University of Chicago - Spring 2026

Probability theory and Brownian motion

Vincent X. Wang

June 6, 2026

These are lecture notes taken from the second half of MATH31400 at University of Chicago, taught by Ewain Gwynne during Spring 2026.

Contents

Probability

We begin with some broad motivations for studying probability theory:

- Sometimes, some random objects can be easier to study than deterministic objects. For example, we can assume that three points on a disk are not collinear with probability 1.
- Quantum mechanics suggests the real world is probabilistic.
- Probabilistic method for combinatorics
- The distribution of prime numbers can be vaguely considered as a probability distribution
- One can construct solutions to PDEs via Brownian motion
- Applications to applied fields like... finance...

Basic notions

Definition 1. A **probability space** is a triple $(\Omega, \mathcal{F}, \mathbb{P})$, where:

1. Ω is the set of possible states,
2. $\mathcal{F}$ is a σ-algebra on Ω consisting of "events"
3. $\mathbb{P}$ is a non-negative measure on $\mathcal{F}$ with total mass 1 (i.e., $\mathbb{P}(\Omega) = 1$).

Remark 2. Recall a σ-algebra is closed under complement, countable unions, and contains Ω.

In general, we will implicitly fix some $(\Omega, \mathcal{F}, \mathbb{P})$ in the backdrop.

Definition 3. We say an event $E \in \mathcal{F}$ occurs **almost surely (a.s.)** if $\mathbb{P}[E] = 1$.

Definition 4. Let $\mathcal{X}$ be a topological space equipped with a Borel σ-algebra. A **random variable** in $\mathcal{X}$ is a measurable function $X : \Omega \to \mathcal{X}$. We identify random variables X, X' that agree almost surely; that is,

$$\mathbb{P}(\{\omega \in \Omega : X(\omega) = X'(\omega)\}) = 1$$

In general, for an event $A \subset \mathcal{X}$ Borel, we will fix the notation

$$\{X \in A\} = \{w \in \Omega : X(w) \in A\}.$$

Remark 5. Thus, X and X' agree almost surely if $\mathbb{P}[X = X'] = 1$.

Definition 6. Let X be a random variable taking values in $\mathcal{X}$. The **distribution of** X is the probability measure

$$\mu_X(A) := \mathbb{P}[X \in A]$$

for all $A \subset X$ Borel. We say two variables X, Y agree in distribution if $\mu_X = \mu_Y$, denoted $X \stackrel{d}{=} Y$.

In probability, we generally only care about the distribution of collections of random variables, and the precise choice of our probability space is not important. As an analogy, picking a probability space is somewhat like picking a chart for a manifold.

Remark 7. Alternatively, one can write that $\mu_X(A)$ is the *pushforward measure* of $\mathbb{P}$ under X, as

$$\mu_X(A) = \mathbb{P}[X^{-1}(A)].$$

We give some examples of random variables:

Example 8 (Trivial probability space)**.** Let $\mathcal{B}$ be the Borel σ-algebra on $\mathcal{X}$, and μ_X a distribution on $\mathcal{X}$. Then, $(\mathcal{X}, \mathcal{B}, \mu_X)$ is a probability space with $\mathrm{id} : \mathcal{X} \to \mathcal{X}$ a random variable with distribution μ_X.

Remark 9. This example demonstrates that we can always construct a probability space from a distribution, but it is not very useful to actually look at.

Example 10 (Rolling fair dice)**.** Let X_1, X_2 be two random variables representing dice rolls, taking values in $\{1, 2, 3, 4, 5, 6\}$. For any reasonable construction, we should have $X_1 \overset{d}{=} X_2$, but $X_1 \neq X_2$. We can explicitly construct X_1 and X_2 by specifying $\Omega = \{1, 2, 3, 4, 5, 6\}^2$, or alternatively $\Omega = \{\text{all possible states of the world}\}$. The latter, however, may not be that helpful.

Example 11 (Countable space)**.** Let $\mathcal{X}$ be countable and equipped with the discrete topology. Then, μ_X is specified by $\mathbb{P}[X = x]$ for all $x \in \mathcal{X}$.

Example 12. Let $\mathcal{X} = \mathbb{R}$. Then, μ_X is specified by the **cumulative distribution function (CDF)** $F(x) := \mathbb{P}[X \leq x] = \mu_X((-\infty, x])$.

The CDF $F : \mathbb{R} \to [0, 1]$ satisfies some important properties:

- F is non-decreasing and continuous from the right (i.e., $F(x) = \lim_{y \searrow x} F(y)$. This follows from the downward continuity of measure.
- $F(x) - \lim_{y \nearrow x} F(y) = \mathbb{P}[X = x]$.
- $\lim_{x \to -\infty} F(x) = 0$ and $\lim_{x \to \infty} F(x) = 1$.

Remark 13. A continuous CDF implies that there are no discrete point masses.

If we can find a function F satisfying all of these properties, then conversely construct a probability measure $\mu((a, b]) = F(b) - F(a)$, which is called a Lebesgue-Stieltjes measure.

Definition 14. Let X be a random variable with values in $\mathbb{R}^d$. If μ_X is absolutely continuous with respect to the Lebesgue measure λ, then the Radon-Nikodym derivative

$$f = \frac{d\mu_X}{d\lambda}$$

is called the **density** of X.

Remark 15. If $d = 1$, then $f = F'$.

Definition 16. Let X be a random variable in $\mathbb{R}$. If it exists, the **expectation of** X is

$$E[X] := \int_\Omega X(w) d\mathbb{P}(\omega) = \int_\mathbb{R} x d\mu_X(x),$$

where the second equality follows from $d\mu_X$ being a pushforward measure.

Definition 17. The **variance** of X (if $\mathbb{E}[X]$ exists) is

$$\mathrm{Var}(X) := \mathbb{E}[(X - \mathbb{E}[X])^2].$$

Equivalently, it is the L^2 norm of $X - \mathbb{E}[X]$.

Remark 18. It is possible that the expectation exists (e.g., $f \in L^1$) but the variance does not (e.g., $f \notin L^2$).

Example 19 (Gaussian)**.** The **Gaussian distribution** with mean $a \in \mathbb{R}$ and variance $\sigma^2 > 0$ is the distribution on $\mathbb{R}$ with density

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}.$$

In this case, we write $X \sim \mathcal{N}(a, \sigma^2)$. We call $\mathcal{N}(0,1)$ the **standard Gaussian**.

As we will see later, the Gaussian distribution is a sort of "canonical" distribution.
The following lemma answers how to find the expectation of $g(X)$ for some measurable g:

Lemma 20 (Law of the Unconscious Statistician)**.** *Let $X : \Omega \to \mathcal{X}$ be a random variable with distribution μ_X and $g : \mathcal{X} \to \mathbb{R}$ a measurable function, then*

$$\mathbb{E}[g(X)] = \int_{\mathcal{X}} g(x) d\mu_X(x).$$

Remark 21. This is essentially just a change of variables formula.

Proof. Assume $g = \sum_{j=1}^{N} a_j \mathbb{1}_{A_j}$ where $A_j \subset \mathcal{X}$ and $a_j \in \mathbb{R}$. Then,

$$\mathbb{E}[g(X)] = \sum_{j=1}^{N} a_j \mathbb{P}[X \in A_j] = \int_{\mathcal{X}} g(y) d\mu_x(y).$$

Lift this by approximating any function via simple functions. □

Definition 22 (σ-algebra of a random variable)**.** Let $X : \Omega \to \mathcal{X}$ be a random variable. The **σ-algebra generated by** X, denoted $\sigma(X)$, is

$$\sigma(X) := \{\{X \in A\} : A \subset \mathcal{X} \text{ Borel}\}.$$

We can observe that this is the pullback of the Borel σ-algebra and inherits being a σ-algebra because of it. Furthermore, $\sigma(X) \subset \mathcal{F}$ because X is $\mathcal{F}$-measurable.

A good way to interpret $\sigma(X)$ is that it is the set of events E that we can differentiable between upon observing X.

Example 23 (Indicator function)**.** Let $E \in \mathcal{F}$. Then,

$$\sigma(\mathbb{1}_E) = \{\varnothing, E, E^C, \Omega\}.$$

Example 24 (Projection)**.** Let $\Omega = \mathbb{R}^2$ and $X(x,y) = x$. Then,

$$\sigma(X) = \{\{A \times \mathbb{R}\} : A \subset \mathbb{R} \text{ Borel}\}.$$

Proposition 25. *Let $g : \mathcal{X} \to \mathcal{Y}$ be a measurable function. Then, $\sigma(g(X)) \subset \sigma(X)$.*

Remark 26. We can think of this as a sort of "information-processing inequality" since we *lose* information by putting it through a function.

Proof. We have that

$$\{g(X) \in B\} = \{X \in g^{-1}(B)\} \in \sigma(X).$$

$\square$

Proposition 27. *Conversely, if X and Y are random variables and Y takes values in $\mathbb{R}$ with $\sigma(Y) \subset \sigma(X)$ (or equivalently, Y is $\sigma(X)$ measurable), then there is a measurable function $g : \mathcal{X} \to \mathbb{R}$ such that $Y = g(X)$ almost surely.*

Proof. Consider the case where Y is a simple function. Then, clearly $Y = \sum_{j=1}^n a_j \mathbb{1}_{X \in A_j}$, which follows because Y is $\sigma(X)$-measurable. More generally, approximate Y by a sequence of random variables Y_n that are each simple and $Y_n \to Y$ almost surely. We can express $Y_n = g_n(X)$ for some functions g_n. Define $g(x) = \liminf_{n\to\infty} g_n(x)$, so $g(X) = Y$ almost surely. $\square$

Independence

Definition 28. Two events E, F are **independent** if $\mathbb{P}[E \cap F] = \mathbb{P}[E]\mathbb{P}[F]$. More generally, a collection of events $\mathcal{E} \subset \mathcal{F}$ is **independent** if for any $E_1, \ldots, E_n \in \mathcal{E}$ distinct,

$$\mathbb{P}[E_1 \cap \ldots E_n] = \prod_{j=1}^n \mathbb{P}[E_j].$$

If $\mathbb{P}[F] \neq 0$, we can define the conditional probability $\mathbb{P}[E|F] = \mathbb{P}[E \cap F]/\mathbb{P}[F]$. In this case, E and F are independent if and only if $\mathbb{P}[E|F] = \mathbb{P}[E]$, which tells us that knowing whether F occurs doesn't tell us anything about if E will occur.

Definition 29. A finite collection of σ-algebras $\mathcal{F}_1, \ldots, \mathcal{F}_n \subset \mathcal{F}$ is **independent** if for any choice of $E_j \in \mathcal{F}_j$ for $j \in \{1, \ldots, n\}$,

$$\mathbb{P}[E_1 \cap \ldots E_n] = \prod_{j=1}^n \mathbb{P}[E_j].$$

Definition 30. A finite collection of random variables $X_1, \ldots, X_n$ is **independent** if $\sigma(X_1), \ldots, \sigma(X_n)$ is independent.

Remark 31 (Functions of independent random variables are independent). If $X_1, \ldots, X_n$ are independent and $g_j : \mathcal{X}_j \to \mathcal{Y}_j$ are measurable functions, then $g_1(X_1), \ldots, g_n(X_n)$ are independent.

Lemma 32. *Let $X_1, \ldots, X_n$ be random variables taking values in $\mathcal{X}_1, \ldots, \mathcal{X}_n$. The following are equivalent:*

1. *$X_1, \ldots, X_n$ are independent (as above).*
2. *For all events $A_j \in \mathcal{X}_j$ for $j \in \{1, \ldots, n\}$, we have*

$$\mathbb{P}[X_1 \in A_1, \ldots, X_n \in A_n] = \prod_{j=1}^n \mathbb{P}[X_j \in A_j].$$

3. *Let* $\overline{X} = (X_1, \ldots, X_n)$ *be a variable in* $\mathcal{X}_1 \times \cdots \times \mathcal{X}_n$. *Then,*

$$\mu_{\overline{X}} = \mu_{X_1} \times \cdots \times \mu_{X_n}$$

as the product measure.

Proof. It is clear from definitions that condition 1 is equivalent to condition 2.

To get from condition 2 to condition 3, observe that a measure is sufficiently defined on rectangles.

To get from condition 3 to condition 2, let $A_j \subset \mathcal{X}_j$ be Borel for $j \in [n]$. If we know that $\mu_{\bar{X}}$ is the product measure, then we know that

$$\mu_{\bar{X}}(A_1 \times \cdots \times A_n) = \prod_{j=1}^{n} \mu_{X_j}(A_j).$$

□

Non-example 33 (Pairwise independence does not imply independence)**.** Let X_1, X_2 be random variables taking values in $\mathbb{Z}/2\mathbb{Z}$ with the uniform distribution. Let $X_3 = X_1 + X_2$ as an element of the group. Then, by definition (X_1, X_2) are independent. One also has that (X_2, X_3) and (X_1, X_3) are independent (by computation). But clearly, (X_1, X_2, X_3) is not independent since X_3 is a function of X_1, X_2.

An easy argument to show that (X_1, X_2, X_3) is not independent is to observe that while there are 8 elements of $(\mathbb{Z}/2\mathbb{Z})^3$, only four of them are "obtainable".

Example 34 (Random variables in $\mathbb{R}$)**.** Let X, Y be random variables taking values in $\mathbb{R}$ and assume that (X, Y) has a density $f : \mathbb{R}^2 \to [0, \infty)$. Then, X, Y are independent iff there exists measurable functions $f_1, f_2 : \mathbb{R} \to [0, \infty)$ such that $f(x, y) = f_1(x) f_2(y)$, where f_1 is the density of X and f_2 is the density of Y. One can see

$$f(x, y) = f_1(x) f_2(y)$$

is equivalent to

$$\mu_{(X,Y)} = \mu_X \times \mu_Y.$$

Lemma 35. *If X and Y are independent real-valued random variables with $E[|X|] < \infty$ and $E[|Y|] < \infty$, then $E[XY] = E[X]E[Y]$.*

Proof. Follows from Fubini's:

$$\begin{aligned} E[XY] &= \int_{\mathbb{R}\times\mathbb{R}} xy \, d\mu_{(X,Y)}(x, y) \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} xy \, d\mu_X(x) \, d\mu_Y(y) \\ &= \int_{\mathbb{R}} y E[X] \, d\mu_Y(y) \\ &= E[X]E[Y]. \end{aligned}$$

□

Lemma 36. *Let $X_1, \ldots, X_n$ be pairwise independent random variables in $\mathbb{R}$, and assume $\mathbb{E}[|X_j|] < \infty$. Then,*

$$\operatorname{Var}(X_1 + \cdots + X_n) = \sum_{j=1}^{n} \operatorname{Var}(X_j).$$

Proof. bleh $\square$

However, this is not true without independence:

Non-example 37. Let X be a random variable with bounded variance. Then,

$$\operatorname{Var}(X + X) = \operatorname{Var}(2X) = 4\operatorname{Var}(X).$$

$$\operatorname{Var}(X - X) = 0.$$

Definition 38. A sequence of random variables $\{X_j\}_{j\geq 1}$ is **identical and independently distributed (i.i.d.)** if for all n, $X_1, \ldots, X_n$ are independent and $\mu_{X_j} = \mu_{X_i}$ for all $i, j \in \{1, \ldots, n\}$.

We have some results on existence of i.i.d. variables.

Theorem 39. *Let μ be a probability measure on $\mathbb{R}$. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be $[0,1]$ equipped with the Borel σ-algebra and the uniform measure. Then, there exists a sequence of i.i.d. random variables on $(\Omega, \mathcal{F}, \mathbb{P})$ with distribution μ.*

Remark 40. This theorem may at first look to be an obvious result, as one can appear to construct independent random variables through a function $\Omega^\infty \to [0,1]$. But the point of this theorem is that each of the i.i.d. variables takes domain in the *same* probability space.

Proof. Relegated to problem set 5. $\square$

More generally,

Theorem 41. *Let μ be the Borel probability measure on a topological space $\mathcal{X}$. Then, there exists $(\Omega, \mathcal{F}, \mathbb{P})$ on which is defined a sequence of i.i.d. random variables with distribution μ.*

Proof. "Proof by abstract nonsense." $\square$

Convergence of random variables

Definition 42. Let $\{X_n\}_{n\geq 1}$ be random variables in a metric space $(\mathcal{X}, d)$, and let X be another random variable in $(\mathcal{X}, d)$. We say that

- X_n **converges to** X **almost surely** if

$$\mathbb{P}\left[\lim_{n\to\infty} X_n = X\right] = 1.$$

- X_n **converges to** X **in probability** if for every $\epsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}\left[d(X_n, X) < \epsilon\right] = 1.$$

- X_n **converges to** X **in distribution** if $\mu_{X_n} \to \mu_X$ weakly (i.e., for every bounded continuous function $f : \mathcal{X} \to \mathbb{R}$, we have $\mathbb{E}[f(X_n)] \to \mathbb{E}[f(X)]$).

Remark 43. The measure-theory analogs are that i) almost sure convergence is convergence pointwise almost everywhere ii) convergence in probability is convergence in measure iii) convergence in distribution is weak convergence in measure.

Remark 44. We have

$$\text{converge a.s} \implies \text{converge in probability} \implies \text{converge in distribution.}$$

Remark 45. Convergence in distribution is different from almost sure convergence and convergence in probability because a set of random variables is not required to be defined on the same probability space.

Remark 46. None of these notions of convergence is enough to imply

$$\mathbb{E}[X_n] \to \mathbb{E}[X].$$

For these, we would need stronger assumptions to deal with integration (e.g., dominated convergence).

Example 47 (Convergence in probability but not a.s.)**.** Let X_n be independent random variables such that $\mathbb{P}[X_n = 1] = 1/n$ and $\mathbb{P}[X_n = 0] = 1 - 1/n$. Let $X = 0$. Then, X_n converges to X in probability because

$$\mathbb{P}[d(X_n, X) < \epsilon] \leq 1 - \frac{1}{n}.$$

However, X_n fails to converge to X almost surely:

$$\begin{aligned}
\mathbb{P}[\lim_{n\to\infty} X_n = X] &= \mathbb{P}[\exists N \forall n \geq N, X_n = 0] \\
&\leq \sum_{N=0}^{\infty} \mathbb{P}[\forall n \geq N, X_n = 0] \\
&= \sum_{N=0}^{\infty} \prod_{n=N}^{\infty} \left(1 - \frac{1}{n}\right)^n \\
&= 0.
\end{aligned}$$

Example 48 (Convergence in distribution but not probability)**.** Let $\{X_n\}_{n\in\mathbb{N}}$ be a sequence of independent random variables, each distributed according to some μ, which is not a point mass. Trivially, $X_n \to X_1$ in distribution. However, X_n fails to converge to X_1 in probability or almost surely.

Lemma 49. *Let $\{X_n\}_{n\geq 1}$ and X be random variables taking values in $\mathbb{R}$. The following are equivalent:*

1. *$X_n \to X$ in distribution.*

2. *$\mathbb{E}[f(X_n)] \to \mathbb{E}[f(X)]$ for all smooth, compactly supported functions f.*

3. *$\mathbb{P}[X_n \in [a, b]] \to \mathbb{P}[X \in [a, b]]$ for all $a < b$ with $\mathbb{P}[X \in \{a, b\}] = 0$.*

Lemma 50. *Let $\{X_n\}_{n\geq 1}$ and X be random variables taking values in $\mathbb{R}$. The following are equivalent:*

1. *$X_n \to X$ in distribution.*

2. $\mathbb{E}[f(X_n)] \to \mathbb{E}[f(X)]$ *for all smooth, compactly supported functions* f.

3. $\mathbb{P}[X_n \in [a,b]] \to \mathbb{P}[X \in [a,b]]$ *for all* $a < b$ *with* $\mathbb{P}[X \in \{a,b\}] = 0$.

Proof. It is clear from definitions that condition 1 implies condition 2, as smooth compactly supported functions are in particular bounded and continuous.

To get from condition 2 to condition 3, let $a < b$ and $\varepsilon > 0$. Let $g_\varepsilon : \mathbb{R} \to [0,1]$ be a smooth function such that $g_\varepsilon \equiv 1$ on $[a+\varepsilon, b-\varepsilon]$ and $g_\varepsilon \equiv 0$ on $\mathbb{R} \setminus [a,b]$. Observe that for any random variable Y, we have

$$\mathbb{P}[Y \in [a+\varepsilon, b-\varepsilon]] \le \mathbb{E}[g_\varepsilon(Y)] \le \mathbb{P}[Y \in [a,b]].$$

If we know that $\mathbb{P}[X \in \{a,b\}] = 0$, then upper continuity of the measure yields

$$\mathbb{P}[X \in [a+\varepsilon, b-\varepsilon]] = \mathbb{P}[X \in [a,b]] - o(1),$$

where $o(1) \to 0$ as $\varepsilon \to 0$. Thus,

$$\begin{aligned}
\liminf_{n\to\infty} \mathbb{P}[X_n \in [a,b]] &\ge \liminf_{n\to\infty} \mathbb{E}[g_\varepsilon(X_n)] \\
&\ge \mathbb{E}[g_\varepsilon(X)] \\
&\ge \mathbb{P}[X \in [a+\varepsilon, b-\varepsilon]] \\
&= \mathbb{P}[X \in [a,b]] - o(1).
\end{aligned}$$

Taking $\varepsilon \to 0$ yields $\liminf_{n\to\infty} \mathbb{P}[X_n \in [a,b]] \ge \mathbb{P}[X \in [a,b]]$. A similar approximation by a function with support slightly larger than the interval shows that $\limsup_{n\to\infty} \mathbb{P}[X_n \in [a,b]] \le \mathbb{P}[X \in [a,b]]$.

To get from condition 3 to condition 1, observe that a bounded continuous function can be approximated by finite linear combinations of indicators of closed intervals. Since μ_X and μ_{X_n} are finite measures, their mass is concentrated in some compact set, so the behavior does not matter. □

Law of large numbers

Lemma 51 (Markov's inequality)**.** *Let* X *be a random variable in* $[0,\infty)$*. Then, for any* $R > 0$*,*

$$\mathbb{P}[X > R] \le \frac{\mathbb{E}[X]}{R}.$$

Proof. Follows from the observation that

$$X \ge R\mathbb{1}_{X>R}.$$

□

Remark 52. It is very easy to get a generalization of Markov's via:

$$\mathbb{P}[X > R] = \mathbb{P}[X^p > R^p] \le \frac{\mathbb{E}[X^P]}{R}.$$

Theorem 53 (Weak law of large numbers)**.** *Let $\{X_j\}_{j\geq 1}$ be iid RVs in $\mathbb{R}$ with finite mean and variance. Then*

$$\frac{1}{n}\sum_{j=1}^{n} X_j \to \mathbb{E}[X_1]$$

in probability.

Proof. Let $a = \mathbb{E}[X_1]$. Then,

$$\mathbb{E}\left[\frac{1}{n}\sum_{j=1}^{n} X_j\right] = a.$$

$$\mathbb{E}\left[\left(\frac{1}{n}\sum_{j=1}^{n} X_j - a\right)^2\right] = \operatorname{Var}\left(\frac{1}{n}\sum_{j=1}^{n} X_j\right) = \frac{1}{n^2}\operatorname{Var}\left(\sum_{j=1}^{n} X_j\right) = \frac{1}{n^2}\sum_{j=1}^{n}\operatorname{Var}(X_1) = \frac{\operatorname{Var}(X_1)}{n}$$

Note that this is going to 0 as n large.

Use Markov's inequality:

$$\mathbb{P}\left[|\frac{1}{n}\sum_{j=1}^{n} X_{j-a}| > \epsilon\right] = \mathbb{P}\left[|\frac{1}{n}\sum_{j=1}^{n} X_{j-a}|^2 > \epsilon^2\right] = \frac{\operatorname{Var}(X_1)}{n\epsilon^2} \to 0.$$

□

Lemma 54 (Borel-Cantelli lemma)**.** *Let $\{E_j\}_{j\geq 1}$ be events. If $\sum_{j=1}^{\infty}\mathbb{P}\left[E_j\right] < \infty$, then, $\mathbb{P}\left[\text{infinitely many } E_j \text{ occur}\right] = 0$.*

Proof. Observe that

$$\{\text{infinitely many } E_j \text{ occur}\} = \bigcap_{n=1}^{\infty}\bigcup_{j=n}^{\infty} E_j.$$

Let $G_n = \bigcup_{j=n}^{\infty} E_j$. Then,

$$\mathbb{P}\left[\bigcap_{n=1}^{\infty} G_n\right] \leq \mathbb{P}\left[G_n\right] \leq \sum_{j=1}^{\infty}\mathbb{P}\left[E_j\right]$$

by the union bound, and this goes to zero as $n \to \infty$. □

Theorem 55 (Strong law of large numbers)**.** *Let $\{X_j\}_{j\geq 1}$ be independent, not necessarily i.i.d. random variables taking values in $\mathbb{R}$, which all have the same mean $a \in \mathbb{R}$. Assume there exists $C > 0$ such that $\mathbb{E}[X_j^4] \leq C$ for all j. Then, $\frac{1}{n}\sum_{j=1}^{n} X_j \to a$ almost surely.*

Remark 56. This is not the strongest statement available. The statement that is more common is: If $\{X_j\}_{j\geq 1}$ are i.i.d. and $\mathbb{E}[|X_1|] < \infty$ then $\frac{1}{n}\sum_{j=1}^{n} X_j \to a$ almost surely. See Theorem 2.4.1 [Dur19].

Proof. Assume WLOG $a = 0$.

Claim: exists $C_1 > 0$ such that

$$\mathbb{E}\left[\left(\sum_{j=1}^{n} X_j\right)^4\right] \leq C_1 n^2.$$

expand...

$$\mathbb{E}\left[\left(\sum_{j=1}^n X_j\right)^4\right] = \sum_{i,j,k,l=1}^n \mathbb{E}[X_iX_jX_kX_l]$$

Since X_j's are independent with mean zero, any term where one of the factors only occurs once is going to be zero. So there are only two types of terms: $\mathbb{E}[X_j^4]$ and $\mathbb{E}[X_i^2]\mathbb{E}[X_j^2]$. The first is dealt with by assumption because there are n terms of the first kind. For the second term, by Cauchy-Schwartz/Jensen's,

$$\mathbb{E}[X_j^2] \leq \mathbb{E}[X_j^4]^{1/2} \leq C^{1/2},$$

and there are less than n^2 terms of the second type.

We want to use Markov's to convert from expectations to probabilites, and Borel-Cantelli to convert to an almost surely statement.

Let

$$E_n = \left\{|\frac{1}{n}\sum_{j=1}^n X_j| > n^{-1/8}\right\} = \left\{|\sum_{j=1}^n X_j| > n^{7/8}\right\}$$

By Markov's:

$$\mathbb{P}\left[E_n\right] \leq \mathbb{P}\left[|\sum_{j=1}^n X_j|^4 > n^{7/2}\right] \leq Cn^2 \cdot n^{-7/2} = cn^{-3/2},$$

so this is summable because $-3/2$ is smaller than 1. By Borel-Cantelli lemma, almost surely as only finitely many of the E_n's occur. This means that for all but finitely many n, we have that $|\frac{1}{n}\sum_{j=1}^n X_j| \leq n^{-1/8} \to 0$ which implies almost surely convergence to zero. □

Fourier analysis and characteristic functions

Definition 57. Let X be a random variable taking values in $\mathbb{R}$. The **characteristic function** is $\varphi_X : \mathbb{R} \to \mathbb{C}$ defined by

$$\varphi_X(t) = \mathbb{E}[e^{itX}] = \int_{\mathbb{R}} e^{ity}\, d\mu_X(y).$$

Remark 58. Equivalently, this is just the Fourier transform of μ_X.

We describe some properties of the characteristic function.

Proposition 59 (Continuity of characteristic function)**.** *The characteristic function is always a continuous function of t.*

Proof. Apply dominated convergence theorem since $|e^{it_nX} - e^{itX}| \leq 2$. □

Proposition 60 (Characteristic function of sum of independents)**.** *If $X_1, \ldots, X_n$ are independent random variables, then $\varphi_{X_1+\cdots+X_n}(t) = \mathbb{E}[e^{it(X_1+\cdots+X_n)}] = \prod_{j=1}^n \mathbb{E}[e^{itX_j}] = \prod_{j=1}^n \varphi_{X_j}(t)$.*

Remark 61. This is an analogous statement to convolution of distributions. The converse is also true (and shown later).

Proposition 62 (Characteristic function of Gaussian)**.** *If $X \sim \mathcal{N}(a, \sigma^2)$, then*

$$\varphi_X = e^{iat - \sigma^2t^2/2}.$$

Proof. Done on problem set 2 via complex analysis. □

Lemma 63 (Moments of a random variable can be expressed in terms of the characteristic function)**.** *Let $N \in \mathbb{N}$ be a positive integer such that $\mathbb{E}[|X|^n] < \infty$. Then, $\varphi_X^{(n)}(0)$ exists and*

$$\varphi_X^{(n)}(0) = i^n \mathbb{E}[X^n].$$

Proof. For all $x \in \mathbb{R}$, we have that

$$\left| e^{ix} - \sum_{k=0}^{n} \frac{(ix)^k}{k!} \right| \leq \min \left\{ \frac{|x|^{n+1}}{(n+1)!}, \frac{2|x|^n}{n!} \right\}$$

from calculus. Set $x = tX$ and take an expectation, so

$$\left| \varphi_X(t) - \sum_{k=0}^{n} \frac{i^k \mathbb{E}[X^k]}{k!} t^k \right| \leq |t|^n \mathbb{E} \left[\min \left\{ \frac{|t||X|^{n+1}}{(n+1)!}, \frac{2|X|^n}{n!} \right\} \right]$$

The RHS goes to zero as $t \to 0$, and it's bounded by $\frac{2|X|^n}{n!}$. By dominated convergence, the right side is $o(t^n)$ as $t \to 0$. So this implies that

$$\sum_{k=0}^{n} \frac{i^k \mathbb{E}[X^k]}{k!} t^k$$

is an n-th order Taylor polynomial for $\varphi_X(t)$, with corresponding derivatives. □

Proposition 64 (Almost everywhere convergence of φ_X implies convergence in distribution)**.** *Let $\{X_j\}$ and X be random variables in $\mathbb{R}$. Assume $\varphi_{X_j} \to \varphi_X$ Lebesgue almost everywhere. Then, $X_j \to X$ in distribution.*

Remark 65. This is an if and only if statement: the converse is automatically true by dominated convergence theorem since e^{itx} is bounded in x.

Proof. Let $g : \mathbb{R} \to \mathbb{R}$ be a smooth function with compact support. To prove convergence in distribution, it suffices to show that $\mathbb{E}[g(X_j)] \to \mathbb{E}[g(X)]$.

Consider the Fourier transform of g:

$$\hat{g}(y) = \int_{\mathbb{R}} e^{-ixy} g(x)\, dx$$

Because g is smooth and has compact support, its Fourier transform $\hat{g}$ is a Schwartz function (i.e., the function is smooth and all of its derivatives decay faster than any polynomial).

By the Fourier inversion formula:

$$g(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{ixy} \hat{g}(y)\, dy$$

Substituting the random variable X_j and taking the expectation yields:

$$\mathbb{E}[g(X_j)] = \frac{1}{2\pi} \mathbb{E} \left[\int_{\mathbb{R}} e^{iX_j y} \hat{g}(y)\, dy \right]$$

Since $\hat{g} \in L^1(\mathbb{R})$, this integral is absolutely convergent. By Fubini's Theorem, we can swap the expectation and the integral:

$$\mathbb{E}[g(X_j)] = \frac{1}{2\pi}\int_{\mathbb{R}} \mathbb{E}[e^{iX_j y}]\hat{g}(y)\,dy = \frac{1}{2\pi}\int_{\mathbb{R}} \varphi_{X_j}(y)\hat{g}(y)\,dy$$

Similarly, for the limiting random variable X:

$$\mathbb{E}[g(X)] = \frac{1}{2\pi}\int_{\mathbb{R}} \varphi_X(y)\hat{g}(y)\,dy$$

Recall that $\varphi_{X_j} \to \varphi_X$ pointwise almost everywhere. Furthermore, characteristic functions are uniformly bounded by 1, meaning $|\varphi_{X_j}(y)\hat{g}(y)| \le |\hat{g}(y)|$. Since $\hat{g}$ acts as an integrable dominating function, we apply the Dominated Convergence Theorem to conclude that:

$$\lim_{j\to\infty} \mathbb{E}[g(X_j)] = \mathbb{E}[g(X)].$$

□

Corollary 66 (Uniqueness of characteristic function)**.** *Let X, Y be random variables with $\varphi_X = \varphi_Y$. Then, X equals Y in distribution.*

Central limit theorem

Theorem 67 (Central limit theorem)**.** *Let $\{X_j\}$ be i.i.d. random variables with $\mathbb{E}[X_j] = 0$ and $\mathrm{Var}(X_j) = 1$. Then,*

$$Z_n := \frac{1}{\sqrt{n}}\sum_{j=1}^{n} X_j \to \mathcal{N}(0,1)$$

in distribution.

Remark 68. The law of large numbers implies that

$$\frac{1}{n}\sum_{j=1}^{n} X_j \to 0$$

almost surely in this setting. Thus, the central limit theorem gives more refined information than the law of large numbers.

Remark 69. The central limit theorem tells us that the Gaussian is universal in some sense.

Proof. Let $Z \sim \mathcal{N}(0,1)$, so $\varphi_Z(t) = e^{-t^2/2}$. It suffices to show that $\varphi_{Z_n}(t)$ converges to $e^{-t^2/2}$ pointwise. Let $\varphi = \varphi_{X_1} = \varphi_{X_j}$. Then,

$$\varphi_{Z_n}(t) = \left[\varphi_{\frac{X_1}{\sqrt{n}}}(t)\right]^n = [\mathbb{E}[e^{itX^1/\sqrt{n}}]]^n = \left[\varphi\left(\frac{t}{\sqrt{n}}\right)\right]^n$$

We have that $\varphi(0) = 1$, $\varphi^{(1)}(0) = i\mathbb{E}[X_1] = 0$ and $\varphi^{(2)}(0) = -\mathbb{E}[X_1^2] = -1$. So we see that

$$\varphi\left(\frac{t}{\sqrt{n}}\right) = 1 - \frac{t^2}{2n} + o\left(\frac{1}{n}\right) \qquad n \to \infty$$

Let Log be the complex branch with ImLog $\in [-\pi, \pi)$. We are interested in values around 1, so this is OK. Then,

$$\operatorname{Log} \varphi_{Z_n}(t) = n \operatorname{Log}\left(1 - \frac{t^2}{2n} + o\left(\frac{1}{n}\right)\right)$$

Taylor expand the Log to find that

$$\operatorname{Log}(\varphi_{Z_n}(t)) = n\left[-\frac{t^2}{2n} + o\left(\frac{1}{n}\right)\right] \to -\frac{t^2}{2}.$$

$\square$

Conditional expectation

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let $\mathcal{G} \subset \mathcal{F}$ be a sub-σ-algebra. Throughout this section, let $\mathbb{E}[|X|] < \infty$.

Definition 70. Let $\mathcal{G} \subset \mathcal{F}$ be a σ-algebra and X a random variable in $\mathbb{R}$. Then, the **conditional expectation of X with respect to $\mathcal{G}$** is the unique (a.s.) random variable $\mathbb{E}[X|\mathcal{G}]$ that satisfies:

1. $\mathbb{E}[X \mid \mathcal{G}]$ is $\mathcal{G}$-measurable
2. $\mathbb{E}[\mathbb{1}_G \mathbb{E}[X|\mathcal{G}]] = \mathbb{E}[\mathbb{1}_G X]$ for all $G \in \mathcal{G}$.

There are a couple of peculiarities with this definition that don't appear to agree with our "elementary" notion of conditioning a variable on an *event* instead of a σ-algebra. My interpretation of this seems to be slightly different from Gwynne's because I wanted to think about what each $\mathbb{E}[X|\mathcal{G}](\omega)$ means: suppose that we only have access to a YES/NO oracle for each of the events in $\mathcal{G}$. Then, $\mathbb{E}[X|\mathcal{G}](\omega)$ is our best "guess" for X given that ω is any element that satisfies all of our YES/NO queries. Here, treat ω as a "representative" of the queries that we are able to answer, as these will be equivalent up to the query answers due to the requirement that $\mathbb{E}[X|\mathcal{G}]$ is $\mathcal{G}$-measurable.

Here are some basic properties, and their interpretations in terms of information. Their proofs follow essentially from definition.

Proposition 71 (Properties of conditional expectation)**.**

1. *If X is $\mathcal{G}$-measurable, then $\mathbb{E}[X \mid \mathcal{G}] = X$.*
2. *If $\sigma(X)$ is independent from $\mathcal{G}$, then $\mathbb{E}[X \mid \mathcal{G}] = \mathbb{E}[X]$.*
3. *(Linearity) For all $a, b \in \mathbb{R}$, then*

$$\mathbb{E}[aX + bY \mid \mathcal{G}] = a\mathbb{E}[X \mid \mathcal{G}] + b\mathbb{E}[Y \mid \mathcal{G}].$$

4. *(Tower property) If $\mathcal{H} \subset \mathcal{G} \subset \mathcal{F}$ are σ-algebras, then*

$$\mathbb{E}[\mathbb{E}[X \mid \mathcal{G}] \mid \mathcal{H}] = \mathbb{E}[X \mid \mathcal{H}]$$

We can interpret the statements as follows:

1. If X is $\mathcal{G}$-measurable, then X is determined by the YES/NO queries to $\mathcal{G}$, so "knowing" $\mathcal{G}$ determines X exactly.
2. If $\sigma(X)$ is independent from $\mathcal{G}$, then the knowing events of $\mathcal{G}$ will not help us determine X anyways.
3. Analog of linearity of expectation.
4. If $\mathcal{H}$ is "coarser" than $\mathcal{G}$, then when we condition on $\mathcal{G}$ and then condition on $\mathcal{H}$ we will end up reaching the same information as if we had originally conditioned on $\mathcal{H}$.

Definition 72. The **conditional expectation of** X **given** Y is

$$\mathbb{E}[X \mid Y] := \mathbb{E}[X \mid \sigma(Y)].$$

To make these definitions more concrete, we can try to "retrieve" the intuitive definition of a conditional expectation.

Proposition 73. *Suppose Y takes values in a countable set A. Then*[1]*,*

$$\mathbb{E}[X \mid Y] = \sum_{a \in A} \frac{\mathbb{E}[X \mathbb{1}_{Y=a}]}{\mathbb{P}\left[Y = a\right]} \mathbb{1}_{Y=a}.$$

[1] One apparent issue with this might be when $\mathbb{P}[Y = a] = 0$. However, we don't really care what happens on measure-zero set here, anyways.

Proof. Let $Z := \sum_{a \in A} \frac{\mathbb{E}[X \mathbb{1}_{Y=a}]}{\mathbb{P}[Y=a]} \mathbb{1}_{Y=a}$. First, observe that Z is $\sigma(Y)$-measurable because it is a function of Y. We verify that Z satisfies the defining property of $\mathbb{E}[X \mid Y]$. Every event in $\sigma(Y)$ takes the form $Y \in B \subset \mathcal{Y}$ for some B. Since expectation is linear, we can assume without loss of generality that the event is of form $Y = b$. Thus,

$$\begin{aligned}
\mathbb{E}[Z \mathbb{1}_{Y=b}] &= \mathbb{E}\left[\sum_{a \in A} \frac{\mathbb{E}[X \mathbb{1}_{Y=a}]}{\mathbb{P}\left[Y = a\right]} \mathbb{1}_{Y=a} \mathbb{1}_{Y=b}\right] \\
&= \sum_{a \in A} \frac{\mathbb{E}[X \mathbb{1}_{Y=a}]}{\mathbb{P}\left[Y = a\right]} \mathbb{E}[\mathbb{1}_{Y=a} \mathbb{1}_{Y=b}] \\
&= \frac{\mathbb{E}[X \mathbb{1}_{Y=b}]}{\mathbb{P}\left[Y = b\right]} \mathbb{E}[\mathbb{1}_{Y=b}] \\
&= \mathbb{E}[X \mathbb{1}_{Y=b}].
\end{aligned}$$

□

This interpretation is somewhat intuitive: when $Y = a$, then we know that $\mathbb{E}[X \mid Y] = \frac{\mathbb{E}[X \mathbb{1}_{Y=a}]}{\mathbb{P}[Y=a]}$. We can specialize further:

Corollary 74. *If $X = \mathbb{1}_E$ and $Y = \mathbb{1}_F$ and $\mathbb{P}(F) > 0$, then*

$$\mathbb{E}[X \mid Y] = \frac{\mathbb{P}\left[E \cap F\right]}{\mathbb{P}[F]}\mathbb{1}_F + \frac{\mathbb{P}\left[E \cap F^C\right]}{\mathbb{P}[F^C]}\mathbb{1}_{F^C}.$$

This retrieves our usual notion of conditional expectation.

We will now extend the idea of conditioning onto a random variable instead:

Definition 75. Let $\mathcal{G} \subset \mathcal{F}$ be a σ-algebra and $E \in \mathcal{F}$ is an event. We define

$$\mathbb{P}\left[E|\mathcal{G}\right] = \mathbb{E}[\mathbb{1}_E|\mathcal{G}].$$

Definition 76. Let $X : \Omega \to \mathcal{X}$ be a random variable. A $\mathcal{G}$-measurable random variable M taking values in $\{$Borel probability measures on $\mathcal{X}\}$ is called the **conditional distribution** of X given $\mathcal{G}$ if for each $A \subset \mathcal{X}$ Borel almost surely, the $\mathbb{P}\left[X \in A|\mathcal{G}\right] = M(A)$.

More concretely, $M(\omega)$ is a measure on $\mathcal{X}$, and $M(A) : \omega \to M(\omega)(A)$ is a random variable taking values in $[0, 1]$.

Theorem 77. *If $\mathcal{X}$ is metrizable with a separable metric, then for any σ-algebra $\mathcal{G} \subset \mathcal{F}$ and any random variable taking values in $\mathcal{X}$, the conditional distribution of $\mathcal{X}$ given $\mathcal{G}$ exists.*

Proof. Abstract nonsense. See Section 5.1.3 of [Dur19]. □

Remark 78. In math, we typically want to decompose objects into ones that are easier to understand. For instance, composition $f = g \circ h$ as functions to study f; quotient groups $G \setminus H$ to study G. Conditioning is the analogous thing to do in probability: given a random variable X, may be sensible to find another random variable Y such that the distribution of Y and the conditional distribution of X given Y are easier to understand.

Example 79. If X and $\mathcal{G}$ are independent, then conditional distribution of X given $\mathcal{G}$ is just the unconditional distribution μ_X.

Example 80. The converse is also true. If the conditional distribution equals the unconditional distribution, then for any $A \subset \mathcal{X}$ Borel, for all $G \in \mathcal{G}$, we have that

$$\mathbb{P}\left[X \in A, G\right] = \mathbb{E}[\mathbb{P}\left[X \in A|\mathcal{G}\right]\mathbb{1}_G] = \mathbb{E}[\mathbb{P}\left[X \in A\right]\mathbb{1}_G] = \mathbb{P}\left[X \in A\right]\mathbb{P}\left[G\right].$$

Example 81. Let X be $\mathcal{G}$-measurable. Then the conditional distribution of X given $\mathcal{G}$ is δ_X.

Example 82. Let $Y \sim \text{Unif}[0, 1]$. Let $X \sim \mathcal{N}(0, Y)$.

$$\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X|Y]] = 0.$$

$$\text{Var}(X) = \mathbb{E}[X^2] = \mathbb{E}[\mathbb{E}[X^2|Y]] = \mathbb{E}[Y] = \frac{1}{2}.$$

Example 83 (Conditioning on countable set)**.** Let X be a random variable in $\mathcal{X}$, and Y a random variable taking values in $\mathcal{Y}$ which is countable and equipped with the discrete topology. For $y \in \mathcal{Y}$, we will define a probability measure M_y on $\mathcal{X}$ by

$$M_y(A) = \frac{\mathbb{P}\left[X \in A, Y = y\right]}{\mathbb{P}\left[Y = y\right]}.$$

Note that if $\mathbb{P}[Y = y] = 0$, we don't really care about $M_y(A)$ in this case anyways, so this is well-defined. We claim that M_Y is the conditional distribution of X given Y.

To establish this, we need to show that for $A \subset \mathcal{X}$ Borel, $M_Y(A) = \mathbb{P}\left[X \in A | Y\right]$. It is easy to check that $M_Y(A)$ is $\sigma(Y)$-measurable since it is a function of Y. Second, it is good enough to check for all $B \subset \mathcal{Y}$ Borel:

$$\mathbb{E}[\mathbb{1}_{Y \in B} M_Y(A)] = \mathbb{E}[\mathbb{1}_{Y \in B}\mathbb{1}_{X \in A}] = \mathbb{P}[Y \in B, X \in A].$$

We verify:

$$\mathbb{E}[\mathbb{1}_{(Y \in B)} M_Y(A)] = \sum_{y \in B} \mathbb{P}\left[Y = y\right] M_y(A) = \mathbb{P}\left[X \in A, Y \in B\right].$$

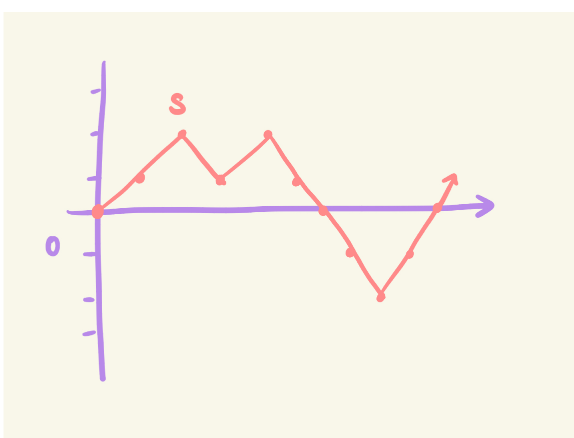
Brownian motion

Intuition and motivations

Brownian motion is motivated by the following question:

Question 84. *What is the most natural random, continuous function from $[0, \infty) \to \mathbb{R}$? More specifically, what is the most natural random variable in $C([0, \infty), \mathbb{R})$ with the topology of uniform convergence on compacts?*

To build up to this, we first consider a random "continuous"[2] $S : \mathbb{N}_0 \to \mathbb{Z}$. We can construct a sensible function as follows: let $\{\xi_j\}_{j \geq 1}$ be i.i.d. where $\mathbb{P}\left[\xi_j = -1\right] = \mathbb{P}\left[\xi_j = 1\right] = \frac{1}{2}$. Set $S_0 = 0$, and let $S_n = \sum_{j=1}^{n} \xi_j$ for $n \geq 1$. We call S_n a **random walk**. Next, extend S from $N_0 \to \mathbb{Z}$ to $[0, \infty) \to \mathbb{R}$ by piecewise linear interpolation in the following sense:



[2] Our notion of "continuous" here is that changing the input changes the output by at most one.

The main idea is that we can rescale S by "squashing" it down along the time axis to simulate a continuous function. Intuitively, we would like some $t \mapsto S_{nt}$ for some large n. However, this map doesn't converge as $n \to \infty$.

Fortunately, by the law of large numbers, we know that

$$\frac{1}{\sqrt{n}} S_n = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} \xi_n \xrightarrow{d} \mathcal{N}(0,1) \quad \text{as } n \to \infty,$$

suggesting that the right scaling is $n^{-1/2} S_{nt}$. Indeed, for any $t > 0$,

$$n^{-1/2} S_{nt} = \frac{(tn)^{1/2}}{n^{1/2}} (tn)^{-1/2} S_{nt} \xrightarrow{d} \mathcal{N}(0,t).$$

Two observations about this construction:

- For any $t > 0$, the function $S_{\cdot + tn} - S_{tn}$ is independent from $S|_{[0,tn]}$.
- For any $0 \leq t_0 < t_1 < \cdots < t_N$,

$$\frac{1}{\sqrt{n}} (S_{t_1 n} - S_{t_0 n}, S_{t_2 n} - S_{t_1 n}, \ldots, S_{t_N n} - S_{t_{N-1}}) \xrightarrow{d} (X_1, \ldots, X_n).$$

Definition

Definition 85. A **random continuous function** is a random variable taking values in $C([0,\infty), \mathbb{R})$ with the topology given by uniform convergence on compacts.

Definition 86. (Standard) Brownian motion is the random continuous function $B : [0,\infty) \to \mathbb{R}$, such that $B_0 = 0$, satisfying

1. For any $s < t$, $B_t - B_s \sim \mathcal{N}(0, t-s)$.
2. For any $0 \leq t_0 < \cdots < t_N$, $B_{t_j} - B_{t_{j-1}}$ for $j \in \{1, \ldots, N\}$ are independent.

We will construct and demonstrate existence of Brownian motion later.

Construction of Brownian motion

Because a Brownian motion $\mathcal{B}$ can be constructed from copies of $\mathcal{B}|_{[0,1]}$, we will just focus on constructing i.i.d. copies of $\mathcal{B}|_{[0,1]}$.

Broadly, our construction will follow two steps:

1. Construct $\mathcal{B}$ for dyadic times.
2. Extend $\mathcal{B}$ to some unique continuous function on $[0,1]$ by showing that it is uniformly continuous on the dyadics.

Define

$$\mathcal{D}_n := \left\{ \frac{j}{2^n} : j = 1, \ldots, 2^n \right\}$$

$$\mathcal{D} = \bigcup_{n=1}^{\infty} D_n.$$

We will let $\{Z_t : t \in \mathcal{D}\}$ be a collection of i.i.d. $\mathcal{N}(0,1)$ random variables. We will define $\mathcal{B}_t$ inductively by "filling in the gaps": let $B_0 := 0$ and $B_1 := Z_1$. For $n \geq 1$, each $t \in \mathcal{D}_n \setminus \mathcal{D}_{n-1}$, we define

$$\mathcal{B}_t := \frac{1}{2}(\mathcal{B}_{t-2^{-n}} + \mathcal{B}_{t+2^{-n}}) + 2^{-(n+1)/2} Z_t.$$

Observe that this makes sense because $t \pm 2^{-n} \in \mathcal{D}_{n-1}$.

Lemma 87. *For each $n \in \mathbb{N}_0$, the random variables $\mathcal{B}_{j/2^n} - \mathcal{B}_{(j-1)/2^n}$ for $j/2^n \in \mathcal{D}_n$ are i.i.d., and each is Gaussian with mean zero and variance 2^{-n}.*

Proof. Induct on n (note $n = 0$ is clear). Assume that $n \geq 1$ and that the statement holds for $n-1$.

Observe that if $j/2^n \in D_n \setminus D_{n-1}$, then

$$B_{j/2^n} - B_{(j-1)/2^n} = 2^{-(n+1)/2} Z_{j/2^n} + \frac{1}{2}\left(B_{(j+1)/2^n} - B_{(j-1)/2^n}\right),$$

and if $j/2^n \in D_{n-1}$, then

$$B_{j/2^n} - B_{(j-1)/2^n} = -2^{-(n+1)/2} Z_{(j-1)/2^n} + \frac{1}{2}\left(B_{j/2^n} - B_{(j-2)/2^n}\right).$$

Hence, $(B_{j/2^n} - B_{(j-1)/2^n})_{j=1}^{2^n}$ is the image of the vector

$$\left((B_{j/2^{n-1}} - B_{(j-1)/2^{n-1}})_{j=1}^{2^{n-1}}, (2^{-(n-1)/2} Z_t)_{t \in D_n \setminus D_{n-1}}\right)$$

under a linear transformation. If we know that the inductive hypothesis holds, then this latter vector consists of 2^n i.i.d. $N(0, 2^{-n+1})$ random variables.

It follows that each increment $B_{j/2^n} - B_{(j-1)/2^n}$ is equal to half the sum of two independent normal random variables with mean zero and variance 2^{-n+1}. By computation, each $B_{j/2^n} - B_{(j-1)/2^n}$ is Gaussian with mean zero and variance

$$\frac{1}{4}(2^{-n+1} + 2^{-n+1}) = 2^{-n}.$$

To show that the increments are independent, observe that for $i \neq j$, we have

$$\mathbb{E}[(B_{i/2^n} - B_{(i-1)/2^n})(B_{j/2^n} - B_{(j-1)/2^n})] = 0.$$

If $|i-j| > 1$, then the two increments are functions of disjoint pairs of independent random variables. Thus, they are independent, so the expectation of their product is the product of their expectations, which is zero.

If i and j are adjacent (i.e., $i = j+1$), then the above formulas give $B_{j/2^n} - B_{(j-1)/2^n} = X + Y$ and $B_{(j+1)/2^n} - B_{j/2^n} = -X + Y$ for two independent $N(0, 2^{-n-1})$ random variables X and Y. It follows that:

$$\begin{aligned}
\mathbb{E}[(B_{i/2^n} - B_{(i-1)/2^n})(B_{j/2^n} - B_{(j-1)/2^n})] &= \mathbb{E}[(X+Y)(-X+Y)] \\
&= -\mathbb{E}[X^2] + \mathbb{E}[Y^2] \\
&= -2^{-n-1} + 2^{-n-1}
\end{aligned}$$

$$= 0.$$

Because the increments are the image of a vector of 2^n i.i.d. Gaussian random variables under a linear transformation, they are jointly Gaussian. That this orthogonality relation implies independence follows from problem set 5. $\square$

Lemma 88.

1. *For all $s, t \in \mathcal{D}$ with $s < t$, it holds that $\mathcal{B}_t - \mathcal{B}_s \sim \mathcal{N}(0, t-s)$*
2. *For any $t_0 \leq t_1 < \cdots < t_n$ in $\mathcal{D}$, it holds that $B_{t_j} - B_{t_{j-1}}$ are independent.*

Proof. Fairly straightforward from just using the lemma shown above.

Let $n \in \mathbb{N}$ be large such that $s, t \in \mathcal{D}_n$. Then,

$$\mathcal{B}_t - \mathcal{B}_s = \sum_{j=2^n s+1}^{2^n t} B_{\frac{j}{2^n}} - B_{\frac{j-1}{2^n}},$$

which is the sum of $2^n(t-s)$ i.i.d. variables distributed according to $\mathcal{N}(0, 2^{-n})$. Thus, $\mathcal{B}_t - \mathcal{B}_s \sim \mathcal{N}(0, t-s)$.

Let $n \in \mathbb{N}$ be large enough such that $t_0, \ldots, t_k \in \mathcal{D}_n$. Then, $B_{t_j} - B_{t_{j-1}}$ are sums of disjoint subsets of $\left\{B_{\frac{i}{2^n}} - B_{\frac{i-1}{2^n}}\right\}_{i=1,\ldots,2^n}$ which are i.i.d., as we can view this as a measurable function of i.i.d. variables.

$\square$

To extend this into $\mathbb{R}$ via uniform continuity, we rely on the following theorem:

Theorem 89 (Kolmogorov continuity theorem)**.** *Let $X : \mathcal{D} \to \mathbb{R}$ be a random function. Assume there exists constants $C, \alpha, \beta > 0$ such that $\mathbb{E}[|X_s - X_t|^\alpha] \leq C|s-t|^{1+\beta}$ for all $s, t \in \mathcal{D}$. Then, for every $\gamma \in [0, \frac{\beta}{\alpha}$, almost surely X is γ-Holder continuous (i.e., there exists $K > 0$ such that $|X_s - X_t| \leq K|s-t|^\gamma$ for all $s, t \in \mathcal{D}$).*

Remark 90. Before beginning the proof, we remark that because we are trying to convert between a statement about expectation and a statement about probability, we probably need to use Markov's and Borel-Cantelli.

Proof. By Markov's inequality, for any $n \in \mathbb{N}$ and $j = 1, ..., 2^n$,

$$\begin{aligned}
\mathbb{P}\left[\left|X_{\frac{j}{2^n}} - X_{\frac{j-1}{2^n}}\right| > 2^{-\gamma n}\right] &= \mathbb{P}\left[\left|X_{\frac{j}{2^n}} - X_{\frac{j-1}{2^n}}\right|^\alpha > 2^{-\gamma n \alpha}\right] \\
&\leq C \cdot 2^{\gamma n \alpha} 2^{-(1+\beta)n} \\
&= C 2^{-(1+\beta-\alpha\gamma)n}.
\end{aligned}$$

Next, apply a union bound over j:

$$\begin{aligned}
\mathbb{P}\left[\max_{1 \leq j \leq 2^n} \left|X_{\frac{j}{2^n}} - X_{\frac{j-1}{2^n}}\right| > 2^{-\gamma n}\right] &\leq C 2^n 2^{-(1+\beta-\alpha\gamma)n} \\
&= C 2^{-(\beta-\alpha\gamma)n}
\end{aligned}$$

By assumption, $\beta - \alpha\gamma > 0$, so the above is summable over n. By Borel-Cantelli, we know that almost surely the event above happens only for finitely many values of n. Equivalently, almost surely, there exists a *random* $N \in \mathbb{N}$ such that

$$\max_{1 \le j \le 2^n} \left| X_{\frac{j}{2^n}} - X_{\frac{j-1}{2^n}} \right| \le 2^{-\gamma n} \quad \forall n \ge N.$$

Because all of the first N maxima are all finite, we can take the existence of a random variable $K > 0$ such that

$$\max_{1 \le j \le 2^n} \left| X_{\frac{j}{2^n}} - X_{\frac{j-1}{2^n}} \right| \le K2^{-\gamma n} \quad \forall n \ge 1.$$

It remains to express $|X_t - X_s|$ in terms of these little dyadic chunks. Fix $s, t \in \mathcal{D}$ and WLOG $s < t$. Let $n \in \mathbb{N}$ such that $2^{-n} \le |s - t| \le 2^{-n+1}$. For $k \ge n$, let $S_k = s$ if $s \in \mathcal{D}_{\|}$ and otherwise S_k is the smallest element $s' \in \mathcal{D}_k$ such that $s' > s$. Similarly, let $t_k = t$ if $t \in \mathcal{D}_k$ and otherwise t_k is the largest element $t' \in D_k$ such that $t' \le t$.

We have:

$$|X_t - X_s| \le \sum_{k=n}^{\infty} |X_{S_k} - X_{S_{k+1}}| + \sum_{k=n}^{\infty} |X_{t_k} - X_{t_{k+1}}|.$$

Observe that on the RHS there are only finitely many non-zero terms since $S_{k+1} = S_k$ for all k that is sufficiently large.

Consider $|X_{s_k} - X_{S_{k+1}}|$. Either the two are equal, or they are consecutive elements of $\mathcal{D}_{k+1}$, and similarly for the other terms. We can therefore bound

$$|X_t - X_s| \le 2K \sum_{k=1}^{\infty} 2^{-\gamma(K+1)} = K'2^{-\gamma n} \le K''|s - t|^{\gamma}.$$

□

Proposition 91. *Almost surely $\{\mathcal{B}_t\}_{t \in \mathcal{D}}$ is γ-Holder for any $\gamma \in (0, \frac{1}{2})$.*

Proof. For all $s, t \in \mathcal{D}$ such that $s < t$, $\mathcal{B}_t - \mathcal{B}_s \sim \mathcal{N}(0, t - s)$. Thus, for any $\alpha > 0$,

$$\mathbb{E}[|\mathcal{B}_t - \mathcal{B}_s|^{\alpha}] = |t - s|^{\alpha/2} \mathbb{E}[|(t - s)^{-1/2} (\mathcal{B}_t - \mathcal{B}_s)|^{\alpha}],$$

where we can let $C_{\alpha} := \mathbb{E}[|(t - s)^{-1/2} (\mathcal{B}_t - \mathcal{B}_s)|^{\alpha}]$ (which is taking the α-th moment of $\mathcal{N}(0, 1)$). Thus,

$$\mathbb{E}[|\mathcal{B}_t - \mathcal{B}_s|^{\alpha}] = |t - s|^{\alpha/2} C_{\alpha}.$$

By Kolmogorov continuity theorem with $\beta = \frac{\alpha}{2} - 1$, we get that $\{\mathcal{B}_t\}_{t \in \mathcal{D}}$ is γ-Holder for all $\gamma < \frac{\frac{\alpha}{2} - 1}{\alpha}$. Send $\alpha \to \infty$ to retrieve the result for all $\gamma \in (0, 1/2)$. □

Theorem 92. *Brownian motion exists and is locally γ-Holder for all $\gamma < \frac{1}{2}$ almost surely.*

Proof. By the previous proposition, plus the fact that uniformly continuous functions on a dense set extend to a unique function on the closure, we see that $\{\mathcal{B}_t\}_{t \in \mathcal{D}}$ extends uniquely to $B : [0, 1] \to \mathbb{R}$. Furthermore, this extension is still γ-Holder for all $\gamma < \frac{1}{2}$ because the extension will have the same modulus of continuity.

Also, note that $\{\mathcal{B}_t\}_{t\in\mathcal{D}}$ have the Gaussian and independent increments properties, so we can can approximate general times by dyadic times to find the same properties. As we recall from earlier, we may construct $B : [0,\infty] \to \mathbb{R}$ by concatenating i.i.d. copies of $B|_{[0,1]}$. $\square$

Remark 93. The globally γ-Holder gets demoted to locally γ-Holder because we concatenate infinitely many copies.

Strong Markov property

Recall the (ordinary) Markov property of Brownian motion, which asserts that $B_{t+\cdot} - \mathcal{B}_t$ is a Brownian motion which is independent from $B|_{[0,t]}$ for a *fixed* time t. A natural question is whether a similar property holds for t being random.

Definition 94. A random time $\tau \in [0,\infty]$ is called a **stopping time** for $\mathcal{B}$ if for every $t > 0$, the event $\{\tau \leq t\} \in \sigma(\mathcal{B}|_{[0,t]})$.

Example 95 (Deterministic time)**.** Any fixed time $\tau = t'$ is a stopping time because $\{t' \leq t\}$ is either $\varnothing$ or Ω.

Example 96 (First time to visit a closed set)**.** Let $A \subset \mathbb{R}$ be a closed set. Then,

$$\tau = \inf\{t > 0 : \mathcal{B}_t \in A\}$$

is a stopping time. This is because

$$\{\tau \leq t\} = \{\mathcal{B}([0,t]) \cap A \neq \varnothing\} \in \sigma(\mathcal{B}|_{[0,t]}).$$

Example 97 (Mins and maxes)**.** If τ_1, τ_2 are stopping times, then $\min\{\tau_1, \tau_2\}$ and $\max\{\tau_1, \tau_2\}$ are stopping times as well.

Non-example 98 (Last time to visit a set)**.** Let $A \subset \mathbb{R}$ be measurable such that $\bar{A} \neq \mathbb{R}$. Then,

$$\tau = \sup\{t \in [0,1] : B_{t\in A}\}$$

is *not* a stopping time. Intuitively, this should be obvious because we can't tell what happens in the future from observing $\mathcal{B}|_{[0,t]}$.

Remark 99. "[A good way to think about stopping times is in terms of driving directions. If Google Maps told you to turn right on Lake Shore Drive at the first stop light, this makes sense. But if Google Maps told you to turn right on the last stoplight on Lake Shore Drive, you wouldn't know when to turn right because at the last stop light you wouldn't be able to know if that was the last one]." - Gwynne, [likely heavily paraphrased]

Theorem 100 (Strong Markov property)**.** *Let τ be a stopping time such that $\mathbb{P}[\tau < \infty] = 1$. Then, $\mathcal{B}_{\cdot+\tau} - \mathcal{B}_\tau$ is a Brownian motion independent from $\sigma(\mathcal{B}|_{[0,t]})$.*

Remark 101. Equivalently, if we condition on $\mathcal{B}_{[0,\tau]}$, then the conditional distribution of $\mathcal{B}_{\cdot+\tau} - \mathcal{B}_\tau$ is that of Brownian motion. This statement can also true if $\mathbb{P}[\tau = \infty] > 0$.

Proof. We set some notation: for $t \geq 0$, set

$$\underline{\mathcal{B}}_s^t = \mathcal{B}_{\min\{s,t\}}$$

$$\overline{\mathcal{B}}_s^t = \mathcal{B}_{s+t} - \mathcal{B}_t.$$

Remark 102. We should think of $\underline{\mathcal{B}}_s^t$ as the being the Brownian motion $\mathcal{B}$ on $s \in [0,t]$, but then it clips off and is constant. We should think of $\overline{\mathcal{B}}_s^t$ as "future" Brownian motion.

Observe that $\sigma(\underline{\mathcal{B}}^t) = \sigma(\mathcal{B}|_{[0,t]})$ because they contain the same information. We want to show that $\overline{\mathcal{B}}^\tau \overset{d}{=} \mathcal{B}$ and $\overline{\mathcal{B}}_s^\tau$ is independent from $\underline{\mathcal{B}}_s^\tau$.

We claim that it suffices to show that

$$\mathbb{P}\left[\overline{B}^\tau \in U, \underline{B}^\tau \in V\right] = \mathbb{P}\left[B \in U\right] \mathbb{P}\left[\underline{B}^\tau \in V\right]$$

for all measurable $U, V \subset C((0,\infty),\mathbb{R})$. To retrieve $\overline{\mathcal{B}}^\tau \overset{d}{=} \mathcal{B}$, we can set $V = \Omega$, and independence follows.

Write out

$$\begin{aligned}\mathbb{P}\left[\overline{B}^\tau \in U, \underline{B}^\tau \in V\right] &= \sum_{t\in T} \mathbb{P}\left[\overline{B}^\tau \in U, \underline{B}^\tau \in V, \tau = t\right] \\ &= \sum_{t\in T} \mathbb{P}\left[\overline{B}^t \in U, \underline{B}^t \in V, \tau = t\right]\end{aligned}$$

Observe that the event

$$\{\tau = t\} = \{\tau \le t\} \setminus \bigcup_{s<t; s\in T} \{\tau \le s\} \in \sigma(\underline{B}^t).$$

Remark 103. Here, we use the property that τ is a stopping time, as all $\{\tau \le s\} \in \sigma(\underline{\mathcal{B}}^t)$.

Furthermore, by the Markov property, $\overline{\mathcal{B}}^t$ is independent from $\underline{B}^t$ (as now we have fixed t) and is distributed as $\mathcal{B}$. So, we can write

$$\begin{aligned}\sum_{t\in T} \mathbb{P}\left[\overline{B}^t \in U, \underline{B}^t \in V, \tau = t\right] &= \sum_{t\in T} \mathbb{P}\left[\overline{B}^t \in U\right] \mathbb{P}\left[\underline{B}^t \in V, \tau = t\right] \\ &= \mathbb{P}\left[\mathcal{B} \in U\right] \sum_{t\in T} \mathbb{P}\left[\underline{B}^t \in V, \tau = t\right] \\ &= \mathbb{P}\left[B \in U\right] \mathbb{P}\left[\underline{B}^\tau \in V\right].\end{aligned}$$

Remark 104. The first equality follows from $\overline{\mathcal{B}}^t$ being independent from the other two events. The second equality holds from the Markov property. The third equality holds because we assumed $\tau \in T$ almost surely.

We now extend this to a general stopping time τ. Let τ_n be the smallest $t \in 2^{-n}\mathbb{Z}$ such that $t \ge \tau$. Since τ is a stopping time, τ_n is also a stopping time. Furthermore, $\tau_n \to \tau$ almost surely. By the continuity of Brownian motion, we must have $\overline{\mathcal{B}}^{\tau_n} \to \overline{\mathcal{B}}^\tau$ and $\underline{\mathcal{B}}^{\tau_n} \to \underline{\mathcal{B}}^\tau$.

From the first limit, we find that $\overline{\mathcal{B}}^\tau \overset{d}{=} \mathcal{B}$. Because independence is preserved under convergence in distribution, we also find that $\overline{\mathcal{B}}^\tau$ is independent from $\underline{\mathcal{B}}^\tau$. □

Non-example 105 (Fails when τ is not stopping)**.** Consider

$$\tau = \max\{t \in [0,1] : \mathcal{B}_t \in A\}$$

as before, which fails to be a stopping time. If we condition on $\mathcal{B}|_{[0,\tau]}$, then the conditional distribution of $\overline{\mathcal{B}}^\tau$ is a Brownian motion conditioned not to hit $A - \mathcal{B}_\tau$ during $[0, 1-\tau]$.

Proposition 106 (Reflection principle)**.** *Let $a > 0$. Then, for any $t \ge 0$,*

$$\mathbb{P}\left[\max_{0\le s\le t} \mathcal{B}_s \ge a\right] = 2\mathbb{P}\left[\mathcal{B}_t \ge a\right].$$

Equivalently,

$$\max_{0\le s\le t} \mathcal{B}_s \overset{d}{=} |\mathcal{B}_t|.$$

Proof. We first show that these statements are equivalent:

$$\mathbb{P}[|\mathcal{B}_t| > a] = \mathbb{P}[\mathcal{B}_t > a] + \mathbb{P}[\mathcal{B}_t < -a] = 2\mathbb{P}[\mathcal{B}_t \geq a].$$

We show the first statement. Let

$$\tau = \min\{t \geq 0 : \mathcal{B}_t = a\}.$$

Note τ is a stopping time. Furthermore,

$$\{\max_{0\leq s\leq t} \mathcal{B}_s \geq a\} = \{\tau \leq t\}.$$

By the law of conditional probability,

$$\mathbb{P}\left[\mathcal{B}_t \geq a\right] = \mathbb{P}\left[\tau \leq t\right] \mathbb{P}\left[\mathcal{B}_t \geq a | \tau \leq t\right].$$

We look at

$$\mathbb{P}\left[\mathcal{B}_t \geq a | \tau \leq t\right].$$

By the strong Markov property, we know that the conditional distribution of $\mathcal{B}_t - \mathcal{B}_\tau$ given $\mathcal{B}|_{[0,\tau]}$ on $\{\tau \leq t\}$ is normal with mean 0 and variance with $t - \tau$. On $\{\tau \leq t\}$, we have that $\mathbb{P}\left[\mathcal{B}_t \geq a|\mathcal{B}|_{[0,\tau]}\right] = \mathbb{P}\left[\mathcal{N}(0, \tau - t) \geq 0\right] = \frac{1}{2}$.

We see that

$$\mathbb{P}\left[\mathcal{B}_t \geq a | \tau \leq t\right] = \frac{1}{2},$$

so we find that

$$\mathbb{P}\left[\mathcal{B}_t \geq a\right] = \frac{1}{2}\mathbb{P}\left[\tau \leq t\right] = \mathbb{P}[\max_{0\leq s\leq t} \mathcal{B}_s \geq a].$$

□

Remark 107. This statement does *not* hold as a function over t. One way to see this is that the LHS is increasing in t, but the RHS is likely not.

As a corollary, we find:

Proposition 108. *Almost surely,*

$$\limsup_{t\to\infty} \mathcal{B}_t = \infty$$

$$\liminf_{t\to\infty} \mathcal{B}_t = -\infty$$

Proof. We know that

$$\max_{0\leq s\leq t} \mathcal{B}_s \stackrel{d}{=} |\mathcal{B}_t| \stackrel{d}{=} t^{1/2}|\mathcal{N}(0,1)|,$$

and this implies that for all $a \in \mathbb{R}$:

$$\lim_{t\to\infty} \mathbb{P}[\max_{0\leq s\leq t} \mathcal{B}_s > a] = 1.$$

Thus, almost surely,

$$\limsup_{t\to\infty} \mathcal{B}_t = \infty.$$

One can derive the other inequality by realizing $-\mathcal{B} \stackrel{d}{=} \mathcal{B}$. □

Brownian motion in $\mathbb{R}^d$

We define Brownian motion in $\mathbb{R}^d$ as one would expect:

Definition 109. Let $d \in \mathbb{N}$. Then, **Brownian motion in** $\mathbb{R}^d$ is the random continuous function $\mathcal{B} = (\mathcal{B}^1, \ldots, \mathcal{B}^d) : [0, \infty] \to \mathbb{R}^d$ where $\mathcal{B}^1, \ldots, \mathcal{B}^d$ are i.i.d. one-dimensional Brownian motions.

Some consequences from the 1D case:

Proposition 110. *Brownian motion on $\mathbb{R}^d$ is the unique (in distribution) random continuous function $\mathcal{B} : [0, \infty] \to \mathbb{R}^d$ such that $\mathcal{B}_0 = 0$ and*

- *For any $s \le t$, $\mathcal{B}_t - \mathcal{B}_s$ is a vector of d i.i.d. $\mathcal{N}(0, t-s)$ random variables.*
- *For any $t_0 \le \cdots < t_N$ we have that $\mathcal{B}_{t_j} - \mathcal{B}_{t_{j-1}}$ are independent.*

Lemma 111 (Brownian scaling)**.** *For any $C > 0$,*

$$\{C^{-1/2}\mathcal{B}_{Ct}\}_{t\ge 0} \stackrel{d}{=} \{\mathcal{B}_t\}_{t \ge 0}.$$

Proposition 112 (Strong Markov property)**.** *Let τ be a stopping time for $\mathcal{B}$, and let $\mathbb{P}[\tau < \infty] = 1$. Then, $\mathcal{B}_{\cdot+\tau} - \mathcal{B}_\tau$ is a d-dimensional Brownian motion and is independent from $\sigma(\mathcal{B}|_{[0,\tau]})$.*

A new property in d-dimensions:

Proposition 113 (Rotational invariance)**.** *Let Q be a $d \times d$ orthogonal matrix (whose columns form an orthonormal basis for $\mathbb{R}^d$). Then $Q\mathcal{B} \stackrel{d}{=} \mathcal{B}$.*

Proof. That $Q\mathcal{B}$ has independent increments is straightforward.

To obtain the Gaussian increments, we use the fact that Q multiplied by a vector of Gaussians yields a vector with the same distribution, provided the components $X_1, \ldots, X_d$ are i.i.d. $\mathcal{N}(0, t-s)$ (as proven on problem set 5). □

Remark 114. This property is conceptually important because when we defined the motion, we implicitly specified a basis. Rotational invariance demonstrates that this choice of basis does not matter, making the object more canonical.

We describe some qualitative properties of d-dimensional Brownian motion. All statements hold almost surely, and proofs can be found in [MP10].

- $d = 2$:
  - Self-intersections in every time interval: for all $a < b$, there exists times $s, t \in [a, b]$ distinct such that $\mathcal{B}_s = \mathcal{B}_t$. There exists points that are hit infinitely many times.
  - Neighborhood recurrent: for every $z \in \mathbb{R}^2$, for every $\epsilon > 0$, there exists arbitrarily large times t such that $\mathcal{B}_t \in D_\epsilon(z)$ [ball].
  - For all $z \in \mathbb{R}^2$ fixed, $\mathbb{P}[z \text{ is hit by } \mathcal{B}] = 0$.

- $d = 3$:
  - Self-intersections in every time interval.
  - Visits each point at most twice.
  - Transience: $\lim_{t\to\infty} |\mathcal{B}_t| = \infty$.
- $d \geq 4$:
  - No self-intersections, so $\mathcal{B}_s \neq \mathcal{B}_t$ whenever $s \neq t$.
  - Transcience.

Remark 115. The range of a Brownian motion has Hausdorff dimension 2 for any $d \geq 2$.

Remark 116. One may wonder if there is an analogue of Brownian motion with d time variables (or equivalently, is there a natural random function $\mathbb{R}^d \to \mathbb{R}$). The analog is called a Gaussian free field (GFF). Unfortunately, this is not a function and is a generalized function (distribution).

Brownian motion and complex analysis

Brownian motion and harmonic functions

Let $U \subset \mathbb{R}^d$ be an open domain. Recall that a C^2 function $u : U \to \mathbb{R}$ is harmonic if $\Delta u = 0$ where

$$\Delta = \sum_{j=1}^{d} \partial_{x_j}^2.$$

Let $\varphi : \partial U \to \mathbb{R}$ be continuous.

Definition 117. The **Dirichlet problem** is given by the following: find $u : \overline{U} \to \mathbb{R}$ which is continuous and satisfies $u|_{\partial U} = \varphi$ and $\Delta u = 0$ on U.

Our goal is to construct a solution to the Dirichlet problem via Brownian motion.

We set some notation: for $z \in \mathbb{R}^d$, let $\mathbb{P}_z$ be the distribution of Brownian motion starting at z, i.e., the distribution of $\mathcal{B}^z = B + z$. Let $\mathbb{E}_z$ be the expectation

$$\mathbb{E}_z[f(B)] := \mathbb{E}_0[f(B + z)]$$

By the strong Markov property, if τ is a stopping time, then the conditional distribution of $\mathcal{B}_{\cdot+\tau}$ given $\mathcal{B}|_{[0,\tau]}$ is $\mathbb{P}_{B_\tau}$ as one can write

$$\mathcal{B}_{\cdot+\tau} = \mathcal{B}_\tau + (\mathcal{B}_{\cdot+\tau} - \mathcal{B}_\tau).$$

Proposition 118. *Let $U \subset \mathbb{R}^d$ be open. Set*

$$\tau := \tau_U := \inf\{t > 0 : \mathcal{B}_t \notin U\}.$$

Assume that for all $z \in U$, $\mathbb{P}_z[\tau_U < \infty] = 1$. Let $\varphi : \partial U \to \mathbb{R}$ be bounded and measurable. Let $u(z) := \mathbb{E}_z[\varphi(\mathcal{B}_\tau)]$ for all $z \in U$. Then, u is harmonic.

Remark 119. A point of confusion on the notation is that τ depends on the starting point z, and we are a little ambiguous on it.

Remark 120. This proposition is not sufficient for a solution to the Dirichlet problem because we did not show that u continuously extends to φ on ∂U, which is the bulk of the technical work.

Remark 121. We can satisfy $\mathbb{P}[\tau_U < \infty] = 1$ when U is bounded because $\limsup_{t\to\infty} |\mathcal{B}_t|$ is ∞.

Proof. It suffices to show that u satisfies the mean value property: if $r > 0$ such that $\overline{D_r(z)} \subset U$, then

$$u(z) = \int_{\partial D_r(z)} u(w) d\sigma(w)$$

where $\sigma(w)$ is the uniform measure on $\partial D_r(z)$.

Let $T := \min\{t > 0 : \mathcal{B}_t \in \partial D_r(z)\}$. Observe that T is a stopping time and almost surely $T < \tau$. By the tower property,

$$u(z) = \mathbb{E}_z[\mathbb{E}_z[\varphi(\mathcal{B}_\tau)|\mathcal{B}|_{[0,T]}]].$$

By the strong Markov property, this is equal to:

$$u(z) = \mathbb{E}_z[\mathbb{E}_{\mathcal{B}_T}[\varphi(\mathcal{B}_\tau)]] = \mathbb{E}_z[u(\mathcal{B}_T)]$$

Because $\mathcal{B}_T$ is uniformly distributed according to the uniform measure by rotational invariance (as there is only one uniform measure here), we conclude that the mean value property holds. □

A question raised by someone in the class:

Question 122. *How can someone guess that Brownian motion is related to harmonic functions?*

The answer is that people saw connections from a particular solution of the heat equation $\partial_t u(t,x) = \Delta u(t,x)$ via Gaussian density, where the solution is given by some

$$u(t,x) = \frac{1}{(2\pi t)^{d/2}} e^{-|x|^2/2t},$$

(up to constants) which inspired connections between probability and PDEs.

Before proceeding onto the technical proof for the Dirichlet problem, we describe the intuition. We would like to show that if z is close to ∂U, then $|\mathcal{B}_\tau - z|$ is small with high probability. Certainly, we should not expect this to hold for *all* open domains, as it is clear from the properties of two-dimensional Brownian motion that the punctured disk $D_1(0) \setminus \{0\}$ will not satisfy this property around 0. It turns out that it suffices that if $d = 2$ and U is simply connected, the extension will be continuous.

Lemma 123. *Assume $d = 2$ and $x \in \partial D_{1/2}(0)$. Let $\mathcal{B}_0 = x$. Then,*

$$\mathbb{P}\left[\ \mathcal{B} \textit{ disconnects } \partial D_{\frac{1}{2}}(0) \textit{ from } \partial D_1(0) \textit{ before leaving } D_1(0)\ \right] > 0.$$

Remark 124. The point of this proof is to understand how to demonstrate that Brownian motion "does something" with non-zero probability, though this proof is very informal.

Proof. Let $\epsilon > 0$ and fix some z while letting τ_1 be the exit time of $\mathcal{B}$ from $D_\epsilon(z)$. The key point is that $\mathcal{B}_{\tau_1} \sim \text{Unif}(\partial D_z(0))$ due to rotational symmetry. Define τ_k similarly. Thus, by the strong Markov property, the conditional distribution of $\mathcal{B}_{\tau_k}$ given $\mathcal{B}|_{[0,\tau_{k-1}]}$ is $\text{Unif}(\partial D_\epsilon(\mathcal{B}_{\tau_{k-1}}))$.

Using this idea, we can "discretize" any curve and argue that there is a finite probability of making some progress along the curve, so for any curve,

$$\mathbb{P}\left[\mathcal{B} \text{ stays } \epsilon\text{-close to given curve until reaching } \epsilon\text{-neighborhood of endpoint}\right] > 0.$$

If $\epsilon > 0$ is small enough, then this implies that $\mathcal{B}$ disconnects $\partial D_{1/2}(0)$ from $\partial D_1(0)$ before exiting $D_1(0)$. $\square$

Lemma 125. *Let $\mathcal{B}$ be a two-dimensional Brownian motion. There exists $C, \alpha > 0$ such that for all $R > r > 0$,*

$$\mathbb{P}_0\left[B \text{ disconnects } \partial D_r(0) \text{ from } \partial D_R(0) \text{ before leaving } D_R(0)\right] \geq 1 - C\left(\frac{r}{R}\right)^{\alpha}.$$

Remark 126. The point of this proof is show how to "upgrade" a result from positive probability to high probability.

Proof. The main idea is to decompose big annulus into logarithmically many small ones, and argue each small annulus independently makes a loop.

For $n \in \mathbb{Z}$, let $\tau_n = \min\{t > 0 : B_t \in \partial D_{2^n}(0)\}$. Let

$$E_n = \{B \text{ makes a loop that disconnects } \partial D_{2^{n-1}}(0) \text{ from } \partial D_{2^n}(0) \text{ between } \tau_{n-1} \text{ and } \tau_n\}$$

Observe that by the strong Markov property,

$$\mathbb{P}\left[E_n | \mathcal{B}|_{[0,\tau_{n-1}]}\right] = \mathbb{P}_{\mathcal{B}_{\tau_{n-1}}}\left[\mathcal{B} \text{ makes a loop that disconnects } \partial D_{2^{n-1}}(0) \text{ from } \partial D_{2^n}(0) \text{ before leaving } D_{2^n}(0)\right]$$

But this is the same as

$$\mathbb{P}_{\left(\frac{1}{2},0\right)}\left[\text{disconnects } \partial D_{1/2}(0) \text{ from } \partial D_1(0) \text{ before leaving } D_1(0)\right] > 0.$$

Also,

$$\{E_j\}_{j \leq n-1} \subset \sigma(B|_{[0,\tau_{n-1}]}).$$

so E_n is independent from $\{E_j\}_{j \leq n-1}$ and. Thus, $\{E_n\}_{n \in \mathbb{Z}}$ are i.i.d.

$$\mathbb{P}\left[\bigcap_{n=n_1+1}^{n_2} E_n^C\right] = (1-p)^{n_2 - n_1}.$$

Choose $n_1 < n_2$ such that $[2^{n_1}, 2^{n_2}] \subset [r, R]$, and $n_2 - n_1 \geq c \log(R/r)$.

$$\mathbb{P}\left[\text{event in lemma}\right] \geq \mathbb{P}\left[\bigcup_{n=n_{1+1}}^{n_2} E_n^C\right] \geq 1 - (1-p)^{c\log(R/r)} = 1 - e^{-\log 1/(1-p) c \log(R/r)}$$
$$= 1 - \left(\frac{r}{R}\right)^{\alpha}$$

where $\alpha = c \log \frac{1}{1-p}$. $\square$

Theorem 127 (Existence of solution to Dirichlet problem)**.** *Let $U \subset \mathbb{R}^2$ be open, simply connected, and $U \neq \mathbb{R}^2$. Let $\tau = \inf\{t > 0 : \mathcal{B}_t \notin U\}$. Let $\varphi : \partial U \to \mathbb{R}$ be bounded and continuous. Let*

$$u(z) = \begin{cases} \mathbb{E}_z[\varphi(\mathcal{B}_\tau)] & z \in U \\ \varphi(z) & z \in \partial U \end{cases}.$$

Then, u is continuous on $\overline{U}$ and harmonic on U.

Proof. First, check that u is well-defined. Claim $\mathbb{P}_z[\tau < \infty] = 1$ for all $z \in U$. By the previous lemma, almost surely there exists arbitrarily large r such that $\mathcal{B}$ disconnects $D_r(0)$ from ∞.

We have that ∂U is connected (in $\mathbb{C}_\infty$) because U is simply connected. If U is unbounded, then ∂U is unbounded. And the above implies that $\mathcal{B}$ must eventually hit ∂U.

We know already that u is harmonic on U, which implies u is continuous on U. Let $x \in \partial U$ and $\epsilon > 0$. Then, there exists $\delta \in (0, \epsilon)$ such that if $y \in \partial U$ and $|x - y| < \delta$, then $|\varphi(x) - \varphi(y)| < \epsilon$. Let $z \in U$ such that $|z - x| < \delta^2$.

By the lemma,

$$\mathbb{P}_z\left[\mathcal{B} \text{ disconnects } D_{\delta^2}(z) \text{ from } \partial D_\delta(z) \text{ before leaving } D_\delta(z)\right] \geq 1 - c\delta^\alpha$$

implies that

$$\mathbb{P}_z\left[\tau < \text{exit time from} D_\delta(z)\right] \geq 1 - c\delta^\alpha$$

So

$$u(z) = \mathbb{E}_z[\varphi(B_\tau)] \leq \varphi(x) + \epsilon + ||\varphi||_\infty C\delta^\alpha$$

where the terms can be made arbitrarily small. Similarly,

$$u(z) = \mathbb{E}_z[\varphi(B_\tau)] \geq \varphi(x) - \epsilon - ||\varphi||_\infty C\delta^\alpha.$$

□

Proposition 128 (Uniqueness of solution)**.** *Suppose U is bounded. Let $u : \overline{U} \to \mathbb{R}$ be continuous, $V = \varphi$ on ∂U, $\Delta v = 0$ on U. Then $v = u$.*

Proof. Note $u - v$ is harmonic on U, continuous on $\overline{U}$, and $u - v = 0$ on ∂U. So by maximum principle, maximum and minima must be on boundary of U, so $u - v = 0$. □

Non-example 129 (Solution is not unique if U unbounded)**.** Let $U = \{\mathrm{Im}(z) \geq 0\}$ and $\phi \equiv 0$. Then, we have $u(z) \equiv 0$ but also $v(z) \equiv e^{\mathrm{Re}(z)} \sin(\mathrm{Im}(z)) = \mathrm{Im} e^z$.

Remark 130*.* The existence of the Dirichlet solution can also be proven for when ∂U is a finite union of non-singleton sets by essentially the same technique.

Remark 131*.* For $d \geq 3$, Brownian motion cannot disconnect $\partial D_r(0)$ from $\partial D_R(0)$. Topological conditions on the domain are *not* sufficient in higher dimensions. e.g., ball with narrow cusp, and start at the center and it kind of winds around the cusp. One condition that is sufficient is that ∂U has finitely many components, each of which is a $d - 1$ dimensional C^1 submanifold of $\mathbb{R}^d$.

Brownian motion and biholomorphisms

Identify $x + iy \in \mathbb{C}$ with $(x, y) \in \mathbb{R}^2$. Let $\mathcal{B} = \mathcal{B}^1 + i\mathcal{B}^2$, where $\mathcal{B}^1$ and $\mathcal{B}^2$ are one-dimensional Brownian motions.

Definition 132. Let $U \subset \mathbb{C}$ be open. Let $\tau = \min\{t \geq 0 : \mathcal{B}_t \notin U\}$. Assume that for all $z \in U$, $\mathbb{P}_z[\tau < \infty] = 1$. The **harmonic measure** on ∂U viewed from $z \in U$ is the distribution of $\mathcal{B}_\tau$ under $\mathbb{P}_z$.

$$\mathrm{hm}_U^z(A) = \mathbb{P}_z[\mathcal{B}_\tau \in A]$$

for all $A \subset \partial U$ Borel.

For simply connected U,

$$u(z) = \mathbb{E}_z[\varphi(\mathcal{B}_\tau)] = \int_{\partial U} \varphi(x) d\, \mathrm{hm}_U^z(x)$$

Example 133 (Harmonic measure on disk). $\mathrm{hm}_{B_1(0)}^0$ is the uniform measure on $\partial B_1(0)$ by rotational invariance.

Theorem 134 (Harmonic measure is conformally invariant). *Suppose $U, V \subset \mathbb{C}$ are open, simply connected, $U, V \neq \mathbb{C}$, U bounded. Let $f : U \to V$ be a biholomorphism. Assume f extends to a continuous function $f : \overline{U} \to \overline{V}$. Then, $\mathrm{hm}_V^z(A) = \mathrm{hm}_U^{f^{-1}(z)}(f^{-1}(A))$ for all $z \in V$ and $A \subset \partial V$ Borel.*

Remark 135. Alternatively, this shows that harmonic measure is pushed forward from U to V by f.

Proof. Let $\varphi : \partial V \to \mathbb{R}$ be continuous and bounded. Let v solve the Dirichlet problem on V with $V|_{\partial V} = \varphi$. We have previously shown that $v \circ f$ is a harmonic function on U, with $v \circ f|_{\partial U} = \varphi \circ f$. By the uniqueness of the Dirichlet problem, for all $w \in U$,

$$(v \circ f)(w) = \mathbb{E}_w[(\varphi \circ f)(\mathcal{B}_{\tau_U})] = \int_{\partial U} \varphi \circ f(x) \, \mathrm{dhm}_U^w(x).$$

Take $w = f^{-1}(z)$, $z \in V$ to find

$$\int_{\partial U} \varphi(f(x)) \, \mathrm{dhm}_U^{f^{-1}(z)}(x) = v(z) = \int_{\partial V} \varphi(y) \, \mathrm{dhm}_v^z(y).$$

Recall that a measure can be defined by its integral against all functions, so $f_* \, \mathrm{hm}_U^{f^{-1}(z)} = \mathrm{hm}_V^z$. □

Example 136 (Positive probability of exiting domain). Let $U \subset \mathbb{C}$ be simply connected and let $f : B_1(0) \to U$ extend continuously to $\overline{B_1(0)} \to \overline{U}$. Let $A \subset \partial U$ be any non-singleton connected Borel set. Let $f^{-1}(A)$ be a union of non-singleton arcs of $\partial B_1(0)$. For $z \in U$, take $f(0) = z$. Then,

$$\mathrm{hm}_U^z(A) = \mathrm{hm}_{B_1(0)}^0(f^{-1}(A)) > 0.$$

This tells us that even in weird domains the probability of leaving from a certain connected set A is positive.

The following examples demonstrate that Brownian motion can tell us what the biholomorphism supplied by the Riemann mapping theorem can "look like".

Example 137 (Conjoined circles)**.** For $\epsilon > 0$, let

$$U_\epsilon = D_1(0) \cup D_1(2) \cup D_\epsilon(1).$$

Let $f_\epsilon : B_1(0) \to U_\epsilon$ be a biholomorphism with $f_\epsilon(0) = 0$. What does $f_\epsilon^{-1}(\partial U_\epsilon \cap \partial D_1(2))$ look like?

Proof. It is kind of tiny because the harmonic measure is small: by conformal invariance,

$$\begin{aligned} \mathrm{hm}^0_{D_1(0)}(f^{-1}(\partial U_\epsilon \cap \partial D_1(2))) &= \mathrm{hm}^0_{U_\epsilon}(\partial U_\epsilon \cap \partial D_1(2)) \\ &\leq \mathrm{hm}^0_{D_1(0)}(\partial D_1(0) \cap \partial D_\epsilon(1)) \\ &= O(\epsilon). \end{aligned}$$

□

Example 138 (Disk with slit)**.** Let $U_\epsilon = D_1(0) \setminus [\epsilon, 1]$, and let $f_\epsilon : D_1(0) \to U_\epsilon$ be a biholomorphism with $f_\epsilon(0) = 0$. Then, $f^{-1}([\epsilon, 1])$ is most of the disk boundary since with high probability, Brownian motion forms a loop before exiting $B_1(0)$. So,

$$\mathrm{hm}^0_{U_\epsilon}([\epsilon, 1]) \to 1 \qquad \text{as } \epsilon \to 0.$$

Remark 139. As a technical remark, we still consider the boundary of U_ϵ to be a curve; it is just not simple.

Example 140 (Explicit Riemann mapping)**.** Let $U \subset \mathbb{C}$ be open and bounded by a simple loop. Fix $z_0 \in U$, $x_0 \in \partial U$. Construct $f : \overline{U} \to \overline{D_1(0)}$ such that $f(z_0) = 0$ and $f(x_0) = 1$ as follows: for $y \in \partial U$, denote $p(y) = \mathrm{hm}^{z_0}_U(\text{CCW arc from } x_0 \text{ to } y)$. Define

$$f(y) = e^{2\pi i p(y)}.$$

On the interior, for $z \in U$ let

$$f(z) := \mathbb{E}_z[f(\mathcal{B}_{\tau_U})],$$

as in the solution to the Dirichlet problem.

We claim that f is a biholomorphism via the Riemann mapping theorem. Let $\tilde{f} : U \to D_1(0)$ be the biholomorphism supplied by the Riemann mapping theorem with $\tilde{f}(z_0) = 0, \tilde{f}(x_0) = 1$. Note that $\tilde{f}$ preserves harmonic measure, which is equivalent to saying that h and f map points on the boundary to the same place. Thus, they have the same boundary data and by uniqueness, $f = \tilde{f}$.

Recurrence and transience

Definition 141. Let $X : [0, \infty) \to \mathbb{R}^d$ be a random continuous function. We say that X is:

- **point recurrent** if for all $z \in \mathbb{R}^d$ and for all $T > 0$,

$$\mathbb{P}[\text{exists } t > T \text{ such that } X_t = z] = 1.$$

- **neighborhood recurrent** if for all $z \in \mathbb{R}^d$, $T > 0$, and $\epsilon > 0$,

$$\mathbb{P}[\text{exists } t > T \text{ such that } |X_t - z| \leq \epsilon] = 1.$$

- **transient** if almost surely,

$$\lim_{t \to \infty} |X_t| = \infty.$$

Remark 142. Should be pretty clear that point recurrent implies neighborhood recurrent implies not transient.

Theorem 143. *Brownian motion is point recurrent in $d = 1$, neighborhood recurrent in $d = 2$, transient in $d \geq 3$.*

Proof. Assume without loss of generality that $\mathcal{B}_0 = 0$. For $d = 1$, we previously saw that $\limsup_{t\to\infty} |\mathcal{B}_t| = \infty$, $\liminf_{t\to\infty} |\mathcal{B}_t| = -\infty$, so it is point current (from continuity).

Let $d \geq 2$. Let $z \in \mathbb{R}^d$ and let $0 < \epsilon < |z| < R$. Let $\tau = \min\{t \geq 0 : |\mathcal{B}_t - z| \in \{\epsilon, R\}\}$. We can supply a solution to the Dirichlet on the annulus via:

Remark 144. Here, τ is the minimum time the Brownian motion that starts in the annulus exits the annulus.

$$u(w) = \begin{cases} \log |w - z| & d = 2 \\ |w - z|^{2-d} & d \geq 3 \end{cases},$$

where one can compute that u is harmonic on $\mathbb{R}^d \setminus \{z\}$.

We know $u(0) = \mathbb{E}[u(\mathcal{B}_\tau)] = u(\epsilon)\mathbb{P}\left[|\mathcal{B}_\tau - z| = \epsilon\right] + u(R)\mathbb{P}\left[|\mathcal{B}_\tau - z| = R\right]$. We can rewrite $\mathbb{P}\left[|\mathcal{B}_\tau - z| = R\right] = 1 - \mathbb{P}\left[|\mathcal{B}_\tau - z| = \epsilon\right]$, and then solve to find

Remark 145. We are cheating a little bit because we didn't deal with the multiple connected components on the boundary, but Dirichlet still works on the two circles/shells.

$$\mathbb{P}\left[|\mathcal{B}_\tau - z| = \epsilon\right] = \frac{u(R) - u(0)}{u(R) - u(\epsilon)}.$$

Plug in for $d = 2$:

$$\mathbb{P}\left[|\mathcal{B}_\tau - z| = \epsilon\right] = \frac{\log R - \log |z|}{\log R - \log \epsilon}.$$

This goes to 1 as $R \to \infty$. This is good enough to imply neighborhood recurrent because it must hit the epsilon neighborhood before going far. Also,

$$\lim_{\epsilon \to 0} \mathbb{P}\left[|\mathcal{B}_\tau - z| = \epsilon\right] = 0,$$

which implies not point recurrent.

Let $d \geq 3$. Then,

$$\mathbb{P}\left[|\mathcal{B}_\tau - z| = \epsilon\right] = \frac{R^{2-d} - |z|^{2-d}}{R^{2-d} - \epsilon^{2-d}}.$$

As $R \to \infty$, this goes to $\frac{|z|^{2-d}}{\epsilon^{2-d}}$.

Let $S > 0$ large, and let $T = \min\{t \geq 0 : |\mathcal{B}_t| = S\}$. This is finite with probability 1. By the strong Markov property, the conditional distribution of the random continuous function $s \mapsto \mathcal{B}_{s+T}$ given $\mathcal{B}|_{[0,T]}$ is $\mathbb{P}_{\mathcal{B}_T}$.

$$\mathbb{P}_{\mathcal{B}_T}\left[\text{Brownian motion ever visits } D_\epsilon(0)\right] = \mathbb{P}_0\left[\text{Brownian motion ever visits } D_\epsilon(-\mathcal{B}_T) | \mathcal{B}_T\right]$$

$$= \frac{|\mathcal{B}_T|^{2-d}}{\epsilon^{2-d}} = \frac{S^{2-d}}{\epsilon^{2-d}}.$$

This goes to 0 as $S \to \infty$, which implies almost surely, there exists some time $t_\epsilon > 0$ such that $\mathcal{B}$ does not reenter $D_\epsilon(0)$ after time t_ϵ. This implies transient (ϵ large).

□

Some open problems

Definition 146. A set $A \subset \mathbb{C}$ is **conformally removable** if whenever $f : \mathbb{C} \to \mathbb{C}$ is a homeomorphism and holomorphic on $\mathbb{C} \setminus A$, then f is holomorphic on $\mathbb{C}$.

Example 147 (Conformally removable)**.** A finite set is conformally removable as a consequence of the Riemann removable singularities theorem.

Non-example 148 (Not conformally removable)**.** If A has non-empty interior, then A is not conformally removable. (Given a homeomorphism, one can tweak values inside the interior.)

Theorem 149 (Jones–Smirnov)**.** *Let $U \subset \mathbb{C}$ be open and simply connected, and assume that the Riemann map $f : B_1(0) \to U$ extends to a Hölder-continuous function $\overline{B_1(0)} \to \overline{U}$. Then, the boundary of U is conformally removable.*

Let $\mathcal{B}$ be a two-dimensional Brownian motion. This leads to the following open questions:

Question 150. *Is $\mathcal{B}[0,1]$ conformally removable?*

Question 151. *Let $z, w \in \mathbb{C} \setminus \mathcal{B}[0,1]$. Does there exist a path from z to w which crosses $\mathcal{B}[0,1]$ only finitely many times?*

Remark 152. The general belief is "no" for both of these questions.

Question 153. *What is the "size" of the smallest path in $\mathcal{B}[0,1]$ from 0 to $\mathcal{B}_1$? What if we require the path to be visited in order by $\mathcal{B}$? (This is interesting in $d = 2, 3$ because in higher dimensions there are no self-intersections.)*

For the lower bound, it is known that the range of Brownian motion does not contain a line segment.

Conjecture 154. *The range does not contain a non-trivial path of finite Euclidean length.*

For the upper bound, it is known that tor $d = 2$, $\mathcal{B}[0,1]$ contains a path of Hausdorff dimension $5/4$ hit in order by $\mathcal{B}$. The general belief is that one can find a path of Hausdorff dimension 1, but that path does not have finite length (with no requirement on hitting order).

References

[Dur19] Rick Durrett. *Probability—theory and examples*, volume 49 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 2019. Fifth edition of [MR1068527].

[MP10] Peter Mörters and Yuval Peres. *Brownian motion*, volume 30 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 2010. With an appendix by Oded Schramm and Wendelin Werner.