.2. *If the increments X_i have a symmetric distribution (that is, $X_i =_d -X_i$) or have mean zero and finite variance, then there are positive constants C_1 and C_2 such that $C_1 n^{-1/2} \leq p_n \leq C_2 n^{-1/2}$ for all $n \geq 1$.*

PROOF. For the general argument, see Feller (1966), Section XII.8. We prove the result here for the simple random walk, that is when each X_i takes values ± 1 with probability half each.

Define the stopping time $\tau_{-1} = \min\{k : S_k = -1\}$. Then

$$p_n = \mathbb{P}(S_n \geq 0) - \mathbb{P}(S_n \geq 0, \tau_{-1} < n).$$

Let $\{S_j^*\}$ denote the random walk reflected at time τ_{-1}, that is

$$\begin{array}{ll} S_j^* = S_j & \text{for } j \leq \tau_{-1}, \\ S_j^* = (-1) - (S_j + 1) & \text{for } j > \tau_{-1}. \end{array}$$

Note that if $\tau_{-1} < n$ then $S_n \geq 0$ if and only if $S_n^* \leq -2$, so

$$p_n = \mathbb{P}(S_n \geq 0) - \mathbb{P}(S_n^* \leq -2).$$

Using symmetry and the reflection principle, we have

$$p_n = \mathbb{P}(S_n \geq 0) - \mathbb{P}(S_n \geq 2) = \mathbb{P}(S_n \in \{0, 1\}),$$

which means that

$$\begin{array}{lllll} p_n & = & \mathbb{P}(S_n = 0) & = & \binom{n}{n/2} 2^{-n} \quad \text{for } n \text{ even}, \\ p_n & = & \mathbb{P}(S_n = 1) & = & \binom{n}{(n-1)/2} \quad \text{for } n \text{ odd}. \end{array}$$

Recall that Stirling's Formula gives $m! \sim \sqrt{2\pi} m^{m+1/2} e^{-m}$, where the symbol $\sim$ means that the ratio of the two sides approaches 1 as $m \to \infty$. One can deduce from Stirling's Formula that

$$p_n \sim \sqrt{\frac{2}{\pi n}},$$

which proves the theorem. ∎

The following theorem expresses, in terms of the p_i, the probability that S_j stays between 0 and S_n for j between 0 and n.

THEOREM 12.3. *We have* $p_n^2 \le \mathbb{P}(0 \le S_j \le S_n \text{ for all } 1 \le j \le n) \le p_{\lfloor n/2 \rfloor}^2$.

PROOF. The two events

$$\begin{aligned} A &= \{0 \le S_j \text{ for all } j \le n/2\} \text{ and} \\ B &= \{S_j \le S_n \text{ for } j \ge n/2\} \end{aligned}$$

are independent, since A depends only on $X_1, \dots, X_{\lfloor n/2 \rfloor}$ and B depends only on the remaining $X_{\lfloor n/2 \rfloor + 1}, \dots, X_n$. Therefore,

$$\mathbb{P}(0 \le S_j \le S_n) \le \mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B) \le p_{\lfloor n/2 \rfloor}^2,$$

which proves the upper bound.

For the lower bound, we follow Peres (1996) and let $f(x_1, \dots, x_n) = 1$ if all the partial sums $x_1 + \ldots + x_k$ for $k = 1, \dots, n$ are nonnegative, and $f(x_1, \dots, x_n) = 0$ otherwise. Also, define $g(x_1, \dots, x_n) = f(x_n, \dots, x_1)$. Then f and g are nondecreasing in each component. Let μ_j be the distribution of X_j, and let $\mu = \mu_1 \times \ldots \times \mu_n$. By Harris' Inequality,

$$\int_{\mathbb{R}^n} fg \, d\mu \ge \left(\int_{\mathbb{R}^n} f \, d\mu \right) \left(\int_{\mathbb{R}^n} g \, d\mu \right) = p_n^2.$$

Also,

$$\begin{aligned} \int_{\mathbb{R}^n} fg \, d\mu &= \int_{\mathbb{R}^n} 1\{ \text{ for all } j, \quad x_1 + \ldots + x_j \ge 0 \text{ and } \quad x_{j+1} + \ldots + x_n \ge 0\} \, d\mu \\ &= \mathbb{P}(0 \le S_j \le S_n \text{ for all } \ 1 \le j \le n), \end{aligned}$$

which proves the lower bound. ∎

13. Points of increase for random walks and Brownian motion

The material in this section has been taken, with only minor modifications, from Peres (1996).

A real-valued function f has a **global point of increase in the interval** $(\mathbf{a}, \mathbf{b})$ if there is a point $t_0 \in (a, b)$ such that $f(t) \le f(t_0)$ for all $t \in (a, t_0)$ and $f(t_0) \le f(t)$ for all $t \in (t_0, b)$. We say t_0 is a **local point of increase** if it is a global point of increase in some interval. In figure 6.1 below, t_0 is a global point of increase for f and t_1 is a local point of increase. Dvoretzky, Erdős and Kakutani (1961) proved that Brownian motion almost surely has no global points of increase in any time interval, or, equivalently, that Brownian motion has no local points of increase. Knight (1981) and Berman (1983) noted that this follows from properties of the local time of Brownian motion; direct proofs were given by Adelman (1985) and Burdzy (1990). Here we show that the nonincrease phenomenon holds for arbitrary symmetric random walks, and can thus be viewed as a combinatorial consequence of fluctuations in random sums.

DEFINITION. Say that a sequence of real numbers $s_0, s_1, \dots, s_n$ has a (global) **point of increase** at k if $s_i \le s_k$ for $i = 0, 1, \dots, k-1$ and $s_k \le s_j$ for $j = k+1, \dots, n$.

THEOREM 13.1. *Let $S_0, S_1, \ldots, S_n$ be a random walk where the i.i.d. increments $X_i = S_i - S_{i-1}$ have a symmetric distribution,* or *have mean 0 and finite variance. Then*

$$\mathbb{P}(S_0, \ldots, S_n \ \textit{has a point of increase}) \leq \frac{C}{\log n} \ ,$$

for $n > 1$, where C does not depend on n.

We will now see how this result implies the following

COROLLARY 13.2. *Brownian motion almost surely has no points of increase.*

PROOF. To deduce this, it suffices to apply Theorem 13.1 to a simple random walk on the integers. Indeed it clearly suffices to show that the Brownian motion $\{B(t)\}_{t\geq 0}$ almost surely has no global points of increase in a fixed rational time interval (a, b). Sampling the Brownian motion when it visits a lattice yields a simple random walk; by refining the lattice, we may make this walk as long as we wish, which will complete the proof. More precisely, for any vertical spacing $h > 0$ define τ_0 to be the first $t \geq a$ such that $B(t)$ is an integral multiple of h, and for $i \geq 0$ let τ_{i+1} be the minimal $t \geq \tau_i$ such that $|B(t) - B(\tau_i)| = h$. Define $N_b = \max\{k \in \mathbb{Z} : \tau_k \leq b\}$. For integers i satisfying $0 \leq i \leq N_b$, define

$$S_i = \frac{B(\tau_i) - B(\tau_0)}{h}.$$

Then, $\{S_i\}_{i=1}^{N_b}$ is a finite portion of a simple random walk. If the Brownian motion has a (global) point of increase in (a, b) at t_0, and if k is an integer such that $\tau_{k-1} \leq t_0 \leq \tau_k$, then this random walk has points of increase at $k-1$ and k. We have, for all n,

$$\begin{aligned} &\mathbb{P}(\text{ Brownian motion has a global point of increase in } (a,b)) \qquad (13.1) \\ \leq \ & \mathbb{P}(N_b \leq n) + \sum_{m=n+1}^{\infty} \mathbb{P}(S_0, \ldots, S_m \text{ has a point of increase, and } N_b = m). \end{aligned}$$

Note that $N_b \leq n$ implies $|B(b) - B(a)| \leq (n+1)h$, so

$$\mathbb{P}(N_b \leq n) \leq \mathbb{P}(|B(b) - B(a)| \leq (n+1)h) = \mathbb{P}\left(|Z| \leq \frac{(n+1)h}{\sqrt{b-a}}\right),$$

where Z has a standard normal distribution. Since $S_0, \ldots, S_m$, conditioned on $N_b = m$ is a finite portion of a simple random walk, it follows from Theorem 13.1 that for some constant C, we have

$$\begin{aligned} &\sum_{m=n+1}^{\infty} \mathbb{P}(S_0, \ldots, S_m \text{ has a point of increase, and } N_b = m) \\ \leq \ & \sum_{m=n+1}^{\infty} \mathbb{P}(N_b = m) \frac{C}{\log m} \leq \frac{C}{\log(n+1)}. \end{aligned}$$

Thus, the probability in (13.1) can be made arbitrarily small by first taking n large and then picking $h > 0$ sufficiently small. ∎

THEOREM 13.3. *For any random walk $\{S_j\}$ on the line,*

$$\mathbb{P}(S_0,\dots,S_n \text{ has a point of increase}) \le 2\frac{\sum_{k=0}^{n} p_k p_{n-k}}{\sum_{k=0}^{\lfloor n/2\rfloor} p_k^2}\,. \tag{13.2}$$

PROOF. The idea is simple. The expected number of points of increase is the numerator in (13.3), and given that there is at least one such point, the expected number is bounded below by the denominator; the ratio of these expectations bounds the required probability.

To carry this out, denote by $I_n(k)$ the event that k is a point of increase for $S_0, S_1, \dots, S_n$ and by $F_n(k) = I_n(k) \setminus \cup_{i=0}^{k-1} I_n(i)$ the event that k is the first such point. The events that $\{S_k$ is largest among $S_0, S_1, \dots S_k\}$ and that $\{S_k$ is smallest among $S_k, S_{k+1}, \dots S_n\}$ are independent, and therefore $\mathbb{P}(I_n(k)) = p_k p_{n-k}$.

Observe that if S_j is minimal among $S_j, \dots, S_n$, then any point of increase for $S_0, \dots, S_j$ is automatically a point of increase for $S_0, \dots, S_n$. Therefore for $j \le k$ we can write

$$\begin{aligned} &F_n(j) \cap I_n(k) = \\ &\quad F_j(j) \cap \{S_j \le S_i \le S_k \text{ for all } i \in [j,k]\} \cap \{S_k \text{ is minimal among } S_k, \dots, S_n\}\,. \end{aligned}$$

The three events on the right-hand side are independent, as they involve disjoint sets of summands; the second of these events is of the type considered in Theorem 12.3. Thus,

$$\begin{aligned} \mathbb{P}(F_n(j) \cap I_n(k)) &\ge \mathbb{P}(F_j(j))\, p_{k-j}^2\, p_{n-k} \\ &\ge p_{k-j}^2\, \mathbb{P}(F_j(j))\, \mathbb{P}\left(S_j \text{ is minimal among } S_j, \dots, S_n\right)\,, \end{aligned}$$

since $p_{n-k} \ge p_{n-j}$. Here the two events on the right are independent, and their intersection is precisely $F_n(j)$. Consequently $\mathbb{P}(F_n(j) \cap I_n(k)) \ge p_{k-j}^2 \mathbb{P}(F_n(j))$.

Decomposing the event $I_n(k)$ according to the first point of increase gives

$$\begin{aligned} \sum_{k=0}^{n} p_k p_{n-k} &= \sum_{k=0}^{n} \mathbb{P}(I_n(k)) = \sum_{k=0}^{n}\sum_{j=0}^{k} \mathbb{P}(F_n(j) \cap I_n(k)) \\ &\ge \sum_{j=0}^{\lfloor n/2\rfloor} \sum_{k=j}^{j+\lfloor n/2\rfloor} p_{k-j}^2 \mathbb{P}(F_n(j)) \ge \sum_{j=0}^{\lfloor n/2\rfloor} \mathbb{P}(F_n(j)) \sum_{i=0}^{\lfloor n/2\rfloor} p_i^2\,. \end{aligned} \tag{13.3}$$

This yields an upper bound on the probability that $\{S_j\}_{j=0}^n$ has a point of increase by time $n/2$; but this random walk has a point of increase at time k if and only if the "reversed" a $\{S_n - S_{n-i}\}_{i=0}^n$ has a point of increase at time $n-k$. Thus, doubling the upper bound given by (13.3) proves the theorem. ∎

PROOF OF THEOREM 13.1. To bound the numerator in (13.3), we can use symmetry to deduce from Theorem 12.2 that

$$\begin{aligned} \sum_{k=0}^{n} p_k p_{n-k} &\le 2 + 2\sum_{k=1}^{\lfloor n/2\rfloor} p_k p_{n-k} \\ &\le 2 + 2C_2^2 \sum_{k=1}^{\lfloor n/2\rfloor} k^{-1/2}(n-k)^{-1/2} \le 2 + 4C_2^2 n^{-1/2} \sum_{k=1}^{\lfloor n/2\rfloor} k^{-1/2}\,, \end{aligned}$$

which is bounded above because the last sum is $O(n^{1/2})$. Since Theorem 12.2 implies that the denominator in (13.2) is at least $C_1^2 \log\lfloor n/2 \rfloor$, this completes the proof. ∎

The following theorem shows that we can obtain a lower bound on the probability that a random walk has a point of increase that differs from the upper bound only by a constant.

THEOREM 13.4. *For any random walk on the line,*

$$\mathbb{P}(S_0, \ldots, S_n \text{ has a point of increase}) \geq \frac{\sum_{k=0}^{n} p_k p_{2n-k}}{2\sum_{k=0}^{\lfloor n/2 \rfloor} p_k^2} . \tag{13.4}$$

In particular if the increments have a symmetric distribution, or *have mean 0 and finite variance, then* $\mathbb{P}(S_0, \ldots, S_n$ *has a point of increase*$) \asymp 1/\log n$ *for* $n > 1$, *where the symbol* $\asymp$ *means that the ratio of the two sides is bounded above and below by positive constants that do not depend on* n.

PROOF. Using (13.3), we get

$$\sum_{k=0}^{n} p_k p_{2n-k} = \sum_{k=0}^{n} \mathbb{P}(I_{2n}(k)) = \sum_{k=0}^{n} \sum_{j=0}^{k} \mathbb{P}(F_{2n}(j) \cap I_{2n}(k)) .$$

Using Theorem 12.3, we see that for $j \leq k \leq n$, we have

$$\begin{aligned} \mathbb{P}(F_{2n}(j) \cap I_{2n}(k)) &\leq \mathbb{P}(F_n(j) \cap \{S_j \leq S_i \leq S_k \text{ for } j \leq i \leq k\}) \\ &\leq \mathbb{P}(F_n(j)) p^2_{\lfloor (k-j)/2 \rfloor} . \end{aligned}$$

Thus,

$$\sum_{k=0}^{n} p_k p_{2n-k} \leq \sum_{k=0}^{n} \sum_{j=0}^{k} \mathbb{P}(F_n(j)) p^2_{\lfloor (k-j)/2 \rfloor} \leq \sum_{j=0}^{n} \mathbb{P}(F_n(j)) \sum_{i=0}^{n} p^2_{\lfloor i/2 \rfloor} .$$

This implies (13.4). The assertion concerning symmetric or mean 0, finite variance walks follows from Theorem 12.2 and the proof of Theorem 13.1. ∎

13.1. Further Discussion and Open Problems. One might be interested in pursuing an analog of the preceding results for two-dimensional random walks. Consider, for instance, a Gaussian random walk, where the increments X_i have a bivariate normal distribution with mean zero and covariance matrix $\mathbf{I}$. The projection of this Gaussian random walk onto any fixed line is a Gaussian random walk in one dimension, and by Theorem 13.4 the probability that it has a point of increase tends to zero at the rate $O(1/\log n)$. One may then ask what is the probability that there exists a line such that the Gaussian random walk, projected onto that line, has a point of increase. Unpublished calculations by R. Pemantle seem to indicate that this probability does not converge to zero at a rate faster than $O(1/\log \log n)$. Although it is conjectured that the probability does converge to zero at the rate $O(1/\log \log n)$, it is an open question whether it converges to zero at all. For three-dimensional random walks, the probability that the random walk, projected onto some line, has a point of increase does not converge to zero as $n \to \infty$.

A continuous analog to this question is whether, for Brownian motion in the plane, there exists a line such that the Brownian motion path, projected onto that line, has a global point of increase. An equivalent question is whether the Brownian motion path admits cut lines. (We say a line l is a **cut line** for the Brownian motion if, for some t_0, $B(t)$ lies on one side of l for all $t < t_0$ and on the other side of l for all $t > t_0$.) To see the equivalence, note that l is a cut line if and only if the Brownian motion, projected onto a line perpendicular to l, has a point of increase. This can be seen from Figure 6.2 below, where l_1 is a cut line, and the Brownian path projected onto the perpendicular line l_2 has a global point of increase where it crosses l_1. It was proved by Bass and Burdzy (1997) that Brownian motion almost surely does not has cut lines. It is still unknown whether a Gaussian random walk in the plane will have cut lines.

Burdzy (1989) showed that Brownian motion in the plane almost surely does have **cut points**, which are points $B(t_0)$ such that the Brownian motion path with the point $B(t_0)$ removed is disconnected. It is conjectured that the Hausdorff dimension of the set of cut points is 3/4. For Brownian motion in three dimensions, there almost surely exist cut planes, where we say P is a **cut plane** if for some t_0, $B(t)$ lies on one side of the plane for $t < t_0$ and on the other side for $t > t_0$.

R. Pemantle has shown that a Brownian motion path almost surely does not cover any straight line segment. Which curves can and can not be covered by a Brownian motion path is, in general, an open question. Also unknown is the minimal Hausdorff dimension of curves in a typical Brownian motion path. Burdzy and Lawler (1990) showed this minimal dimension to be at most $3/2 - 1/4\pi^2 \approx 1.47$.

14. Frostman's Lemma, energy, and dimension doubling

In this section, we prove a Lemma due to Frostman (1935) and show how it can be used to prove a converse to the energy theorem. We then show how the energy theorem can be used to deduce a result concerning the Hausdorff dimension of $B(K)$ for closed sets K.

LEMMA 14.1. *If $K \subset \mathbb{R}^d$ is a closed set such that $H^\alpha(K) > 0$, then there exists a nonzero Borel probability measure μ supported on K such that $\mu(D) < C|D|^\alpha$ for all Borel sets D, where C is a constant not depending on D and $|D|$ denotes the diameter of D.*

We will give a proof discovered in the 1980s that is much simpler than Frostman's original proof. It is based on the MaxFlow-MinCut Theorem of Ford and Fulkerson (1956) which we state below.

Consider a tree with vertex set V, directed edge set E, and root r. To each $e \in E$, we assign a bound $b(e)$. A **flow** is a function $f : E \to \mathbb{R}^+$ such that for all $v \in V$ such that v is not the root or a terminal node we have $f(e_v) = \sum_{e \in B_v} f(e)$, where e_v denotes the edge ending at v and B_v denotes the set of vertices emanating from e_v. Thus, the flow into and out of each vertex other than the root is conserved. The **strength** of the flow, denoted $\|f\|$, is defined to be $\sum_{e \in B_r} f(e)$, or the total flow out of the root. We say $\pi \subset E$ is a **cutset** if any path starting from the root that is infinite or ends at a terminal node must traverse an edge in π. In figure 6.3, the edges ending at the black vertices form a cutset, and the numbers beside the edges are possible values for a flow. The MaxFlow-Mincut Theorem, which is proven, for example, in Polya, et. al. (1983), says that

$$\sup_{f \leq b} \|f\| = \inf_{cutsets\ \pi} \sum_{e \in \pi} b(e),$$

so the strength of the strongest flow equals the total size of the minimal cut.

PROOF OF FROSTMAN'S LEMMA. We may assume $K \subset [0,1]^d$. Any cube in $\mathbb{R}^d$ of side length s can be split into 2^d cubes of side length $s/2$. We first create a tree with a root that we associate with the cube $[0,1]^d$. Every vertex in the tree has 2^d edges emanating from it, each leading to a vertex that is associated with one of the 2^d sub-cubes with half the side length of the original cube. We then erase the edges ending in vertices associated with sub-cubes that do not intersect K. For any edge e at level n that remains, define $b(e) = 2^{-n\alpha}$. Note that if $x \in K$, then there is an infinite path emanating from the root, all of whose vertices are associated with cubes that contain x and thus intersect K. Any cutset must contain one of these vertices, which means the cubes associated with the endpoints of the edges in any cutset must must cover K. Since $H^\alpha(K) > 0$, we have

$$\inf \left\{ \sum_j |A_j|^\alpha : K \subset \cup_j A_j \right\} > 0.$$

Thus, by the MaxFlow-MinCut Theorem, there exists a flow f of positive strength that meets the bound.

We now show how to define a suitable measure on the space of infinite paths. Define $\tilde{\mu}(\{\text{ all paths through } e\}) = f(e)$. It is easily checked that the collection $\mathcal{C}$ of sets of the form $\{\text{ all paths through } e\}$ is a semi-algebra. (Recall this means that if $S, T \in \mathcal{C}$, then $S \cap T \in \mathcal{C}$ and S^c is a finite disjoint union of sets in $\mathcal{C}$. Because the flow through any vertex is preserved, $\tilde{\mu}$ is countably additive. Thus, using Theorem A1.3 of Durrett (1996), we can extend $\tilde{\mu}$ to a measure μ on $\sigma(\mathcal{C})$. We can interpret μ as a Borel measure on $[0,1]^d$ satisfying $\mu(C_v) = f(e_v)$, where C_v is the cube associated with the vertex v and e_v is the edge ending at v. Since K is closed, any $x \in K^c$ is in one of the sub-cubes removed during the construction. Hence, μ is supported on K. Suppose D is a Borel subset of $\mathbb{R}^d$. Let n be the integer such that $2^{-n} < |D \cap [0,1]^d| \le 2^{-(n-1)}$. Then, $D \cap [0,1]^d$ can be covered with 3^d of the cubes in the above construction having side length 2^{-n}. Using the bound, we have $\mu(D) \le 3^d 2^{-n\alpha} \le 3^d |D|^\alpha$, so we have a finite measure μ satisfying the conclusion of the Lemma. Renormalizing to get a probability measure completes the proof. ∎

The following result is a converse to the Frostman's Energy Theorem 6.6.

THEOREM 14.2 (Frostman, 1935). *If $K \subset \mathbb{R}^d$ is closed and $\dim_{\mathcal{H}}(K) > \alpha$, then there exists a Borel probability measure μ on K such that*

$$\mathcal{E}_\alpha(\mu) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \frac{d\mu(x) d\mu(y)}{|x-y|^\alpha} < \infty.$$

PROOF. Since $\dim_{\mathcal{H}}(K) > \alpha$, there exists a $\beta > \alpha$ such that $H^\beta(K) > 0$. By Frostman's Lemma, there exists a nonzero Borel probability measure μ on K and constant C such that $\mu(D) \le C|D|^\beta$ for all Borel sets D. By restricting μ to a smaller set if necessary, we can make the support of μ have diameter less than one. Fix $x \in K$, and for $k \ge 1$ let $S_k(x) = \{y : 2^{-k} < |x-y| \le 2^{1-k}\}$. Since μ has no atoms, we have

$$\int_{\mathbb{R}^d} \frac{d\mu(y)}{|x-y|^\alpha} = \sum_{k=1}^\infty \int_{S_k(x)} \frac{d\mu(y)}{|x-y|^\alpha} \le \sum_{k=1}^\infty \mu(S_k(x)) 2^{k\alpha},$$

where the equality follows from the Monotone Convergence Theorem and the inequality holds by the definition of the S_k. Also,

$$\sum_{k=1}^{\infty} \mu(S_k(x)) 2^{k\alpha} \leq C \sum_{k=1}^{\infty} |2^{2-k}|^{\beta} 2^{k\alpha} = C' \sum_{k=1}^{\infty} 2^{k(\alpha-\beta)},$$

where $C' = 2^{2\beta} C$. Since $\beta > \alpha$, we have

$$\mathcal{E}_\alpha(\mu) \leq C' \sum_{k=1}^{\infty} 2^{k(\alpha-\beta)} < \infty,$$

which proves the theorem. ∎

DEFINITION. The α-capacity of a set K, denoted $\mathrm{Cap}_\alpha(K)$, is

$$\left[\inf_{\mu} \mathcal{E}_\alpha(\mu)\right]^{-1},$$

where the infimum is over all Borel probability measures supported on K. If $\mathcal{E}_\alpha(\mu) = \infty$ for all such μ, then we say $\mathrm{Cap}_\alpha(K) = 0$.

THEOREM 14.3 (McKean, 1955). *Let B denote Brownian motion in $\mathbb{R}^d$. Let $A \subset [0, \infty)$ be a closed set such that* $\dim_{\mathcal{H}}(A) \leq d/2$. *Then, almost surely* $\dim_{\mathcal{H}} B(A) = 2\dim_{\mathcal{H}}(A)$.

REMARK. Theorem 14.3 requires A to be fixed. If we allow a random A depending on the Brownian path, then the conclusion still holds if $d \geq 2$. However, for $d = 1$, suppose $A = Z_B = \{t : B_1(t) = 0\}$. We have shown that $\dim_{\mathcal{H}}(Z_B) = 1/2$ almost surely, but $\dim_{\mathcal{H}}(B_1(Z_B)) = \dim_{\mathcal{H}}(\{0\}) = 0$.

PROOF OF THEOREM 14.3. Let $\alpha < \dim_{\mathcal{H}}(A)$. By Theorem 14.2, there exists a Borel probability measure μ on A such that $\mathcal{E}_\alpha(\mu) < \infty$. Denote by μ_B the random measure on $\mathbb{R}^d$ defined by

$$\mu_B(D) = \mu(B_d^{-1}(D)) = \mu(\{t : B_d(t) \in D\})$$

for all Borel sets D. Then

$$\mathbb{E}[\mathcal{E}_{2\alpha}(\mu_B)] = \mathbb{E}\left[\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \frac{d\mu_B(x)d\mu_B(y)}{|x-y|^{2\alpha}}\right] = \mathbb{E}\left[\int_{\mathbb{R}}\int_{\mathbb{R}} \frac{d\mu(t)d\mu(s)}{|B_d(t)-B_d(s)|^{2\alpha}}\right],$$

where the second equality can be verified by a change of variables. Note that the denominator on the right hand side has the same distribution as $|t-s|^\alpha |Z|^{2\alpha}$, where Z is a d-dimensional standard normal random variable. We have

$$\mathbb{E}[|Z|^{-2\alpha}] = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} |y|^{-2\alpha} e^{-|y|^2/2}\, dy < \infty.$$

Hence, using Fubini's Theorem to take the expectation inside the integral, we have

$$\begin{aligned} \mathbb{E}[\mathcal{E}_{2\alpha}(\mu_B)] &= \int_{\mathbb{R}}\int_{\mathbb{R}} \mathbb{E}[|Z|^{-2\alpha}] \frac{d\mu(t)d\mu(s)}{|t-s|^\alpha} \\ &\leq \mathbb{E}[|Z|^{-2\alpha}] \mathcal{E}_\alpha(\mu) < \infty. \end{aligned}$$

Thus, $\mathbb{E}[\mathcal{E}_{2\alpha}(\mu_B)] < \infty$, and μ_B is supported on $B_d(A)$ since μ is supported on A. It follows from the Energy Theorem 6.6 that $\dim_{\mathcal{H}} B_d(A) \geq 2\alpha$. By letting $\alpha \to \dim_{\mathcal{H}}(A)$, we see that $\dim_{\mathcal{H}}(B_d(A)) \geq 2\dim_{\mathcal{H}}(A)$ almost surely.

Using that B_d is almost surely γ-Hölder for all $\gamma < 1/2$, we have shown in Corollary 5.7 that $\dim_{\mathcal{H}}(B_d(A)) \le 2\dim_{\mathcal{H}}(A)$. ■

REMARK. Suppose $2\alpha < d$. Our proof of Theorem 14.3 shows that if $\mathrm{Cap}_\alpha(A) > 0$, then $\mathrm{Cap}_{2\alpha}(B_d(A)) > 0$ almost surely. The converse of this statement is also true, but much harder to prove.

The following are from week 7, (W 3/11, F 3/13, 1998).
Notes taken by Diego Garcia, `diego@stat.berkeley.edu`.

15. The Law of the Iterated Logarithm

THEOREM 15.1 (The Law of the Iterated Logarithm). *For* $\psi(t) = \sqrt{2t\log\log t}$

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

REMARK. By symmetry it follows that

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

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

PROOF. The main idea is to scale by a geometric sequence. We will first prove the upper bound. Fix $\epsilon > 0$ and $q > 1$. Let

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

If the events A_n would happen infinitely often, then the law of the iterated logarithm would be violated. By Theorem 8.6 the maximum of Brownian motion up to a fixed time t has the same distribution as $|B(t)|$. Therefore

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

We can use the tail estimate $\mathbb{P}(Z > x) \le e^{x^2/2}$ to conclude that for large n:

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

which is summable in n. Since $\sum_n \mathbb{P}(A_n) < \infty$, by the Borel-Cantelli Lemma we get that only finitely many of these events occur. For large t write $q^{n-1} \le t < q^n$. We have

$$\frac{B(t)}{\psi(t)} = \frac{B(t)}{\psi(q^n)}\frac{\psi(q^n)}{\psi(t)} \le (1+e)q,$$

since $\psi(t)/t$ is decreasing in t. Thus

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

but since this holds for any $\epsilon > 0$ and $q > 1$ we have proved that $\limsup B(t)/\psi(t) \le 1$.

For the upper bound, fix $q > 1$. In order to use the Borel-Cantelli lemma in the other direction, we need to find a set of *independent* events. Let

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

we will now use the normal tail estimate from a previous section:

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

Using this estimate we get

$$\mathbb{P}(D_n) = \mathbb{P}\left[Z \geq \frac{\psi(q^n - q^{n-})}{\sqrt{q^n - q^{n-1}}}\right] \geq \frac{ce^{-n\log(n\log q)}}{\sqrt{2\log(n\log q)}} > \frac{c'}{n\log n}$$

and therefore $\sum_n \mathbb{P}(D_n) = \infty$. 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 $-B(q^{n-1})$. From the above we get that for infinitely many n:

$$\frac{B(q^n)}{\psi(q^n)} \geq \frac{-2\psi(q^{n-1}) + \psi(q^n - q^{n-1})}{\psi(q^n)} \geq \frac{-2}{\sqrt{q}} + \frac{q^n - q^{n-1}}{q^n} \tag{15.1}$$

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.

Now (15.1) implies that

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

and letting $q \uparrow \infty$ concludes the proof of the upper bound. ■

COROLLARY 15.2. *If $\{\lambda_n\}$ is a sequence of random times (not necessarily stopping times) satisfying $\lambda_n \to \infty$ and $\lambda_{n+1}/\lambda_n \to 1$ a.s., then*

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

Furthermore, if $\lambda_n/n \to a$ a.s., then

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

PROOF. The proof of the upper bound goes through under the above conditions. 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 possibility that λ_n is a sequence of times where the value of Brownian motion is too small. To get around this problem define

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

Note that D_k and Ω_k are independent events. Moreover, by scaling, $\mathbb{P}(\Omega_k)$ is a positive constant (c) depending on q but not on k. Therefore $\mathbb{P}(D_k^*) = c\mathbb{P}(D_k)$, and so the sum of these probabilities is infinite. $\{D_k^*\}$ itself is not an independent sequence, but its even-indexed subsequence is. So by the Borel-Cantelli lemma, we get that for infinitely many k

$$\min_{q^k \le t \le q^{k+1}} B(t) \ge \psi(q^k)\left(1 - \frac{1}{q} - \frac{2}{\sqrt{(q)}}\right) - \sqrt{q^k}$$

Now define $n(k) = \min\{n \;:\; \lambda_n > q^k\}$. Since the ratio of λ's converges to 1, it follows that $\lambda_{n(k)}/q^k \to 1$ a.s. as $k \to \infty$, and so $q^k \le n(k) < q^{k+1}$ for all large k. Thus for infinitely many k:

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

But since $\sqrt{q^k}/\psi(q^k) \to 0$ we conclude that

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

and since the left hand side does not depend on q we arrive at the desired conclusion.

For the last part, note that if $\lambda_n/n \to a$ then $\psi(\lambda_n)/\psi(an) \to 1$ and the above proof goes through. ■

COROLLARY 15.3. *If $\{S_n\}$ is a simple random walk on $\mathbb{Z}$, then a.s.*

$$\limsup_{n\to\infty} \frac{S_n}{\psi(n)} = 1.$$

This immediately follows from the previous corollary by setting:

$$\lambda_0 = 0, \qquad \lambda_n = \min\{t > \lambda_{n-1} : |B(t) - B(\lambda_{n-1})| = 1\}.$$

The waiting times $\{\lambda_n - \lambda_{n-1}\}$ are i.i.d. random variables with mean 1 by Wald's equation. By the law of large numbers λ_n/n will converge to 1, and the corollary follows.

The following are from week 9-10, (W 3/25 - W 4/8, 1998).
Notes taken by Ye Xia, `xia@eecs.berkeley.edu`.

16. Skorokhod's representation

Our goal here is to find a stopping time with finite expected value at which Brownian motion has a given mean 0 distribution. If the distribution is on two points $a < 0 < b$, then this is easy. Define

$$\tau_{a,b} = min\{t : B_t \in \{a, b\}\} \tag{16.1}$$

Let $\tau = \tau_{a,b}$. Then, by Wald's Lemma,

$$0 = \mathbb{E}B_\tau = a\mathbb{P}(B_\tau = a) + b\mathbb{P}(B_\tau = b)$$

Then,

$$\mathbb{P}(B_\tau = a) = \frac{b}{|a| + b}$$

$$\mathbb{P}(B_\tau = b) = \frac{|a|}{|a|+b}$$

By the corollary of Wald's Lemma,

$$\mathbb{E}\tau = \mathbb{E}B_\tau^2 = \frac{a^2 b}{|a|+b} + \frac{b^2 a}{|a|+b} = |a|b$$

Let the distribution of the random variable X_1 be $\mu_{a,b}$, such that,

$$X_1 = \begin{cases} a & \text{with probability } \frac{b}{|a|+b} \\ \\ b & \text{with probability } \frac{|a|}{|a|+b} \end{cases} \tag{16.2}$$

Then, we have

$$B_\tau \stackrel{\mathrm{d}}{=} X_1.$$

THEOREM 16.1 (Skorokhod's Representation, 1965). *Let B be the standard Brownian motion on $\mathbb{R}$.*

(i) If X is a real random variable, then, there exists a stopping time τ, which is finite a.s., such that B_τ has the same distribution as X.

(ii) If $\mathbb{E}X = 0$ and $\mathbb{E}X^2 < \infty$, then τ can be chosen to have finite mean.

Only part (ii) of the theorem is useful.

PROOF.

(i) Pick X according to its distribution. Define $\tau = \min\{t : B(t) = X\}$. It is clear τ is finite a.s.

(ii) Let X have distribution ν on $\mathbb{R}$. We can assume ν has no mass on $\{0\}$, i.e., $\nu(\{0\}) = 0$. For, suppose $\nu(\{0\}) > 0$. Write $\nu = \nu(\{0\})\delta_0 + (1 - \nu(\{0\})\tilde{\nu}$, where the distribution $\tilde{\nu}$ has no mass on $\{0\}$. Let stopping time $\tilde{\tau}$ be the solution of the problem for the distribution $\tilde{\nu}$. The solution for the distribution ν is,

$$\tau = \begin{cases} \tilde{\tau} & \text{with probability } 1 - \nu(\{0\}) \\ 0 & \text{with probability } \nu(\{0\}) \end{cases}$$

Then, $\mathbb{E}\tau = (1 - \nu(\{0\}))\mathbb{E}\tilde{\tau} < \infty$ and $B(\tau)$ has distribution ν. From now on, we assume $\nu(\{0\}) = 0$. It follows that

$$M \stackrel{\text{def}}{=} \int_0^\infty x\, d\nu = -\int_{-\infty}^0 x\, d\nu.$$

Let $\phi : \mathbb{R} \longrightarrow \mathbb{R}$ be a non-negative measurable function. Then

$$\begin{aligned} M\int_{-\infty}^{\infty} \phi(x)\, d\nu &= M\int_0^\infty \phi(y)\, d\nu(y) + M\int_{-\infty}^0 \phi(z)\, d\nu(z) \\ &= \int_{-\infty}^0 (-z)\, d\nu(z) \int_0^\infty \phi(y)\, d\nu(y) + \int_0^\infty y\, d\nu(y) \int_{-\infty}^0 \phi(z)\, d\nu(z) \\ &= \int_{-\infty}^0 \int_0^\infty (y\phi(z) - z\phi(y))\, d\nu(y) d\nu(z) \end{aligned}$$

In the last step, we applied Fubini to the second integral. By the definition of the distribution $\mu_{z,y}$ in (16.2), we can write

$$y\phi(z) - z\phi(y) = (|z| + y) \int_{\{z,y\}} \phi(x)\, d\mu_{z,y}(x)$$

Then,

$$\int_{-\infty}^{\infty} \phi(x)\, d\nu = \frac{1}{M} \int_{-\infty}^{0} \int_{0}^{\infty} \left(\int_{\{z,y\}} \phi(x)\, d\mu_{z,y}(x) \right) (|z| + y)\, d\nu(y) d\nu(z) \tag{16.3}$$

Consider the random variable (Z, Y) on the space $(-\infty, 0) \times (0, \infty)$. Define its distribution by

$$\mathbb{P}((Z, Y) \in A) \stackrel{\text{def}}{=} \frac{1}{M} \int_{A} (|z| + y)\, d\nu(y) d\nu(z) \tag{16.4}$$

for all Borel set A on $(-\infty, 0) \times (0, \infty)$. It is easy to verify that (16.4) defines a probability measure. In particular, let $\phi(x) = 1$, and by (16.3),

$$\frac{1}{M} \int_{-\infty}^{0} \int_{0}^{\infty} (|z| + y)\, d\nu(y) d\nu(z) = 1.$$

Once (Z, Y) is defined, (16.3) can be rewritten as

$$\mathbb{E}\phi(X) = \int_{-\infty}^{\infty} \phi(x)\, d\nu = \mathbb{E}\left[\int_{\{Z,Y\}} \phi\, d\mu_{Z,Y} \right]. \tag{16.5}$$

In the last term above, the expectation is taken with respect to the distribution of (Z, Y). The randomness comes from (Z, Y). When ϕ is any bounded measurable function, apply (16.5) to the positive and negative part of ϕ separately. We conclude that (16.5) holds for any bounded measurable function.

The stopping time τ is defined as follows. Let the random variable (Z, Y) be independent of the Brownian motion B. Now let $\tau = \tau_{Z,Y}$ be as in (16.1). In words, the stopping rule is to first pick the values for Z, Y independent of the Brownian motion, according to the distribution defined by (16.4). Stop when the Brownian motion reaches either Z or Y for the first time. Notice that τ is a stopping time with respect to the Brownian filtration $\mathcal{F}_t = \sigma\{\{B(s)\}_{s \le t}, Z, Y\}$.

Next, we will show $B(\tau) \stackrel{\text{d}}{=} X$. Indeed, for any bounded measurable function ϕ:

$$\begin{aligned} \mathbb{E}\phi(B(\tau)) &= \mathbb{E}[\mathbb{E}[\phi(B(\tau_{Z,Y}))|Z, Y]] \\ &= \mathbb{E}[\int_{\{Z,Y\}} \phi\, d\mu_{Z,Y}] = \mathbb{E}\phi(X). \end{aligned}$$

Here the second equality is due to the definition of $\tau_{Z,Y}$, the third one, to (16.5).

The expectation of τ can be computed similarly:

$$\mathbb{E}\tau = \mathbb{E}[\mathbb{E}[\tau_{Z,Y}|Z, Y]] = \mathbb{E}[\int_{\{Z,Y\}} x^2\, d\mu_{Z,Y}] = \int x^2\, d\nu(x).$$

The second equality follows from the corollary of Wald's Lemma, and the third one, from (16.5), by letting $\phi(x) = x^2$. ∎

16.1. Root's Method. Root (1969) showed for random variable X with $\mathbb{E}X = 0$ and $\mathbb{E}X^2 < \infty$, there exits a closed set $A \subset \mathbb{R}^2$, such that $B(\tau) \stackrel{\mathrm{d}}{=} X$ and $\mathbb{E}\tau = \mathbb{E}X^2$, for $\tau = \min\{t : (t, B(t)) \in A\}$. In words, τ is the first time the Brownian graph hits the set A (see figure 16.1). This result is not useful in practice since the proof is based on a topological existence theorem, and does not provide a construction of the set A.

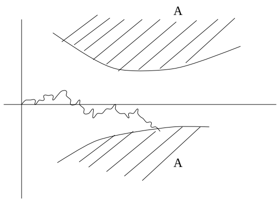

FIGURE 16.1. Root's Approach - the shaded area is the set A

To illustrate the difference between Skorokhod's method and Root's method, let the random variable X take values in $\{-2, -1, 1, 2\}$, each with probability $1/4$. Since this is a very simple case, it is not necessary to go through the procedure shown in the proof of the theorem. The Skorokhod's stopping rule simply says the Brownian motion should stop at the first time $|B(t)| = 1$ or 2. Figure 16.2 illustrates both the Skorokhod's stopping rule and the Root's stopping rule. In Root's, the two dimensional set A consists of four horizontal lines represented by $\{(x, y) : x \geq M, |y| = 1\} \cup \{(x, y) : x \geq 0, |y| = 2\}$, for some $M > 0$. This is intuitively clear by the following argument. Let M approaches 0. The Brownian motion takes value of 1 or -1, each with probability $1/2$, at the first time the Brownian graph hits the set A. Let M approaches ∞. The Brownian motion takes value of 2 or -2, each with probability $1/2$, at the first time the Brownian graph hits the set A. Since the probability assignment is a continuous function of M, by the intermediate value theorem, there exists an $M > 0$ such that

$$\mathbb{P}(B(\tau) = 2) = \mathbb{P}(B(\tau) = 1) = \mathbb{P}(B(\tau) = -1) = \mathbb{P}(B(\tau) = -2) = 1/4$$

However, it is difficult to compute M explicitly. The third graph in figure 16.2 shows an alternative way to define the two-dimensional set A.

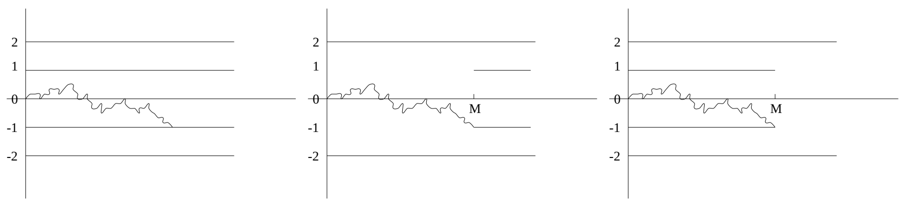

FIGURE 16.2. Comparison of Skorokhod's and Root's Approach - From left to right: Skorokhod's; Root's I; and Root's II

16.2. Dubins' Stopping Rule. Skorokhod's stopping rule depends on random variables (i.e., Z, Y in the proof of the theorem) independent of the Brownian motion. Since the Brownian motion contains a lot of randomness, it seems possible not to introduce the extra randomness. Dubins (1968) developed a method for finding the stopping time following this idea. We use the same X above as an example. First, run the Brownian motion until $|B(t)| = 3/2$. Then, stop when it hits one of the original four lines. Figure 16.3 gives the graphical demonstration of this procedure. To generalize it to the discrete case, let X have discrete distribution ν. Suppose $\nu(\{0\}) = 0$. First, find the centers of mass for the positive and negative part of the distribution separately. For example, for the positive part, the center of mass is,

$$\frac{\int_0^\infty x \, d\nu}{\nu([0,\infty])}$$

Run the Brownian motion till it reaches one of the centers of mass, either positive or negative. Then shift the distribution so that the center of mass is at 0. Normalize the distribution (the positive or negative part corresponding to the center of mass). Then repeat the procedure until exactly one line lies above the center of mass and another one lies below it, or until the center of mass overlaps with the last line left. Stop the Brownian motion when it hits one of these two lines in the former case, or when it hits the last center of mass.

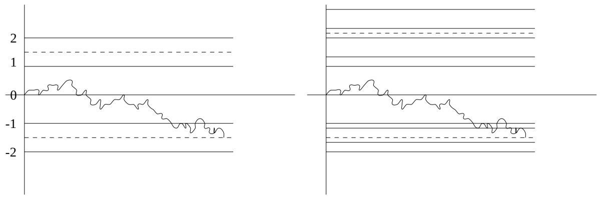

FIGURE 16.3. Dubins' Approach - From left to right: random variable X; general discrete random variable. The dotted lines are the centers of mass for the positive or negative part.

In the case where X has a continuous distribution, it needs to be approximated by discrete distributions. See Dudley (1989) for details.

16.3. Azéma-Yor's Stopping Rule. See chapter VI section 5 of Revuz and Yor (1994) for reference.

16.4. Skorokhod's representation for a sequence of random variables. Let $\{X_i\}_{i\geq 1}$ be independent random variables with mean 0 and finite variances. Let τ_1 be a stopping time with $\mathbb{E}\tau_1 = \mathbb{E}X_1^2$ and $B(\tau_1) \stackrel{\mathrm{d}}{=} X_1$. $\{B(\tau_1+t) - B(\tau_1)\}_{t\geq 0}$ is again a Brownian motion. Then, we can find a stopping time τ_2 with $\mathbb{E}\tau_2 = \mathbb{E}X_2^2$, and $B(\tau_1+\tau_2) - B(\tau_1) \stackrel{\mathrm{d}}{=} X_2$ and is independent of $\mathcal{F}_{\tau_1}$. Repeat the procedure for $\tau_3, \tau_4 \cdots$, etc. Define $T_1 = \tau_1$, and $T_n = \tau_1 + \tau_2 + \cdots + \tau_n$. Then, $B(T_k + \tau_{k+1}) - B(T_k) \stackrel{\mathrm{d}}{=} X_k$ and is independent of $\mathcal{F}_{T_k}$. We get,

$$\begin{aligned} B(T_n) &\stackrel{\mathrm{d}}{=} X_1 + X_2 + \cdots + X_n \\ \mathbb{E}T_n &= \sum_{i=1}^{n} \mathbb{E}\tau_i = \sum_{i=1}^{n} \mathbb{E}X_i^2 \end{aligned}$$

This is a very useful formulation. For example, if $\{X_i\}_{i\geq 1}$ is an i.i.d. sequence of random variables, then $\{\tau_i\}_{i\geq 1}$ is also i.i.d. By the Strong Law of Large Numbers, $\frac{T_n}{n} \longrightarrow \mathbb{E}\tau_1 = \mathbb{E}X_1^2$ almost surely, as $n \longrightarrow \infty$. Let $S_n = \sum_{i=1}^{n} \mathbb{E}X_i = B(T_n)$. By the Corollary 15.2 of the Law of Iterated Logarithm (LIL) for the Brownian motion, where the sequence of random times T_n has the property $\frac{T_n}{n} \longrightarrow \alpha$ as $n \longrightarrow \infty$, for some constant $\alpha, 0 < \alpha < \infty$, we have,

$$\limsup_{n\to\infty} \frac{S_n}{\sqrt{2n \log\log n}\sqrt{\mathbb{E}X_1^2}} = 1$$

This was first proved by Strassen (1964).

17. Donsker's Invariance Principle

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 $\{S_t\}_{t\geq 0}$ (Figure 17.1).

THEOREM 17.1 (Donsker's Invariance Principle). *As* $n \longrightarrow \infty$

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

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* $n \longrightarrow \infty$

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

REMARK. The proof of the theorem shows it suffices that ψ is bounded and continuous at almost all Brownian paths.

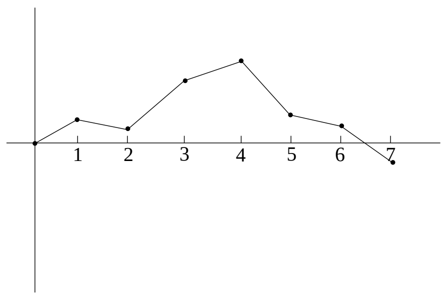

FIGURE 17.1. Derivation of S_t by linear interpolation of the random walk S_n

17.1. Applications of Donsker's Theorem.

EXAMPLE 17.2. As $n \longrightarrow \infty$

$$\frac{\max_{1\le k\le n} S_k}{\sqrt{n}} \overset{\text{in law}}{\Longrightarrow} \max_{0\le t\le 1} B(t)$$

i.e., for any constant a

$$\mathbb{P}(\max_{1\le k\le n} S_k \ge a\sqrt{n}) \longrightarrow \frac{2}{\sqrt{2\pi}}\int_a^\infty e^{-u^2/2}\,du,$$

because by Theorem 8.6

$$\mathbb{P}(\max_{0\le t\le 1} B(t) \ge a) = 2\mathbb{P}(B(1) \ge a).$$

To prove this, let $\phi : \mathbb{R} \longrightarrow \mathbb{R}$ be a bounded continuous function. Take the function $\psi(f) = \phi(\max_{[0,1]} f)$. Then ψ is a bounded and continuous function on $\tilde{C}[0,1]$. By the construction of $\{S_t\}_{t\ge 0}$, we have

$$\mathbb{E}\psi(\{\frac{S_{tn}}{\sqrt{n}}\}_{0\le t\le 1}) = \mathbb{E}\phi(\max_{0\le t\le 1}\{\frac{S_{tn}}{\sqrt{n}}\}) = \mathbb{E}\phi(\frac{\max_{1\le k\le n} S_k}{\sqrt{n}}).$$

Also,

$$\mathbb{E}\psi(\{B(t)\}_{0\le t\le 1}) = \mathbb{E}\phi(\max_{0\le t\le 1} B(t)).$$

Then, by Donsker's Theorem,

$$\mathbb{E}\phi(\frac{\max_{1\le k\le n} S_k}{\sqrt{n}}) \longrightarrow \mathbb{E}\phi(\max_{0\le t\le 1} B(t)).$$

EXAMPLE 17.3. As $n \longrightarrow \infty$,

$$\frac{1}{n}\max\{1 \le k \le n : S_k S_{k-1} \le 0\} \overset{\text{in law}}{\Longrightarrow} \max\{0 \le t \le 1 | B(t) = 0\} \tag{17.1}$$

The left hand side is the last time between 1 to n, scaled by n, that the random walk crosses 0. The probability on the right hand can be explicitly calculated.

To prove (17.1), define the function ψ by $\psi(f) = \max\{t \le 1 : f(t) = 0\}$. ψ is not a continuous function on $\tilde{C}[0,1]$. But it is continuous at every $f \in \tilde{C}[0,1]$ with the property that

$$f\left(\psi(f) - \varepsilon, \psi(f) + \varepsilon\right)$$

contains a neighborhood of 0 for every $\varepsilon > 0$. To elaborate this, suppose $f(t) > 0$ for $\psi(f) < t \le 1$. For any given $\varepsilon > 0$, let $\delta_0 = \min_{[\psi(f)+\varepsilon,1]} f(t)$. Choose $\delta_1 > 0$ so that $(-\delta_1, \delta_1) \subseteq f(\psi(f)-\varepsilon, \psi(f)+\varepsilon)$. Choose a positive $\delta < \min\{\delta_0, \delta_1\}$. Then, $\psi(f-\delta)-\psi(f) < \varepsilon$, and $\psi(f) - \psi(f+\delta) < \varepsilon$ (Figure 17.2). Let $\tilde{f} \in \tilde{C}[0,1]$ such that $||\tilde{f} - f||_\infty < \delta$. Then, for every t, $f(t) - \delta \le \tilde{f}(t) \le f(t) + \delta$. Hence, $|\psi(\tilde{f}) - \psi(f)| < \varepsilon$. That is, ψ is continuous at f. Since the last zero of a Brownian path on $[0,1]$ almost surely is strictly less than 1, and is an accumulation point of zeroes from the left, the Brownian path almost surely has the property that f has. Hence, ψ is continuous at almost all Brownian paths.

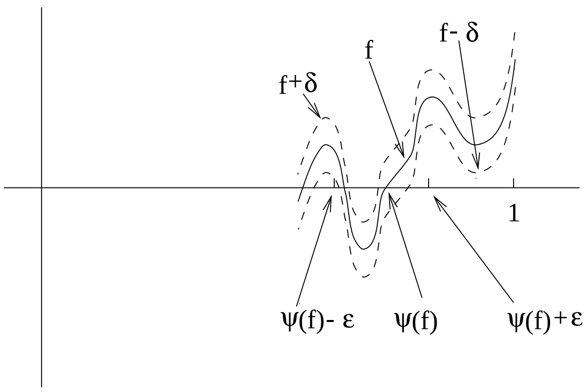

FIGURE 17.2. Illustration that shows ψ is continuous at almost all Brownian paths

The following are from week 11, (W 4/8, F 4/10, 1998).
Notes taken by Yoram Gat, `yoram@stat.berkeley.edu`.

PROOF OF THEOREM 17.1. Let

$$F_n(t) = \frac{S_{tn}}{\sqrt{n}}, \ 0 \le t \le 1.$$

By Skorokhod embedding, there exist stopping times $T_k,\ k = 1, 2, \ldots$ for some standard Brownian motion B such that $S(k) = B(T_k)$. Define $W_n(t) = \frac{B(nt)}{\sqrt{n}}$. Note that W_n is also a standard Brownian motion. We will show that as $n \to \infty$:

$$\mathbb{P}(\sup_{0 \le t \le 1} |F_n - W_n| > \epsilon) \to 0. \tag{17.2}$$

The theorem will follow since by (17.2) if $\Psi : C(0,1) \to \mathbb{R}$ is bounded by M and is continuous on almost every Brownian motion path, then

$$\mathbb{P}(\exists f : ||W - f|| < \epsilon, |\Psi(W) - \Psi(f)| > \delta) \tag{17.3}$$

converges to 0 as $\epsilon \to 0$. Now

$$|\mathbb{E}\Psi(F_n) - \mathbb{E}\Psi(W_n)| \le \mathbb{E}|\Psi(F_n) - \Psi(W_n)|$$

and the right hand side is bounded above by

$$\begin{aligned} &2M\mathbb{P}\left(||W_n - F_n|| \ge \epsilon\right) \ + \\ &2M\mathbb{P}\left(||W_n - F_n|| < \epsilon\right), |\Psi(W_n) - \Psi(F_n)| > \delta) \ + \ \delta. \end{aligned}$$

The second term is bounded by (17.3), so by setting δ small, then setting ϵ small and then setting n large, the three terms of the last expression may be made arbitrarily small.

To prove (17.2), let A_n be the event that $|F_n(t) - W_n(t)| > \epsilon$ for some t. We will show that $\mathbb{P}(A_n) \to 0$.

Let $k = k(t)$ designate the integer such that $\frac{k-1}{n} \le t < \frac{k}{n}$. Then, since $F_n(t)$ is linearly interpolated between $F_n(\frac{k-1}{n})$ and $F_n(\frac{k}{n})$,

$$A_n \subset \{\exists t : \left|\frac{S(k)}{\sqrt{n}} - W_n(t)\right| > \epsilon\} \cup \{\exists t : \left|\frac{S(k-1)}{\sqrt{n}} - W_n(t)\right| > \epsilon\}.$$

Writing $S(k) = B(T_k) = \sqrt{n}W_n(T_k/n)$, we get

$$A_n \subset \{\exists t : \left|W_n(\frac{T_k}{n}) - W_n(t)\right| > \epsilon\} \cup \{\exists t : \left|W_n(\frac{T_{k-1}}{n}) - W_n(t)\right| > \epsilon\}.$$

Given $\delta \in (0,1)$, the event on the right implies that either

$$\{\exists t : |T_k/n - t| \vee |T_{k-1}/n - t| \geq \delta n\} \tag{17.4}$$

or

$$\{\exists s, t \in [0,2] : |s-t| < \delta, |W_n(s) - W_n(t)| > \epsilon\}.$$

Since each W_n is uniformly continuous on $[0,2]$, by choosing δ small and then n large, the probability of the later event can be made arbitrarily small.

To conclude the proof, all we have to show is that for each δ, the probability of (17.4) converges to 0. In fact, this event only happens for finitely many n a.s. Since we chose k so that t is in the interval $[(k-1)/n, k/n]$, the absolute differences in (17.4) are bounded above by $[T_k/n - (k-1)/n] \vee [k/n - T_{k-1}/n]$ for t fixed. This implies that (17.4) is a subset of the union of the events

$$\{\sup_{0 \leq k \leq n} \frac{|T_k - k + c|}{n} > \delta\} \tag{17.5}$$

for $c = 1$ and $c = -1$. Note the deterministic fact that if a real sequence $\{a_n\}$ satisfies $\lim a_n/n \to 1$, then $\sup_{0 \leq k \leq n} |a_k - k|/n \to 0$. Since T_n is a sum of i.i.d. mean 1 random variables, the Law of Large Numbers enables us to apply this to $a_n = T_n + c$, and conclude that (17.5) happens only finitely many times, as desired. ∎

18. Harmonic functions and Brownian motion in $\mathbb{R}^d$

Definition 18.1. Let $D \subset \mathbb{R}^d$ be a domain (a connected open set). $u : D \to \mathbb{R}$ is *harmonic* if it is measurable, locally bounded (i.e. bounded on closed balls in D) and for any ball $B = B(x,r) \subset D$,

$$u(x) = \frac{1}{\mathcal{L}_d(B)} \int_B u(y)dy.$$

If u is harmonic then it is continuous in D: If $x_n \to x$ then

$$\mathbf{1}(B(x_n, r))(y) \xrightarrow[n\to\infty]{a.e.} \mathbf{1}(B(x,r))(y),$$

thus, by the Dominated Convergence Theorem, $u(x_n) \to u(x)$.

Theorem 18.2. *u is harmonic in D iff u is continuous in D and*

$$u(x) = \frac{1}{\sigma_{d-1}(S(x,r))} \int_{S(x,r)} u(y) d\sigma_{d-1}(y), \tag{18.1}$$

where $S(x,r) = \{y : |y-x| = r\}$, and σ_{d-1} is the $d-1$ dimensional Hausdorff measure.

Proof. Define:

$$\int_{S(x,r)} u(y) d\sigma_{d-1}(y) = \Psi(r) r^{d-1}.$$

Then,

$$R^d \mathcal{L}(B(x,1))u(x) = \int_{B(x,r)} u(y)dy = \int_0^R \Psi(r)r^{d-1}dr.$$

Differentiate w.r.t. R to obtain:

$$d\mathcal{L}(B(x,1))u(x) = \Psi(R).$$

$\Rightarrow \Psi(R) = \text{const}$. By continuity as $R \to 0$, (18.1) holds. ∎

Equivalent definition: u is harmonic if u is continuous, twice differentiable, and $\Delta u = \sum_i \frac{\partial^2 u}{(\partial x_i)^2} = 0$.

DEFINITION 18.3.

$$G(x,y) = \int_0^\infty p(x,y,t)dt, \;\; x,y \in \mathbb{R}^d$$

is the Green function in $\mathbb{R}^d$, where $p(x,y,t)$ is the Brownian transition kernel, $p(x,y,t) = (2\pi t)^{d/2} \exp -\frac{|x-y|^2}{2t}$.

$G(x,y)$ is finite iff $x \neq y$ and $d > 2$.

THEOREM 18.4. *$x \mapsto G(x,0)$ is harmonic on $\mathbb{R}^d \setminus \{0\}$.*

PROOF. Define $F_r(x) = \int_{B(x,r)} G(x,z)dz$. We prove that for all ϵ small enough, $F_\epsilon(x)$ is harmonic, i.e.,

$$F_\epsilon(x) = \int_{B(x,r)} F_\epsilon(y)dy. \tag{18.2}$$

The theorem follows by dividing both sides of (18.2) by $\mathcal{L}(B(0,\epsilon))$ and taking the limit as $\epsilon \to 0$.

By Fubini's Theorem and the definition of the Markov kernel p:

$$F_\epsilon(x) = \int_0^\infty \int_{B(x,\epsilon)} p(x,z,t)dzdt = \int_0^\infty \mathbb{P}_x(W(t) \in B(0,\epsilon))dt$$

Applying Fubini another time,

$$F_\epsilon(x) = \mathbb{E}_x \int_0^\infty \mathbf{1}(W_t \in B(0,\epsilon))dt. \tag{18.3}$$

Denote by ν_{d-1} the uniform probability measure on the unit sphere in $\mathbb{R}^d$ (i.e. $\nu_{d-1} = \sigma_{d-1}/||\sigma_{d-1}||$). Fix x, let $r < x$ and let $\tau = \min\{t : |W(t) - x| = r\}$. Let $\epsilon < x - r$. Since W spends no time in $B(0,\epsilon)$ before time τ, we can write $F_\epsilon(x)$ as

$$\mathbb{E}_x \int_\tau^\infty \mathbf{1}(W_t \in B(0,\epsilon))dt = \mathbb{E}_x \mathbb{E}_x \left[\int_\tau^\infty \mathbf{1}(W_t \in B(0,\epsilon))dt | W_\tau \right]$$

By the Markov property and since W_τ is uniform on the sphere of radius r about x by rotational symmetry, we conclude:

$$F_\epsilon(x) = \mathbb{E}_x F_\epsilon(W_\tau) = \int_{S(0,1)} F_\epsilon(x+ry)d\nu_{d-1}(y). \quad ∎$$

$G(x,0)$ can be presented as an explicit formula:

$$G(x,0) = \int_0^\infty (2\pi t)^{-d/2} \exp -\frac{|x|^2}{2t} dt.$$

Substituting $s = \frac{|x|^2}{2t}$, we obtain:

$$G(x,0) = \int_0^\infty (\frac{\pi |x|^2}{s})^{-d/2} e^{-s} \frac{|x|^2}{2s^2} ds.$$

Integrating:

$$G(x,0) = |x|^{2-d} \frac{\pi^{-d/2}}{2} \int_0^\infty e^{-s} s^{\frac{d}{2}-2} ds.$$

(The integral is known as $\Gamma(\frac{d}{2}-1)$.) Thus we obtained:

$$G(x,0) = c_d |x|^{2-d}.$$

We have therefore proved that $x \mapsto \frac{1}{|x|^{d-2}}$ is harmonic in $\mathbb{R}^d \setminus \{0\}$, $d \geq 3$.

For $d \geq 3$, the time Brownian motion spends in any ball around the origin has expectation $F_R(0)$ by (18.3). By the definition of $F_R(0)$, this expectation can be written as

$$\int_{B(0,R)} G(0,x)dx = c_d \int_{B(0,R)} |x|^{2-d} dx = \tilde{c}_d \int_0^R r^{d-1} r^{2-d} dr = c'_d R^2,$$

in particular, it is finite.

We now wish to show that Brownian motion in $\mathbb{R}^d$, $d \geq 3$ is transient. Define

$$h_r(x) = \mathbb{P}_x(\exists t \geq 0 : W_t \in B(0,r)).$$

Recalling the definition of $F_r(x)$

$$F_r(x) = \int_{B(0,r)} G(x,z)dz = \int_{B(0,r)} G(x-z,0)dz \tag{18.4}$$

note that it depends only on $|x|$, and define $\tilde{F}_r(|x|) = F_r(x)$.

Suppose $|x| > r$. Since $F_r(x)$ is the expected time spent in $B(0,r)$ starting from x, it must equal the probability of hitting $S(0,r)$ starting from x times the expected time spent in $B(0,r)$ starting from the hitting point of this sphere. Therefore $F_r(x) = h_r(x)\tilde{F}_r(r)$.

Since G is harmonic, from (18.4) we have

$$F_r(x) = \mathcal{L}(B(0,r))G(x,0) = \mathcal{L}(B(0,r))c_d|x|^{2-d}.$$

Therefore: $h_r(x) = \left(\frac{r}{|x|}\right)^{d-2}$.

To conclude, we use the fact that $\limsup_{t\to\infty} |W(t)| = \infty$ a.s. Therefore, for any r, R,

$$\begin{aligned} &\mathbb{P}(\text{BM visits } B(0,r) \text{ for arbitrarily large t}) \\ &\quad \leq \mathbb{P}(\text{BM visits } B(0,r) \text{ after hitting } S(0,R)) \\ &\quad = \left(\frac{r}{R}\right)^{d-2}. \end{aligned}$$

We are now also able to calculate the probability that a Brownian motion starting between two spheres will hit the smaller one before hitting the larger one.

PROPOSITION 18.5. *Define*

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

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

$$a = \frac{\left(\frac{r}{|x|}\right)^{d-2} - \left(\frac{r}{R}\right)^{d-2}}{1 - \left(\frac{r}{R}\right)^{d-2}}. \tag{18.5}$$

PROOF.

$$\begin{aligned} a &= \mathbb{P}_x(\text{BM hits } S(0,r)) - \mathbb{P}_x(\text{BM hits } S(0,R) \text{ first and then hits } S(0,r)) \\ &= \left(\frac{r}{|x|}\right)^{d-2} - (1-a)\left(\frac{r}{R}\right)^{d-2}. \end{aligned}$$

Solving we get (18.5). ∎

Since a is fixed under scaling, the visits of Brownian motion to $S(0,e^k),\ k \in \mathbb{Z}$ form a discrete random walk with constant probability to move up (k increasing) or down (k decreasing). The probability to move down is

$$\frac{e^{2-d} - e^{4-2d}}{1 - e^{4-2d}}.$$

This probability is less than half by the convexity of $x \mapsto e^{-x}$.

The following are from week 12, (W 4/15, F 4/17, 1998).
Notes taken by Xiaowen Zhou, `xiaowen@stat.berkeley.edu`.

19. Maximum principle for harmonic functions

PROPOSITION 19.1 (Maximum Principle). *Suppose that u is harmonic in $D \subset \mathbb{R}$ where D is a connected open set.*

(i) If u attains its maximum in D, then u is a constant.

(ii) If u is continuous on $\bar{D}$ and D is bounded, then $\max_{\bar{D}} u = \max_{\partial D} u$.

(iii) Assume that D is bounded, u_1 and u_2 are two functions harmonic on D and continuous on $\bar{D}$. If u_1 and u_2 take the same values on ∂D, then they are identical on D.

PROOF.

(i) Set $M = \sup_D u$. Note that $V = \{x \in D : u(x) = M\}$ is relatively closed in D. Since D is open, for any $x \in V$, there is a ball $B(x,r) \subset D$. By the mean-value property of u,

$$u(x) = \frac{1}{\mathcal{L}(B(x,r))} \int_{B(x,r)} u(y)dy \le M$$

Equality holds iff $u(y) = M$ almost everywhere on $B(x,r)$, or, by continuity, $B(x,r) \subset V$ This means that V is also open. This means $V = D$ because D is connected, so u is constant on D.

(ii) $\sup_{\bar{D}} u$ is attained on $\bar{D}$ since u is continuous and $\bar{D}$ is closed and bounded. The conclusion is straightforward from (i).

(iii) Consider $u_1 - u_2$. It follows from (ii) that

$$\sup_{\bar{D}}(u_1 - u_2) = \sup_{\partial D}(u_1 - u_2) = 0.$$

Similarly $\sup_{\bar{D}}(u_2 - u_1) = 0$. So $u_1 = u_2$ on $\bar{D}$. ■

COROLLARY 19.2. *Suppose that u is a radial harmonic function in $D \equiv \{r < |x| < R\} \subset \mathbb{R}^d$ ("radial" means $u(x) = \tilde{u}(|x|)$ for some function $\tilde{u}$ and all x) and u is continuous on $\bar{D}$.*

If $d \geq 3$, there exist constants a and b such that $u(x) = a + b|x|^{2-d}$.

If $d = 2$, there exist constants a and b such that $u(x) = a + b\log|x|$.

PROOF. For $d \geq 3$, choose a and b such that

$$a + br^{2-d} = \tilde{u}(r),$$

and

$$a + bR^{2-d} = \tilde{u}(R).$$

Notice that harmonic function $u(x) = \tilde{u}(|x|)$ and the harmonic function $x \to a + b|x|^{2-d}$ agree on ∂D. They also agree on D by Proposition 19.1. So $u(x) = a + b|x|^{2-d}$. By similar consideration we can show that $u(x) = a + b\log|x|$ in the case $d = 2$. ■

20. The Dirichlet problem

DEFINITION. Let $D \subset \mathbb{R}^d$ be a domain. We say that D satisfies the **Picare Cone Condition** if there is a cone $C_x(\alpha, h)$ of fixed height $h(x)$ and angle $\alpha(x)$ in D^c based at each point of ∂D.

PROPOSITION 20.1 (Dirichlet Problem). *Suppose $D \subset \mathbb{R}^d$ is a bounded domain with boundary ∂D, and f is a continuous function on ∂D. Then there is a unique function u which is harmonic on D and is an "extension" of f in the sense that*

$$\lim_{x \to a, x \in D} u(x) = f(a)$$

for each $a \in \partial D$. If D satisfies the Picare Cone Condition, then u is continuous on $\bar{D}$.

PROOF. Let W be a Brownian motion in $\mathbb{R}^d$. Define

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

for any Borel set $A \subset \mathbb{R}^d$. For a ball $B(x,r) \subset D$, strong Markov property implies that

$$u(x) = \mathbb{E}_x[\mathbb{E}_x[f(W_{\tau_D})|\mathcal{F}_{\tau_{B(x,r)}}]] = \mathbb{E}_x[u(W_{\tau_{B(x,r)}})] = \int_{S(x,r)} u(x+y)\mu_r(dy),$$

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.

Now suppose that the Picare Cone Condition holds. 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

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

For any positive integer k and $h' > 0$, using scaling and strong Markov property we have

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

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

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

$$|u(x) - u(z)| = |\mathbb{E}_x f(W_{\tau_D}) - f(z)| \le \mathbb{E}_x |f(W_{\tau_D}) - f(z)| \tag{20.1}$$

If the Brownian motion hits the cone $C_z(\alpha, \delta)$, which is outside the domain D, before the sphere $S(z, \delta)$, then $|z - W\tau_D| < \delta$, and $f(W_{\tau_D})$ is close to $f(z)$. The complement has small probability. More precisely, (20.1) is bounded above by

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

Hence u is continuous on $\bar{D}$. ∎

21. Polar points and recurrence

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

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

$|x| = 1$ implies $a = 0, |x| = R$ 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 $0 < r < R$,

$$\mathbb{P}_x\{\tau_{S(0,R)} < \tau_{S(0,r)}\} = \frac{\log \frac{|x|}{r}}{\log \frac{R}{r}}. \tag{21.1}$$

Notice that that above probability is $\frac{1}{2}$ if $\frac{R}{|x|} = \frac{|x|}{r}$.

Definition 21.1. A set A is polar for a Markov process X if for all x

$$\mathbb{P}_x\{X_t \in A \text{ for some } t > 0\} = 0$$

Define the image of Brownian motion as the random set

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

Proposition 21.2. *Points are polar for a planar Brownian motion, that is for all $z \in \mathbb{R}^2$ we have $\mathbb{P}_0\{z \in BM(0, \infty)\} = 0$.*

Proof. Take $z \ne 0$ 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

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

Given $\delta > 0$, by Markov property and state homogeneity of Brownian motion,

$$\mathbb{P}_0\{0 \in BM(\delta,\infty)\} = \mathbb{E}_0[\mathbb{P}_{X_\delta}\{0 \in BM(0,\infty)\}] = 0.$$

Let $\delta \to 0+$, we have

$$\mathbb{P}_0\{0 \in BM(0,\infty)\} = 0.$$

Hence any fixed single point is a polar set for a planar Brownian motion. ∎

The expected area $\mathbb{E}_0[\mathcal{L}(BM(0,\infty))]$ of planar Brownian motion can be written as

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

where the first equality is by Fubini, the second from the previous theorem. So almost surely, planar Brownian motion lives on a set with zero Lebesgue measure.

PROPOSITION 21.3. *Planar Brownian motion is neighborhood recurrent. i.e.*

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

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. ∎

Question: Let W be a Brownian motion in $\mathbb{R}^3$. Is there an infinite cylinder avoided by W? Or equivalently, what is the value of

$$\mathbb{P}_0\{\text{All orthogonal projections of } W \text{ are neighborhood recurrent}\}?$$

The following are from week 13-14, (W 4/22, F 4/24, W 4/29, F 5/1, 1998).
Notes taken by Bálint Virág, `balint@stat.berkeley.edu`.

22. Capacity and harmonic functions

In this section we will characterize the sets that BM hits, and give bounds on the hitting probabilities in terms of capacity.

The central question of this section is the following: which sets $\Lambda \subset \mathbb{R}^d$ does Brownian motion hit with positive probability? We will show how this is related to the question: for which $\Lambda \subset \mathbb{R}^d$ are there bounded harmonic functions on $\mathbb{R}^d \setminus \Lambda$?

Consider the simplest case first. When Λ is the empty set, the answer to the first question is trivial, whereas the answer to the second one is provided by Liuville's theorem. We will give a probabilistic proof of this theorem.

THEOREM 22.1. *For $d \geq 1$ any bounded harmonic function on $\mathbb{R}^d$ is constant.*

PROOF. Let $u : \mathbb{R}^d \to [-M, M]$ be a harmonic function, x, y two distinct points in $\mathbb{R}^d$, and H the hyperplane so that the reflection in H takes x to y.

Let W_t be Brownian motion started at x, and $\overline{W}_t$ its reflection in H. Let $\tau_H = \min\{t : W(t) \in H\}$. Note that

$$\{W_t\}_{t \geq \tau_H} \stackrel{\mathrm{d}}{=} \{\overline{W}_t\}_{t \geq \tau_H}. \tag{22.1}$$

Harmonicity implies that

$$\mathbb{E}u(W_t) = \mathbb{E}(\mathbb{E}(u(W_t)|\ |W(t) - x|)) = \mathbb{E}(u(x)) = u(x)$$

since the conditional expectation above is just the average u on a sphere about x of radius $|W(t) - x|$. Decomposing the above into $t < \tau_H$ and $t \geq \tau_H$ we get

$$u(x) = \mathbb{E}_x u(W_t \mathbf{1}(t < \tau_H)) + \mathbb{E}_x u(W_t \mathbf{1}(t \geq \tau_H)).$$

A similar equality holds for $u(y)$. Now using (22.1):

$$|u(x) - u(y)| = |\mathbb{E}u(W_t \mathbf{1}(t < \tau_H)) - \mathbb{E}u(\overline{W}_t \mathbf{1}(t < \tau_H))| \leq 2M\mathbb{P}(t < \tau_H) \to 0$$

as $t \to \infty$. Thus $u(x) = u(y)$, and since x and y were chosen arbitrarily, u must be constant. ■

A stronger result is also true.

THEOREM 22.2. *For $d \geq 1$, any positive harmonic function on $\mathbb{R}^d$ is constant.*

PROOF. Let $x, y \in \mathbb{R}^d$, $a = |x - y|$. Suppose u is a positive harmonic function. Then $u(x)$ can be written as

$$\frac{1}{\mathcal{L}_d B_R(x)} \int_{B_R(x)} u(z)dz \leq \frac{\mathcal{L}_d B_{R+a}(y)}{\mathcal{L}_d B_R(x)} \frac{1}{\mathcal{L}_d B_{R+a}(y)} \int_{B_{R+a}(y)} u(z)dz = \frac{(R+a)^d}{R^d} u(y)$$

This converges to $u(y)$ as $R \to \infty$, so $u(x) \leq u(y)$, and by symmetry, $u(x) = u(y)$ for all x, y. Hence u is constant. ■

Nevanlinna (about 1920) proved that for $d \geq 3$ there exist non-constant bounded harmonic functions on $\mathbb{R}^d \setminus \Lambda$ iff $\mathrm{Cap}_G(\Lambda) > 0$. Here G denotes the Green function $G(x, y) = c|x - y|^{2-d}$. It was proved later that $\dim \Lambda > d - 2$ implies existence of such functions, and $\dim \Lambda < d - 2$ implies nonexistence. Kakutani (1944) showed that there exist such functions iff $\mathbb{P}(\text{BM hits } \Lambda) > 0$. Note that the Green function is translation invariant, while the hitting probability of a set is invariant under scaling. It is therefore better to estimate hitting probabilities by a capacity function with respect to a scale-invariant modification of the Green kernel, called the Martin kernel:

$$K(x, y) = \frac{G(x, y)}{G(0, y)} = \frac{|y|^{d-2}}{|x - y|^{d-2}}$$

for $x \neq y$ in $\mathbb{R}^d$, and $K(x, x) = \infty$. The following theorem shows that Martin capacity is indeed a good estimate of the hitting probability.

THEOREM 22.3 (Benjamini, Pemantle, Peres 1995). *Let Λ be any closed set in $\mathbb{R}^d$, $d \geq 3$. Then*

$$\frac{1}{2}\operatorname{Cap}_K \Lambda \leq \mathbb{P}(\exists t > 0 \ : W(t) \in \Lambda) \leq \operatorname{Cap}_K(\Lambda) \tag{22.2}$$

Here

$$\operatorname{Cap}_K(\Lambda) = \left[\inf_{\mu(\Lambda)=1} \int_\Lambda \int_\Lambda K(x,y)d\mu(x)d\mu(y)\right]^{-1}.$$

PROOF. To bound the probability of ever hitting Λ from above, consider the stopping time $\tau = \min\{t \ : \ W_t \in \Lambda\}$. The distribution of W_τ on the event $\tau < \infty$ is a possibly defective distribution ν satisfying

$$\nu(\Lambda) = \mathbb{P}(\tau < \infty) = \mathbb{P}(\exists t > 0 \ : W(t) \in \Lambda). \tag{22.3}$$

Now recall the standard formula, valid when $0 < \epsilon < |y|$:

$$\mathbb{P}(\exists t > 0 : |W_t - y| < \epsilon) = \frac{\epsilon^{d-2}}{|y|^{d-2}}. \tag{22.4}$$

By a first entrance decomposition, the probability in (22.4) is at least

$$\mathbb{P}(|W_\tau - y| > \epsilon \text{ and } \exists t > \tau : |W_t - y| < \epsilon) = \int_{x:|x-y|>\epsilon} \frac{e^{d-2}d\nu(x)}{|x-y|^{d-2}}.$$

Dividing by ϵ^{d-2} and letting $\epsilon \to 0$ we obtain

$$\int_\Lambda \frac{d\nu(x)}{|x-y|^{d-2}} \leq \frac{1}{|y|^{d-2}},$$

i.e. $\int_\Lambda K(x,y)d\nu(x) \leq 1$ for all $y \in \Lambda$. Therefore $I_K(\nu) \leq \nu(\Lambda)$ and thus

$$\operatorname{Cap}_K(\Lambda) \geq [I_K(\nu/\nu(\Lambda))]^{-1} \geq \nu(\Lambda),$$

which by (22.3) yields the upper bound on the probability of hitting Λ.

To obtain a lower bound for this probability, a second moment estimate is used. It is easily seen that the Martin capacity of Λ is the supremum of the capacities of its compact subsets, so we may assume that Λ is itself compact. For $\epsilon > 0$ and $y \in \mathbb{R}^d$ let $B_\epsilon(y)$ denote the Euclidean ball of radius ϵ about y and let $h_\epsilon(|y|)$ denote the probability that the standard Brownian path hits this ball, that is $(\epsilon/r)^{d-2}$ if $r > \epsilon$, and 1 otherwise.

Given a probability measure μ on Λ, and $\epsilon > 0$, consider the random variable

$$Z_\epsilon = \int_\Lambda \mathbf{1}(\exists t > 0 : W_t \in B_\epsilon(y))h_\epsilon(|y|)^{-1}d\mu(y).$$

Clearly $\mathbb{E}Z_\epsilon = 1$. By symmetry, the second moment of Z_ϵ can be written as:

$$\begin{aligned}
\mathbb{E}Z_\epsilon^2 &= 2\mathbb{E}\int_\Lambda\int_\Lambda \mathbf{1}(\exists t > 0 : W_t \in B_\epsilon(x), \exists s > t : W_s \in B_\epsilon(y))\frac{d\mu(x)d\mu(y)}{h_\epsilon(|x|)h_\epsilon(|y|)} \\
&\leq 2\mathbb{E}\int_\Lambda\int_\Lambda \mathbf{1}(\exists t > 0 : W_t \in B_\epsilon(x))\frac{h_\epsilon(|y-x|-\epsilon)}{h_\epsilon(|x|)h_\epsilon(|y|)}d\mu(x)d\mu(y) \\
&= 2\int_\Lambda\int_\Lambda \frac{h_\epsilon(|y-x|-\epsilon)}{h_\epsilon(|y|)}d\mu(x)d\mu(y) \qquad (22.5)
\end{aligned}$$

The last integrand is bounded by 1 if $|y| \leq \epsilon$. On the other hand, if $|y| > \epsilon$ and $|y - x| \leq 2\epsilon$ then $h_\epsilon(|y-x| - \epsilon) = 1 \leq 2^{d-2} h_\epsilon(|y-x|)$, so that the integrand on the right hand side of (22.5) is at most $2^{d-2}K(x,y)$. Thus

$$\mathbb{E}Z_\epsilon^2 \leq 2\mu(B_\epsilon(0)) + 2^{d-1}\int\int \mathbf{1}(|y-x| \leq 2\epsilon)K(x,y)d\mu(x)d\mu(y)$$
$$+2\int\int \mathbf{1}(|y-x| > 2\epsilon)\left(\frac{|y|}{|y-x|-\epsilon}\right)^{d-2} d\mu(x)d\mu(y). \tag{22.6}$$

Since the kernel is infinite on the diagonal, any measure with finite energy must have no atoms. Restricting attention to such measures μ, we see that the first two summands in (22.6) drop out as $\epsilon \to 0$ by dominated convergence. This leaves

$$\lim_{\epsilon \downarrow 0} \mathbb{E}Z_\epsilon^2 \leq 2I_K(\mu). \tag{22.7}$$

Clearly the hitting probability $\mathbb{P}(\exists t > 0, y \in \Lambda\ : W_t \in B_\epsilon(y))$ is at least

$$\mathbb{P}(Z_\epsilon > 0) \geq \frac{(\mathbb{E}Z_\epsilon)^2}{\mathbb{E}Z_\epsilon^2} = (\mathbb{E}Z_\epsilon^2)^{-1}.$$

Transience of Brownian motion implies that if the Brownian path visits every ϵ-neighborhood of the compact set Λ then it almost surely intersects Λ itself. Therefore, by (22.6):

$$\mathbb{P}(\exists t > 0, y \in \Lambda\ : W_t \in B_\epsilon(y)) \geq \lim_{\epsilon \downarrow 0}(\mathbb{E}Z_\epsilon^2)^{-1} \geq \frac{1}{2I_K(\mu)}.$$

Since this is true for all probability measures μ on Λ, we get the desired conclusion:

$$\mathbb{P}(\exists t > 0, y \in \Lambda\ : W_t \in B_\epsilon(y)) \geq \frac{1}{2}\operatorname{Cap}_K(\Lambda). \quad \blacksquare$$

REMARK. The right-hand inequality in (22.2) can be an equality– a sphere centered at the origin has hitting probability and capacity both equal to 1. The next exercise shows that the constant 1/2 on the left cannot be increased.

EXERCISE 22.4. Consider the spherical shell

$$\Lambda_R = \{x \in \mathbb{R}^d\ :\ 1 \leq |x| \leq R\}.$$

Show that $\lim_{R \to 0} \operatorname{Cap}_K(\Lambda_R) = 2$.

23. Kaufman's theorem on uniform dimension doubling

In the previous sections we have seen the dimension doubling property of Brownian motion: for any set A, we have $\dim W(A) = 2\dim(A)$ a.s. For $d \geq 2$ this is true for *all closed sets* a.s. To see the distinction, consider the zero set of 1 dimensional Brownian motion.

LEMMA 23.1. *Consider a cube $Q \subset \mathbb{R}^d$ centered at a point x and having diameter $2r$. Let W be Brownian motion in $\mathbb{R}^d$, with $d \geq 3$. Define recursively*

$$\begin{aligned}\tau_1 &= \inf\{t \geq 0 :\ W(t) \in Q\}\\ \tau_{k+1} &= \inf\{t \geq \tau_k + r^2 :\ W(t) \in Q\}\end{aligned}$$

with the usual convention that $\inf \emptyset = \infty$. *Brownian motion will leave the cube forever after a geometric number of* τ_n*-s, more precisely there exists a positive* $\theta = \theta_d < 1$ *which does not depend on* r *and* z, *such that* $\mathbb{P}_z(\tau_{n+1} < \infty) \le \theta^n$.

PROOF. It is sufficient to show that for some θ as above,

$$\mathbb{P}_z(\tau_{k+1} = \infty \mid \tau_k < \infty) > 1 - \theta.$$

But the quantity on the right can be bounded below by

$$\mathbb{P}_z(\tau_{k+1} = \infty \| W(\tau_k + r^2) - x| > 2r, \tau_k < \infty)\mathbb{P}_z(|W(\tau_k + r^2) - x| > 2r|\tau_k < \infty)$$

The second factor is clearly positive, and the first is also positive since W is transient. Both probabilities are invariant under changing the scaling factor r. ■

COROLLARY 23.2. *Let* D_m *denote the set of binary cubes of side length* 2^{-m} *inside* $[-\frac{1}{2}, \frac{1}{2}]^d$. *A.s. there exists a random variable* $C(\omega)$ *so that for all* m *and for all cubes* $Q \in D_m$ *we have* $\tau^Q_{K+1} = \infty$ *with* $K = C(\omega)m$.

PROOF.

$$\sum_m \sum_{Q \in D_m} \mathbb{P}(\tau^Q_{\lceil cm+1 \rceil} < \infty) \le \sum_m 2^{dm}\theta^{cm}$$

Choose c so that $2^d\theta^c < 1$. Then by Borel-Cantelli, for all but finitely many m-s we have $\tau^Q_{\lceil cm+1 \rceil + 1} = \infty$ for all $Q \in D_m$. The number of the finitely many exceptions is random. Finally, we can choose a random $C(\omega) > c$ to handle the exceptional cubes. ■

THEOREM 23.3 (Kaufman's Uniform Dimension Doubling for $d \ge 3$).

$$\mathbb{P}(\dim W(A) = 2 \dim A \textit{ for all closed } A \in [0,\infty]) = 1 \tag{23.1}$$

PROOF. The $\le$ direction holds in all dimensions by the Hölder property of Brownian motion (see Corollary 4 and Lemma 5.11).

Fix L. We will show that with probability one, all closed subsets S of $[-L, L]^d$ we have

$$\dim W^{-1}(S) \le \frac{1}{2} \dim S. \tag{23.2}$$

Applying this to $S = W(A) \cap [-L, L]$ for a countable unbounded set of L we get the desired conclusion. By scaling, it is sufficient to prove (23.2) for $L = \frac{1}{2}$. We will verify this for the paths satisfying Corollary 23.2; these have full measure. The rest of the argument is deterministic, we fix an ω belonging to such a path.

For $\beta > \dim S$ and for all ϵ there exist covers of S by binary cubes $\{Q_j\}$ in $\bigcup_m D_m$ so that $\sum |Q_j|^\beta < \epsilon$. If N_m denotes the number of cubes from D_m in such a cover, then

$$\sum_m N_m 2^{-m\beta} < \epsilon.$$

Consider the W-inverse image of these cubes. Since we chose ω so that Corollary 23.2 is satisfied, this yields a cover of $W^{-1}(S)$, which for each $m \ge 1$ uses at most $C(\omega)mN_m$ intervals of length $r^2 = d2^{-2m}$.

For $\beta_1 > \beta$ we can bound the $\beta_1/2$ dimensional Hausdorff content of W^{-1} above by

$$\sum_{m=1}^{\infty} C(\omega) m N_m (d2^{-2m})^{\beta_1/2} = C(\omega) d^{\beta_1/2} \sum_{m=1}^{\infty} m N_m 2^{-m\beta_1}.$$

This can be made small by choosing a suitable ϵ. Thus $W^{-1}(S)$ has Hausdorff dimension at most $\beta/2$ for all $\beta > \dim S$, and thus, at most $\dim S/2$. ∎

In two dimensions we cannot rely on transience of Brownian motion. To get around this problem, we can look at the Brownian path up to a stopping time. A convenient one is

$$\tau_R^* = \min\{t \;:\; |W(t)| = R\}.$$

For the two dimensional version of Kaufman's theorem it is sufficient to show that

$$\mathbb{P}(\dim W(A) = 2\dim(A \cap [0, \tau_R^*]) \text{ for all closed } A \in [0,\infty]) = 1.$$

Lemma 23.1 has to be changed accordingly. Define τ_k as in (23.1), and assume that the cube Q is inside the ball of radius R about 0. Then we have

LEMMA 23.4.

$$\mathbb{P}_z(\tau_k < \tau_R^*) \le \left(1 - \frac{c}{m}\right)^k \le e^{cn/m} \tag{23.3}$$

Here $c = c(R) > 0$, $2^{-m-1} < r < 2^{-m}$, and z is any point in $\mathbb{R}^d$.

PROOF. It suffices to bound $\mathbb{P}_z(\tau_{k+1} \ge \tau_R^* \mid \tau_k < \tau_R^*)$ from below by

$$\mathbb{P}_z(\tau_{k+1} \ge \tau_R^* |\; W(\tau_k + r^2) > 2r, \tau_k < \tau_R^*) \mathbb{P}_z(|W(\tau_k + r^2) - x| > 2r \;|\tau_k < \tau_R^*)$$

The second factor does not depend on r and R, and it can be bounded below by a constant. The first factor is bounded below be the probability that planar Brownian motion started at distance $2r$ from the origin hits the sphere of radius $2R$ before the sphere of radius r (both centered at the origin). Using (21.1), this is given by

$$\frac{\log_2 \frac{2r}{r}}{\log_2 \frac{2R}{r}} \ge \frac{1}{\log_2 R + 1 + m}.$$

This is at least c_1/m for some c_1 which depends on R only. ∎

The bound (23.3) on $P(\tau_k < \infty)$ in two dimensions is worse by a linear factor than the bound in higher dimensions. This, however, does not make a significant difference in the proof of the two dimensional version of Theorem 23.3.

EXERCISE 23.5. Prove Theorem 23.3 for two dimensions.

The following are from week 15, (W 5/6, F 5/8, 1998).
Notes taken by Yoram Gat, `yoram@stat.berkeley.edu`.

24. Packing dimension

Tricot (1980, 1982) introduced **packing dimension**, which plays a dual role to Hausdorff dimension in many settings. For our present purpose, the representation of packing dimension which is convenient to use as a definition, is as a regularization of upper Minkowski dimension:

$$\dim_{\mathrm{p}}(A) = \inf_{A \subset \cup_j A_j} \sup_j \overline{\dim}_{\mathcal{M}}(A_j)\,,$$

where the infimum is over all countable covers of A. (See Tricot (1982), Proposition 2, or Falconer (1990), Proposition 3.8.) The infimum may also be taken only over countable covers of A with closed sets since $\overline{\dim}_{\mathcal{M}}(A_j) = \overline{\dim}_{\mathcal{M}}(\overline{A_j})$.

Note that for packing dimension, unlike the Minkowski dimension, the following property holds for any partition A_j, $j = 1, 2, \ldots$ of A:

$$\dim_{\mathrm{p}} A = \sup_j \dim_{\mathrm{p}} A_j.$$

Part (i) of the next lemma is due to Tricot (1982) (see also Falconer 1990); Part (ii) for trees can be found in Benjamini and Peres (1994, Proposition 4.2(b)); the general version given is in Falconer and Howroyd (1994) and in Mattila and Mauldin (1994).

LEMMA 24.1. *Let A be a closed subset of $\mathbb{R}$.*

(i) *If any open set V which intersects A satisfies $\overline{\dim}_{\mathcal{M}}(A \cap V) \geq \alpha$, then $\dim_{\mathrm{p}}(A) \geq \alpha$.*

(ii) *If $\dim_{\mathrm{p}}(A) > \alpha$, then there is a (relatively closed) nonempty subset $\widetilde{A}$ of A, such that $\dim_{\mathrm{p}}(\widetilde{A} \cap V) > \alpha$ for any open set V which intersects $\widetilde{A}$.*

PROOF. (i) If A_j is closed and for any V open s.t. $V \cap A \neq \emptyset$, it holds that $V \not\subset A_j$, then A_j^c is a dense open set relative to A. By Baire's category theorem $A \cap \bigcap_j A_j^c \neq \emptyset$, i.e., $A \not\subset \bigcup_j A_j$. Therefore, if $A \subset \bigcup_j A_j$ then there exists an open set, V, s.t. $V \cap A \neq \emptyset$ and s.t. $V \subset A_j$.

(ii) Define

$$\widetilde{A} = A \setminus \bigcup \{J \text{ rational interval } : \dim_{\mathcal{M}}(J \cap A) \leq \alpha\}.$$

Then, any countable cover of $\widetilde{A}$ together with the sets removed on the right yields a countable cover of A, giving $\dim_{\mathrm{p}} \widetilde{A} \vee \alpha \geq \dim_{\mathrm{p}} A > \alpha$. Since $\widetilde{A} \subset A$, we conclude that $\dim_{\mathrm{p}} \widetilde{A} = \dim_{\mathrm{p}} A > \alpha$. If for some V open, $V \cap \widetilde{A} \neq \emptyset$ and $\dim_{\mathrm{p}}(\widetilde{A} \cap V) \leq \alpha$ then V contains some rational interval J s.t. $\widetilde{A} \cap J \neq \emptyset$. For that J,

$$\begin{aligned}\dim_{\mathrm{p}}(A \cap J) &\leq \max(\dim_{\mathrm{p}}(A \setminus \widetilde{A}), \dim_{\mathrm{p}}(\widetilde{A} \cap J)) \\ &\leq \max(\alpha, \dim_{\mathrm{p}}(\widetilde{A} \cap V)) \;=\; \alpha,\end{aligned}$$

contradicting the construction of $\widetilde{A}$. ∎

25. Hausdorff dimension and random sets

Fix p. For $n \geq 0$, define $Q^n(p)$, to be random subsets of $[0,1]$ as follows. For $n > 1$, $0 \leq i < 2^i$ let U_i^n be i.i.d. random variables uniform on $[0,1]$. Let

$$Q^n(p) = \bigcup_{i\,:\,U_i^n \leq p} \big(i2^{-n}, (i+1)2^{-n}\big)$$

Define $Q(p) = \bigcap_n Q^n(p)$.

We call an interval I a **binary interval** if its endpoints are $i2^{-n}$, $(i+1)2^{-n}$ for some integers i, n. Note that each interval I can be covered by two closed binary intervals of length at most $|I|$. Thus we can require covers to be made of binary intervals in the definition Hausdorff dimension for subsets of $\mathbb{R}$.

LEMMA 25.1. *If $A \subset [0,1]$ intersects the random set $Q(2^{-\alpha})$ with positive probability then* $\dim_{\mathcal{H}}(A) \geq \alpha$.

PROOF. Let $b = \mathbb{P}(A \cap Q(2^{-\alpha}))$. Then for any countable collection of binary subintervals of $[0,1]$ s.t. $A \subset \bigcup_j I_j$ we have

$$b \leq \sum_j \mathbb{P}(I_j \cap Q(2^{-\alpha}) \neq \emptyset) \leq \sum_j \mathbb{P}(I_j \cap Q^{n_j}(2^{-\alpha}) \neq \emptyset)$$

Here n_j is defined so that I_j is of length 2^{-n_j}. Thus the summand on the right equals $(2^{-\alpha})^{n_j} = |I_j|^\alpha$. Therefore $b \leq \sum_j |I_j|^\alpha$. However, if $\dim_{\mathcal{H}}(A) = \alpha_1 < \alpha$, then there exists a collection of binary intervals $\{I_j\}$ s.t. $A \subset \bigcup_j I_j$ and s.t. $\sum_j |I_j|^\alpha < b$. ■

COROLLARY 25.2. *If a random set $A \subset [0,1]$ intersects the independent random set $Q(2^{-\alpha})$ with positive probability then* $||\dim_{\mathcal{H}}(A)||_\infty \geq \alpha$.

PROOF. By assumption

$$0 < \mathbb{P}(A \cap Q \neq \emptyset) = \mathbb{E}\mathbb{P}(A \cap Q \neq \emptyset | A)$$

By the previous lemma, when $\mathbb{P}(A \cap Q \neq \emptyset | A) > 0$ we also have $\dim_{\mathcal{H}} \geq A$, so the right hand side equals

$$\begin{aligned} \mathbb{E}[\mathbf{1}(\{\dim_{\mathcal{H}} A \geq \alpha\})\mathbb{P}(A \cap Q \neq \emptyset | A)] &\leq \mathbb{E}\mathbf{1}(\{\dim_{\mathcal{H}} A \geq \alpha\}) \\ &= \mathbb{P}(\dim_{\mathcal{H}} A \geq \alpha). \quad ■ \end{aligned}$$

LEMMA 25.3. *If a random set A has $\mathbb{P}(A \cap K \neq \emptyset) > 0$ for all fixed K s.t. $\dim_{\mathcal{H}} K \geq \beta$, then* $||\dim_{\mathcal{H}}(A)||_\infty \geq 1 - \beta$.

PROOF. We show that $||\dim_{\mathcal{H}} Q(2^{-\alpha})||_\infty \geq 1 - \alpha$. The hypothesis of the lemma will then imply that for Q independent of A, $\mathbb{P}(A \cap Q(2^{-(1-\beta-\epsilon)} \neq \emptyset) > 0$ and thus by Corollary 25.2 $||\dim_{\mathcal{H}}(A)||_\infty \geq \beta$.

Let Q_1, Q_2 be two independent random sets s.t. $Q_1 \overset{\mathrm{d}}{=} Q(2^{-\alpha})$ and $Q_2 \overset{\mathrm{d}}{=} Q(2^{\alpha-1+\epsilon})$. Then $Q_1 \cap Q_2 \overset{\mathrm{d}}{=} Q(2^{\epsilon-1})$. By the theory of branching processes, $\mathbb{P}(Q(2^{\epsilon-1}) \neq \emptyset) > 0$. Thus, using Corollary 25.2 again, and taking $\epsilon \to 0$,

$$\big|\big|\dim_{\mathcal{H}} Q(2^{-\alpha})\big|\big|_\infty \geq 1 - \alpha. \quad ■$$

26. Lévy's modulus of continuity

The following technical lemma is closely related to the discussion in the previous section. After stating it, we show how it implies a lower bound on the modulus of continuity for Brownian motion.

LEMMA 26.1. *Suppose that for each n and open binary interval $I \subset [0,1]$ of length 2^{-n} we have an indicator random variable Z_I with $\mathbb{P}(Z_I = 1) = p_n$. Suppose further that if $|I| = |J|$ and $\operatorname{dist}(I,J) > \frac{cn}{2^n}$, then Z_I and Z_J are independent.*

$$\begin{aligned} A_n &= \bigcup\{I : |I| = 2^{-n}, Z_I = 1\}, \\ A &= \limsup A_n = \textstyle\bigcap_n \bigcup_{m\geq n} A_n. \end{aligned}$$

If we have

$$\liminf_{n\to\infty} \frac{\log_2 p_n}{n} \geq -\theta,$$

then for any fixed $E \subset [0,1]$ closed, the following statements hold:
(i) *If $\dim_{\mathrm{p}} E > \theta$ then $\mathbb{P}(E \cap A \neq \emptyset) = 1$,*
(ii) *$||\dim_{\mathcal{H}} A||_\infty \geq 1 - \theta$.*

We will now use this lemma to prove the lower bound for Lévy's modulus of continuity.

THEOREM 26.2. *Let $\varphi(h) = (2h\log(1/h))^{1/2}$, and $E \subset [0,1]$. Then*

$$\sup_{t\in E}\ \limsup_{h\downarrow 0} \frac{|B_{t+h} - B_t|}{\varphi(h)} \geq \sqrt{\dim_{\mathrm{p}} E}.$$

In particular, for $E = [0,1]$ the lower bound is 1. Moreover, the (random) Hausdorff dimension of the set of times $t \in [0,1]$ where

$$\limsup_{h\downarrow 0} \frac{|B_{t+h} - B_t|}{\varphi(h)} \geq \lambda$$

satisfies $\|\dim_{\mathcal{H}}\|_\infty \geq 1 - \lambda^2$.

PROOF. For each binary interval $I = (a2^{-n}, (a+1)2^{-n})$, set the indicator variable $Z_I = 1$ iff the following two conditions hold:

$$\sup_{t\in I} \left|B(t) - B((a+1)2^{-n})\right| \leq 2^{-n/2}, \tag{26.1}$$

$$\left|B((a+n)2^{-n}) - B((a+1)2^{-n})\right| \geq \lambda\varphi((n-1)2^{-n}). \tag{26.2}$$

Suppose that $Z_I = 1$. For $t \in I$, set $h = (a+n)2^{-n} - t$, so that $h \in ((n-1)2^{-n}, n2^{-n})$. Using the triangle inequality we get

$$|B(t+h) - B(t)| \geq \lambda\varphi((n-1)2^{-n}) - 2^{-n/2}$$

dividing by $\varphi(h) \leq \varphi(n2^{-n})$:

$$\frac{|B(t+h) - B(t)|}{\varphi(h)} \geq \left(\varphi((n-1)2^{-n}) - 2^{-n/2}\right)/\varphi(n2^{-n}) \to \lambda$$

as $n \to \infty$. Therefore, if t is contained in the intersection of a sequence I_j of binary intervals with $|I_j| \downarrow 0$ and $Z_{I(j)} = 1$, then

$$\limsup_{h \downarrow 0} \frac{|B_{t+h} - B_t|}{\varphi(h)} \geq \lambda. \tag{26.3}$$

The events (26.1) and (26.2) are independent, so by scaling we can write

$$\mathbb{P}(Z_I = 1) = \mathbb{P}\left(\sup_{t \in (0,1)} |B_t - B_1| < 1 \right) \mathbb{P}\left(|B_1|^2 \geq \lambda \sqrt{2 \log(2^n/(n-1))} \right).$$

The first probability is a nonzero constant. For the second, we can use the usual tail estimate $\mathbb{P}(|B_1| > x) > cx^{-1}e^{-x/2}$ for $x > 1$ (see Lemma 2.5) to conclude that

$$p_n \stackrel{\text{def}}{=} \mathbb{P}(Z_I = 1) \geq \left(c_\lambda \left[\log(2^n/(n-1)) \right]^{-1/2} (n-1)^{\lambda^2} \right) 2^{-n\lambda^2}.$$

In particular, $\liminf(\log_2 p_n)/n \geq -\lambda^2$. Finally, note that the random variables Z_I satisfy the conditions of Lemma 26.1 with $\theta = \lambda^2$, and that the theorem follows from the conclusions of the Lemma. ∎

Proof of Lemma 26.1. Let $E \subset [0,1]$ be closed with $\dim_{\mathrm{p}} E > \theta$. Let $\tilde{E}$ be defined as in Lemma 24.1 (ii), i.e.,

$$\tilde{E} = E \setminus \bigcup \{ J \text{ rational interval} : \dim_{\mathcal{M}}(J \cap E) < \theta \}.$$

From the proof of Lemma 24.1 we have $\dim_{\mathrm{p}} E = \dim_{\mathrm{p}} \tilde{E}$. Define

$$A_n^* = \textstyle\bigcup_{m \geq n} A_m$$

By definition $A_n^* \cap \tilde{E}$ is open in $\tilde{E}$. We will show that it is also dense in $\tilde{E}$. This, by Baire's category theorem will imply (i).

To show that $A_n^* \cap \tilde{E}$ is dense in $\tilde{E}$, we need to show that for any binary interval J which intersects $\tilde{E}$, $A_n^* \cap \tilde{E} \cap J$ is a.s. non-empty.

For the rest of the proof, take n large enough so that $\tilde{E} \cap J$ intersects more than $2^{n(\theta+2\epsilon)}$ binary intervals of length 2^{-n}, and so that $(\log p_n)/n > -\theta - \epsilon$. Let S_n be the set of the intervals intersected. Define $T_n = \sum_{I \in S_n} Z_I$, so that

$$\mathbb{P}(A_n \cap \tilde{E} \cap J = \emptyset) = \mathbb{P}(T_n = 0).$$

To show that this probability converges to zero, it suffices to prove that $\operatorname{Var} T_n/(\mathbb{E}T_n)^2$ does. The first moment of T_n is given by

$$\mathbb{E}T_n = |S_n| p_n > e^{(\theta+2\epsilon)n} e^{-\theta-\epsilon n} = e^{\epsilon n}$$

The variance can be written as

$$\operatorname{Var} T_n = \operatorname{Var} \sum_{I \in S_n} Z_I = \sum_{I \in S_n} \sum_{J \in S_n} \operatorname{Cov}(Z_I Z_J).$$

Here each summand is at most p_n, and the summands for which I and J are far are 0 by assumption. Thus the above is at most

$$p_n \#\{(I,J) \in S_n \times S_n : |I| - |J| \leq cn2^{-n}\} \leq p_n c' n |S_n| = c' n \mathbb{E}T.$$

This means that $\operatorname{Var} T_n/(\mathbb{E}T_n)^2 \to 0$. Thus we showed that A_n^* is a.s. an open dense set, concluding the proof of (i).

Claim (ii) follows from (i) and Lemma 25.3. ∎

References

Adelman, O. (1985). Brownian motion never increases: a new proof of a theorem of Dvoretzky, Erdős and Kakutani. *Israel J. of Math.* **50**, 189–192.

Barlow, M.T. and Perkins, E. (1984). Levels at which every Brownian excursion is exceptional. *Lecture Notes in Math.* **1059**, Springer, Berlin-New York, 1–28.

Bass, R. (1995). *Probabilistic Techniques in Analysis*, Springer-Verlag, New York.

Bass, R. and Burdzy, K. (1997). Cutting Brownian Paths. *Memoir AMS.*

Benjamini, I., Pemantle, R. and Peres, Y. (1995). Martin capacity for Markov chains. *Ann. Prob.* **23**, 1332–1346

Benjamini, I. and Peres, Y. (1994). Tree-indexed Random walks on groups and first passage percolation. *Probab. Th. Rel. Fields.* **98**, 91–112.

Berman, S. M. (1983). Nonincrease almost everywhere of certain measurable functions with applications to stochastic processes. *Proc. Amer. Math. Soc.* **88**, 141–144.

Bertoin, J. (1991). Increase of a Lévy process with no positive jumps. *Stochastics* **37**, 247–251.

Bickel, P. (1967). Some contributions to the theory of order statistics. *Proc. Fifth Berkeley Sympos. Math. Statist. and Probability, Vol I: Statistics*, 575–591.

Burdzy, K. (1990). On nonincrease of Brownian motion. *Ann. Probab.* **18**, 978–980.

Burdzy, K. and Lawler, G. (1990). Non-intersection exponents for Brownian paths. Part II: Estimates and Applications to a random fractal. *Ann. Probab.* **18**, 981–1009.

Burdzy, K. (1989). Cut points on Brownian paths. *Ann. Probab.* **17**, 1012–1036.

Christensen, J.P.R. (1972). On sets of Haar measure zero in abelian Polish groups. *Israel J. Math.* **13**, 255–260.

Doney, R. A. (1994). Increase of Lévy processes. *Preprint.*

Dubins, L.E. (1968). On a theorem of Skorokhod *Ann. Math. Statist.* **39** 2094–2097.

Dudley, R.M. (1989) *Real Analysis and Probability.* Wadsworth & Brooks, Pacific Grove, CA.

Durrett, R. (1996). *Probability: Theory and Examples.* 2nd ed. Duxbury Press, Belmont CA

Dvoretzky, A., Erdős, P. and Kakutani, S. (1961). Nonincrease everywhere of the Brownian motion process. *Proc. Fourth Berkeley Symposium,* Vol. **II**, 103–116.

Falconer, K.J. (1990). *Fractal geometry: mathematical foundations and applications.* John Wiley & Sons.

Falconer, K.J. and Howroyd, J. D. (1994). Projection theorems for Box and Packing dimension. *To appear in Math. Proc. Camb. Phil. Soc.*

Feller, W. (1966). *An introduction to probability theory and its applications*, Volume II. John Wiley and Sons: New York.

Ford, L.R. Jr., and Fulkerson, D.R. (1956). Maximal flow through a network. *Canad. J. Math.* **8** 399–404.

Freedman, D. (1971). *Brownian motion and diffusion.* Springer-Verlag, New York.

Frostman, O. (1935). Potentiel d'equilibre et capacité des ensembles. Thesis.

Frostman, O. (1935). Potential d'équilibre et capacité des ensembles avec quelques applications a la théorie des fonctions, Medd Lund. Univ. Math. Seminar **3**.

Harris, T. E. (1960). A lower bound for the critical probability in a certain percolation process. *Proc. Camb. Phil. Soc.* **56**, 13–20.

Hausdorff, F. (1919). Dimension und äusseres Mass, *Math. Annalen* **79**, 157-179

Hewitt, E. and Stromberg, K. (1965). *Real and abstract analysis.* Springer-Verlag, New York.

Hunt, G. (1956). Some theorems concerning Brownian motion. *Trans. Amer. Math. Soc.* **81**, 294–319.

Hunt, B.R., Sauer, T. and Yorke, J.A. (1992). Prevalence: a translation-invariant "almost every" on infinite-dimensional spaces. *Bull. Amer. Math. Soc.* **27** (2), 217–238.

Hunt, B.R., Sauer, T. and Yorke, J.A. (1993). An addendum to: "Prevalence: a translation-invariant 'almost every' on infinite dimensional spaces". *Bull. Amer. Math. Soc.* **28** (2), 306–307.

Kakutani, S. (1944). Two dimensional Brownian motion and harmonic functions. *Proc. Imp. Acad. Tokyo* **20**, 648–652.

Kaufman, R. (1969), Une propriété métrique du mouvement brownien *C. R. Acad. Sci. Paris* **268**, 727-728.

Lehmann, E. (1966). Some concepts of dependence. *Ann. Math. Statist.* **37**, 1137–1153.

Lévy, P. (1948) *Processus Stochastiques et Mouvement Brownien.* Gauthier-Villars, Paris.

Lévy, P. (1954). La mesure de Hausdorff de la courbe du mouvement brownien. *Giorn. Ist. Ital. Attuari* **16**, 1–37.

Mattila, P. and Mauldin, R. D. (1994). Measure and dimension functions: measurability and densities. *University of Jyväskylä Preprint* **175**.

Knight, F. B. (1981). *Essentials of Brownian motion and diffusion.* American Math. Soc., Providence, Rhode Island.

McKean, H.P. (1955). Hausdorff-Besicovitch dimension of Brownian motion paths. *Duke Math J.* **22**, 229–234.

Paley R.E.A.C., Wiener N. and Zygmund A. (1933) Notes on random functions. *Math. Z.* **37**, 647–668

Peres, Y. (1996). Points of Increase for Random Walks. *Israel J. of Math.* **95**, 341–347.

Polya, G., Tarjan, R., and Woods, D. (1983). *Notes on Introductory Combinatorics.* Birkhäuser, Boston.

Przytycki F. and Urbanski M. (1989). On the Hausdorff dimension of some fractal sets *Studia Math.* **93**, 155–186.

Revuz D. and Yor M. (1994). *Continuous Martingales and Brownian Motion.* Springer-Verlag, Berlin, New York, 2nd edition.

Root D.H. (1969). The existence of certain stopping times on Brownian Motion. *Ann. Math. Statist.* **40** 715–718.

Skorokhod A. (1965). *Studies in the Theory of Random Processes.* Addison-Wesley, Reading, MA.

Strassen, V. (1964). An invariance principle for the law of the iterated logarithm. *Z. Warsch. Verw Gebiete* **3**, 211–226.

Taylor S.J (1953). The Hausdorff α-dimentional measure of Brownian paths in n-space, *Proc. Cambridge Philos. Soc.* **49**, 31–39.

Tricot, C. (1980). Rarefaction indices. *Mathematika* **27**, 46–57.

Tricot, C. (1982). Two definitions of fractal dimension. *Math. Proc. Camb. Phil. Soc.* **91**, 57–74.

Tricot, C. (1991). Rectifiable and fractal sets. In *Fractal Geometry and Analysis*, NATO ASI Series C: v. 346, pp. 532–541, Edited by J. Bélair and S. Dubuc, Kluwer Academic Publishers.

Wiener N. (1923). Differential Spaces, *J. Math. Phys.*