Math 275D Lecture Notes

Professor: Jun Yin
Scribe: Daniel Raban

Contents

1 Introduction to Brownian Motion

1.1 Definition of Brownian motion

Here is a heuristic definition of Brownian motion, describing its properties.

Definition 1.1. **Brownian motion** is a "random" function $B(t)$ for $t \geq 0$, $t \in \mathbb{R}$ such that

1. $B(0) = 0$
2. (Independence on disjoint intervals) $B(t_2) - B(t_1)$ is independent of $B(t_4) - B(t_3)$ if $[t_3, t_4] \cap [t_2, t_1] = \varnothing$ for any $t_1, t_2, t_3, t_4 \in \mathbb{R}$ with $t_i \geq 0$.
3. For all $t > 0$, $B(t) \sim N(0, t)$.
4. $B(t)$ is continuous a.s.

We can ask many questions about Brownian motion. For example, we can ask what is $\sup_{0 \leq t \leq 1} |B(t)|$?

1.2 Comparison with the Poisson process

This is similar to a Poisson Process. Recall the Poisson Process:

Definition 1.2. A **Poisson process** is a "random" function $N(t)$ for $t \geq 0$, $t \in \mathbb{R}$ such that

1. $N(0) = 0$
2. $N(t_4) - N(t_3)$ is independent of $N(t_2) - N(t_1)$ if $[t_3, t_4] \cap [t_2, t_1] = \varnothing$ for any $t_1, t_2, t_3, t_4 \in \mathbb{R}$ with $t_i \geq 0$.
3. $N(t) \sim \text{Pois}(t)$.

How do we know such a thing actually exists? We need to have a probability space $(\Omega, \mathcal{F}, \mathbb{P})$? What are the sample space and measurable sets we want to talk about? To talk about a Poisson process, we only need to know when the times of the jumps are. So we can take Ω to be the set of step functions and $\mathcal{F} = \sigma(N(t) : t \geq 0)$.

We can define $\Omega, \mathcal{F}$ for Brownian motion similarly. But how do we define $\mathbb{P}$? The Heuristic definition only defines $\mathbb{P}$ on some special events! Next time, we will make this rigorous. Later, we will define it in yet another, better way and show that the methods are equivalent.

2 Independence Properties and Construction of Brownian Motion

2.1 Independence of sections of Brownian motion

Denote independence by $\perp$. We know that $(B(t_2) - B(t_1)) \perp (B(s_2) - B(s_2))$ if $[s_1, s_2] \cap [t_1, t_2] = \varnothing$. However, this does not directly imply that the random variables $B(x)$ for $x \in [t_1, t_2]$ and $B(y)$ for $y \in [s_1, s_2]$ are independent. To prove this, we recall a consequence of the π-λ lemma.

Lemma 2.1. *Suppose $\mathcal{T}_i = \sigma(\mathcal{A}_i)$, where $\mathcal{A}_i$ is a π-system for $i = 1, 2$. If $A_1 \perp A_2$ for all $A_1 \in \mathcal{A}_1$ and $A_2 \in \mathcal{A}_2$, then $\mathcal{T}_1 \perp \mathcal{T}_2$.*

Proposition 2.1. *Let $t_1 < t_2 < s_1 < s_2$, and let $f(a) = B(t_1+a) - B(t_1)$ for $a \in [0, t_2 - t_1]$ and $g(b) = B(s_1+b) - B(s_1)$ for $b \in [0, s_2 - s_1]$. These two random functions are independent of each other.*

Remark 2.1. This is stronger than the fixed coordinates being independent because we can say things like $\max_a f(a) \perp \max_b g(b)$.

An important consequence of this is that for any t_0, Brownian motion from 0 to t_0 is independent of what happens after t_0.

Proposition 2.2. *Let $a > 0$. Then $B(at) \sim \sqrt{a}N(0, t)$.*

2.2 Difficulty in construction of Brownian motion

How can we construct Brownian motion? Recall that constructing $X \sim U[0, 1]$ is difficult; we have to talk about σ-fields and Lebesgue measure. A main difficulty is that not all sets are measurable. So we need to find a decent collection of measurable sets of functions for Brownian motion.

If we want to construct random vectors (X, Y), then we have to have $\mathcal{F} = \sigma(\mathcal{F}_X \times \mathcal{F}_Y)$. If we have a random sequence $(X_1, X_2, \dots)$, we have the σ-field $\sigma(\bigcup \mathcal{F}_i)$, but we need to use the Kolmogorov extension theorem[1] to construct $\mathbb{P}$.

With the Poisson process, we only needed to look at jumping times to understand the whole process. So we only need a sequence $(T_1, T_2, T_3, \dots)$. So we do not run into the same problem there we have with Brownian motion.

One idea (which does not work): Define $B(t), t \in \mathbb{Q}$ using the Kolmogorov extension theorem and extend the values continuously. But it is difficult to show that $\lim_{s \in \mathbb{Q} \to s_0} B(s)$ exists. So the correct idea is that we only get $B(t)$ for $t \in \mathbb{Z}[\frac{1}{2}]$ first, where $\mathbb{Z}[\frac{1}{2}] = \{m/2^{-n} : m, n \in \mathbb{Z}\}$ is the set of **dyadic rational numbers**.

Step 1: Using the Kolmogorov extension theorem, we can create a random list $C(x)$ for $x \in \mathbb{Q}_2$ such that

[1] Kolmogorov created the foundations for probability theory at the young age of 33.

- $C(0) = 0$,
- separate intervals are independent,
- $C(y) - C(x) \sim N(0, y - x)$ for $x, y \in \mathbb{Z}[\frac{1}{2}]$.

Theorem 2.1. *$C(x)$ is uniformly continuous.*

We will prove this next time. Using this, the next step is as follows.

Step 2: Let $\psi : UCF(\mathbb{Z}[\frac{1}{2}]) \to C[0, 1]$ send $C(x)$ to its unique continuous extension. Then let $\mathbb{P}_{BM} = \mathbb{P}_{CM} \circ \psi^{-1}$.

2.3 Gaussian random vectors

Before we construct Brownian motion, we need to understand a notion related to Gaussian random variables.

Definition 2.1. A **Gaussian random vector** is a random vector $X = (X_1, \dots, X_n)$ such that for all $y \in \mathbb{R}^n$, $X \cdot y$ is a Gaussian random variable.

The reason we care about this is that $(B(1), B(2), B(3), \dots)$ is a Gaussian random vector (i.e. its finite dimensional projections are Gaussian random vectors).

Proposition 2.3. *Let X be a Gaussian random vector with $\mathbb{E}[X] = 0$. If $\mathbb{E}[X_i X_j] = 0$, then $X_i \perp X_j$.*

This does not hold for general random vectors and will be important for us in our construction of Brownian motion.

3 Uniform Continuity of Brownian Motion

3.1 Brownian Motion is Uniformly Continuous on $\mathbb{Z}[\frac{1}{2}]$

Recall that we wanted to define Brownian motion by defining it on $\mathbb{Z}[\frac{1}{2}]$, the dyadic rational numbers. We have $B(x) - B(y) \sim N(0, y - x)$ if $y, x \in \mathbb{Z}[\frac{1}{2}]$ and $B(0) = 0$. Since $(B(t_1), \ldots, B(t_n))$ is a Gaussian vector, to show that $(B(2) - B(1)) \perp (B(3) - B(2))$, for example, we need only show that these are uncorrelated.

Lemma 3.1. *B on $\mathbb{Z}[\frac{1}{2}]$ is uniformly continuous.*

Here is the idea. Look at the interval $[0, 1]$, and say $[a, b] \subseteq [0, 1]$ is an interval of length ε. We know that $|B(b) - B(a)| \sim N(0, \varepsilon) \sim \sqrt{\varepsilon}$.

We need to prove that

$$\lim_{\varepsilon \downarrow 0} \sup_{\substack{|t-s| \le \varepsilon \\ t,s \in \mathbb{Z}[\frac{1}{2}] \cap [0,1]}} |B(t) - B(s)| = 0.$$

The "bad event" where this does not happen is the union of small/basic bad events $\{|B(t) - B(s)| \ge \delta\}$, but we can show that the probability of the big bad event is $\ll \sum \mathbb{P}(\text{small bad event})$.

Proof. For each m and $\gamma > 0$, define $G_m^\gamma = \{\nexists k \text{ s.t.} |B(k/2^m) - B((k-1)/2^m)| \ge 2^{-\gamma m}\}$. Let $H^N = \bigcap_{m \ge N}^\infty G_m^\gamma$.

Now let's bound $|B(y) - B(x)|$. Find an interval $I_1 = [a, b] \subseteq [x, y]$ of length 2^{-N}. If we are in H^N, then $|B(b) - B(a)| \le 2^{-\gamma N}$. Now let $I_2 = [a - 1/2^{N+1}, a]$ and $I_3 = [b, b + 1/2^{N+1}]$ if these are contained in $[x, y]$ and let them be $\varnothing$ otherwise. Then if we are in H^N, $|B(a) - B(a - 1/2^{N+1})| \le 2^{-\gamma(N+1)}$ and similarly for I_3. Proceeding like this, we can split the interval $[x, y]$ into intervals I_k of length 2^{-n} for $n \ge N$ with at most 2 of each kind. So $[x, y] = \bigcup_{n=1}^\infty I_n$, and we get

$$|B(y) - B(x)| \le 2 \sum_{m \ge N}^\infty 2^{m\gamma}.$$

So in H^N, if $|x - y| \le 2^{-(N-1)}$, then $|B(y) - B(x)| \le C 2^{-N\gamma}$ for some constant $C > 0$.

To get uniform continuity, we don't want to have fixed H^N. Using Markov's inequality with the fact that $\mathbb{E}[X]^4] = 1/2^m$ when $X \sim N(0, 1/2^m)$,

$$1 - \mathbb{P}(G_m^\gamma) = 2^m \mathbb{P}(|B(1/2^m) - B(0)| \ge 2^{-\gamma m}) \le \frac{2^{-2m}}{2^{-4\gamma m}} \le 2^{(-1+4\gamma)m}.$$

So $1 - \mathbb{P}(H^N)$ is bounded by a convergent geometric series. We will finish the proof next time. $\square$

4 Holder Continuity and Non-Lipschitz Continuity of Brownian Motion

4.1 Brownian Motion is Hölder Continuous on $\mathbb{Z}[\frac{1}{2}]$

Let's finish our proof of the following.

Lemma 4.1. *B on $\mathbb{Z}[\frac{1}{2}]$ is uniformly continuous a.s.*

Last time, we had a sequence of events H^N on which we had a condition bounding $|B(x) - B(y)|$.

Proof. In H^N, if $x, y \in \mathbb{Z}[\frac{1}{2}]$, and $|x-y| \leq 2^{-N}$, then $|B(x) - B(y)| \leq C2^{-\gamma N}$. We also know that $\mathbb{P}(H^N) \approx 1 - C2^{-\gamma N}$. Since $H^N = \bigcap_{m \geq N}^{\infty} G_m$, $H^N \subseteq H^{N+1}$. In H^N, for all $k \geq N$, if $|x-y| \in \mathbb{Z}[\frac{1}{2}]$ and $|x-y| \leq 2^{-k}$, then $|B(x) - B(y)| \leq 2^{-\gamma K}$. So in H^N, $B(t)$ is a uniformly continuous function.

Now, since $H^N \subseteq H^{N+1}$, we get $\mathbb{P}(\bigcup_{N=1}^{\infty} H^N) = \lim_{N \to \infty} \mathbb{P}(H^N) = 1$. So with probability 1, $B(t)$ is uniformly continuous on $\mathbb{Z}[\frac{1}{2}]$. □

The proof actually shows the following:

Corollary 4.1. *For any $\gamma < \frac{1}{2}$, $B(t)$ on $\mathbb{Z}[\frac{1}{2}]$ is a γ-Hölder continuous function a.s.*

Proof. In H^N, if $|x-y| \leq 2^{-N}$, then $|B(x) - B(y)| \leq |x-y|^{\gamma} C$. So in H^N for any $x, y \in [0,1]$ with $x, y \in \mathbb{Z}[\frac{1}{2}]$, $|B(x) - B(y)| \leq |x-y|^{\gamma} C_N$. So in $\bigcup H_n$, $B(t)$ is γ-Hölder continuous.

How did we find γ? We had $\mathbb{E}[B(t)^4] \sim t^2$, so we chose $\gamma < 1/4$. We can do better by using the $2p$-th moment: $\mathbb{E}[B(t)^{2p}] \sim t^p$, so we can pick $\gamma < (p-1)/(2p)$. □

Now that we have a uniformly continuous $B(t)$ on $\mathbb{Z}[\frac{1}{2}]$, we can extend it continuously to the entire positive real line.

4.2 Brownian motion is a.s. not Lipschitz anywhere

Lemma 4.2. *$B(t)$ is not Lipschitz at any point with probability 1.*

Here is the idea: Split $[0,1]$ into n intervals of length $1/n$. When the distance between two points is $1/n$, then $B(x) - B(y)$ should be about $1/\sqrt{n}$

Proof. Fix a constant C. Let $A_n := \{\exists x \in [0,1] \text{s.t. if } |y-x| \leq 3/n, \text{ then } |\frac{B(y)-B(x)}{y-x}| \leq C\}$. We want to show that $\mathbb{P}(\bigcap_n A_n) = 0$, and we have that $A_{n+1} \subseteq A_n$. So we want to prove that $\lim_n \mathbb{P}(A_n) = 0$. Define $A_n^{(x)}$ to be the event $\{|y-x| \leq 3/n, |\frac{B(y)-B(x)}{y-x}| \leq C\}$ In $A_n^{(x)}$, $|B(x+1/n) - B(x)| \leq C/n$; call this event $A_{n,1}^{(x)}$. Similarly, $|B(x+2/n) - B(x+1/n)| \leq 3C/n$;

call this event $A_{n,2}^{(x)}$. Going to the left of x, we also get $|B(x-1/n)-B(x)| \leq C/n$; call this event $A_{n,3}^{(x)}$. So we have $A_n^{(x)} \subseteq \bigcap_{i=1}^3 A_{n,i}^{(x)}$. Since the $A_{n,i}^{(x)}$ are independent,

$$\mathbb{P}(A_n) \leq \prod_{i=1}^{3} \mathbb{P}(A_{n,i}^{(x)})$$

We have that $\mathbb{P}(A_{n,i}^{(x)}) \sim 1/\sqrt{n}$, so $\mathbb{P}(A_n^{(x)}) \leq (1/n)^{3/2}$.

Let's extend this to all of A_n, not just a specific point. Let $a_{k,n} = k/n$. Then

$$\mathbb{P}\left(\bigcup_{k=1}^{n} A_n^{(a_{n,k})}\right) \leq n\mathbb{P}(A_n^{(a_{n,k})}) \sim \frac{1}{\sqrt{n}}.$$

The set of points where x is Lipschitz is open, so if $B(t)$ is Lipschitz at x, then it is Lipschitz at $a_{n,k}$ for some n, k. So we can bound the probability of A_n; the details are left as an exercise.[2] □

4.3 σ-fields for Brownian motion

Which σ-field do we use for Brownian motion? We have a few choices we can use:

$$\mathcal{F}_s^0 = \sigma(B(t), t \leq s)$$

$$\mathcal{F}_s^+ = \bigcap_{t>s} \mathcal{F}_t^0$$

The second of these choices contains the first, and it contains events that allow you to "see into the future" a little bit. But we will see that the only extra events $\mathcal{F}_s^+$ contains are null sets. Here is an example of an event in $\mathcal{F}_s^+$ but not in $\mathcal{F}_s^0$:

$$A = \left\{\limsup_{t\to s^+} \frac{B(t)-B(s)}{t-s} \geq \frac{1}{2}\right\}.$$

[2] :(

5 Brownian Motion Filtrations and Markov Property

5.1 Two filtrations for Brownian motion

Last time, we introduced the σ-fields

$$\mathcal{F}_s^+ := \bigcap_{t>s} \mathcal{F}_t^0, \qquad \mathcal{F}_s^0 := \sigma(B(s'), s' \leq s).$$

We have that $\mathcal{F}_s^0 \subseteq \mathcal{F}_s^+$, and these are not the same because we have an event in $\mathcal{F}_s^+ \setminus \mathcal{F}_s^0$:

$$\left\{ \limsup_{t \to s^+} \frac{B(t) - B(s)}{t - s} \geq \frac{1}{2} \right\}.$$

Later, we will show that the sets in $\mathcal{F}_s^+ \setminus \mathcal{F}_s^0$ differ from sets in $\mathcal{F}_s^0$ by null sets. In applications, we can use both of these σ-fields interchangeably. Usually, we use $\mathcal{F}_s^+$ because $\mathcal{F}_s^+ = \lim_{t \to s^+} \mathcal{F}_t^+ = \lim_{t \to s^+} \mathcal{F}_t^0$. That is, we want a right continuous filtration.

How do we show that $\mathcal{F}_s^+$ is almost the same as $\mathcal{F}_s^0$? For any bounded random variable Z on Ω, we can define the conditional expectations

$$\mathbb{E}[Z \mid \mathcal{F}_s^+], \qquad \mathbb{E}[Z \mid \mathcal{F}_s^0].$$

We will show that these are equal a.s. for any such Z.

Here is an application of this result (which is very difficult to prove without it).

Example 5.1. Fix s. What is

$$A_s = \mathbb{P}(\inf\{t > s : B(t) > B(s)\} = s\}?$$

This should be the same as

$$A_0 = \mathbb{P}(\inf\{t > 0 : B(t) > 0\} = 0).$$

Naively, you may assume that the answer should be $1/2$; but in fact, it is 1.

Proposition 5.1. $\mathbb{P}(\inf\{t > 0 : B(t) > 0\} = 0) = 1$.

Proof. Let B_0 be the event $\inf_{B(t)>0}\{t : t > 0\} = 0$, so $A_0 = \mathbb{P}(B_0)$. Then B_0 is $\mathcal{F}_0^+$-measurable. Then

$$\mathbb{E}[\mathbb{1}_{B_0} \mid \mathcal{F}_0^0] = \mathbb{E}[\mathbb{1}_{B_0} \mid \mathcal{F}_0^+] = \mathbb{1}_{B_0}.$$

But $\mathcal{F}_0^0 = \{\varnothing, \Omega\}$, since $B(0) = 0$. So $\mathbb{E}[\mathbb{1}_{B_0} \mid \mathcal{F}_0^0] = \mathbb{E}[\mathbb{1}_{B_0}]$. This gives us

$$\mathbb{1}_{B_0} = \mathbb{E}[\mathbb{1}_{B_0}].$$

So $\mathbb{P}(B_0) = 1$ or 0.

To show that $\mathbb{P}(B_0) \neq 0$, let C_0 be the event $\inf\{t > 0 : B(t) < 0\}$. Then $\mathbb{P}(C_0) = \mathbb{P}(B_0)$, and $\mathbb{P}(B_0 \cup C_0) = 1$. So $\mathbb{P}(B_0) > 0$. $\square$

5.2 Markov property of Brownian motion

Now let's prove this crucial result about $\mathcal{F}_s^+$ and $\mathcal{F}_s^0$. Here is some notation.

Let $Y : C_B(\mathbb{R}) \to [-M, M]$ for some $M > 0$. Then define $\mathbb{E}_x[Y] := Y(B(\cdot) + x)$ for $x \in \mathbb{R}$; this says that we input a Brownian motion with $B(0) = x$. Similarly, we define $\mathbb{E}_x[Y \mid \mathcal{F}] := \mathbb{E}[Y(B(\cdot) + x) \mid \mathcal{F}]$.

Recall the **shift operator** $\theta_s : C(\mathbb{R}) \to C(\mathbb{R})$ given by $\theta_s(f)(x) = f(x + s)$.

Lemma 5.1 (Markov property of Brownian motion)**.** *With the same notation as above,*

$$\mathbb{E}_x[Y \circ \theta_s \mid \mathcal{F}_s^+] = \mathbb{E}_{B(s)}[Y].$$

Remark 5.1. The right hand side is the expectation of $Y(\tilde{B}(\cdot))$, where $\tilde{B}$ is a Brownian motion independent of B such that $\tilde{B}(0) = B(s)$. So the right hand side is $\mathcal{F}_s^0$-measurable, while the left hand side is $\mathcal{F}_s^+$-measurable. The difficulty of proving this statement comes from this aspect.

Remark 5.2. Recall the similarity to the Markov property for Markov chains. There is a strong version of this property akin to the strong Markov property.

6 Markov Property of Brownian Motion

6.1 Proof of the Markov property of Brownian motion

We want to prove the following.

Lemma 6.1 (Markov property of Brownian motion)**.** *Let $Y : C(\mathbb{R}) \to [-M, M]$ be a functional, and let θ_s be the shift by s. Then*

$$\mathbb{E}_x[Y \circ \theta_s \mid \mathcal{F}_s^+] = \mathbb{E}_{B(s)}[Y].$$

Recall the following lemma:

Theorem 6.1 (Monotone class argument)**.** *Let $\mathcal{A}$ be a π-system, and let $\mathcal{H}$ be a collection of functions that satisfies:*

1. *If $A \in \mathcal{A}$, then $1_A \in \mathcal{H}$.*
2. *If $f, g \in \mathcal{H}$, then $f + g \in \mathcal{H}$.*
3. *If $f_n \uparrow f$ with $0 \le f_n \in \mathcal{H}$, then $f \in \mathcal{H}$.*

Then $\mathcal{H}$ has all bounded measurable functions with respect to $\sigma(\mathcal{A})$.

Now on to the proof of the Markov property.

Proof. We need to show that for any $A \in \mathcal{F}_s^+$ and for any Y, $\mathbb{E}[\mathbb{1}_A(Y \circ \theta_s)] = \mathbb{E}[\mathbb{1}_A \, \mathbb{E}_{B(s)}[Y]]$. By the monotone class argument, we only need to prove this for Ys of the form $Y = \prod_{k=1}^n f_k(B(t_k))$; here, $\mathcal{A}$ is the algebra generated by the $B(t)$s. Then

$$Y \circ \theta_s(B(\cdot)) = \prod_{k=1}^n f_k(B(t_k + s)).$$

The point of this trick now is that the difference between $\mathcal{F}_s^0$ and $\mathcal{F}_s^+$ doesn't matter because there is a gap between s and $s + t_1$.

For this Y, let's prove that that the result holds for any $A \in \mathcal{F}_{S+\delta}^0$ with $\delta \le t_1/2$. So we want to show that

$$\mathbb{E}[Y \circ \theta_s \mid \mathcal{F}_{s+\delta}^0] = \varphi_{B(s+\delta)}^Y$$

for some φ. We only need to show that this holds for $A \in \mathcal{A}_\delta$, where $\mathcal{A}_\delta = \{\bigcap_{k=1}^n \{B(a_k) \in R_k\} : a_k \in [0, s+\delta], R_k \text{ is Borel}\}$, because $\sigma(\mathcal{A}_\delta) = \mathcal{F}_{s+\delta}^0$. So we can calculate

$$\mathbb{E}[1_A(Y \circ \theta_s)] = \mathbb{E}\left[\prod_{k=1}^n 1_{\{B(a_k) \in R_k\}} \cdot \prod_{k=1}^n f_k(B(t_k + s))\right]$$

$$= \int_{R_1} p_{a_1}(0, \hat{a}_1)\, d\hat{a}_1 \int_{R_2} p_{a_2-a_1}(\hat{a}_1, \hat{a}_2)\, d\hat{a}_2 \cdots \int_{R_n} p_{a_n-a_{n-1}}(\hat{a}_{n-1}, \hat{a}_n)\, d\hat{a}_n$$
$$\cdot \int_R p_{s+\delta-a_n}(\hat{a}_n, \xi)\, d\xi \int_R p_{t-\delta}(\xi, \hat{t}_1) f_1(\hat{t}_1)\, dt_1 \cdot \int_R p_{t_2-t_1}(\hat{t}_1, \hat{t}_2) f_2(\hat{t}_2)\, d\hat{t}_2$$
$$\cdots \int_R p_{t_m-t_{m-1}}(\hat{t}_{m-1}, \hat{t}_m) f_m(\hat{t}_m) d\hat{t}_m,$$

where $p_t(x, y)$ is the pdf of a $N(0, t)$ random variable, $p_t(x, y) = \frac{1}{\sqrt{2\pi t}} e^{-|y-x|^2/2t}$. These integrals should be nested (evaluate them in reverse order).

$$= \mathbb{E}[\mathbb{1}_A \varphi^Y_{s+\delta}],$$

where $\varphi^Y_{s+\delta} \in \mathcal{F}^0_{s+\delta}$. In particular, this depends only on $B(s+\delta)$. If we send $\delta \downarrow 0$, this φ has a limit. This limit is exactly $\varphi^Y(B_s) = \mathbb{E}_{B(s)}[Y]$. □

6.2 Events at 0 and ∞ are trivial

Corollary 6.1. *$\mathcal{F}^+_0$ is trivial.*

Proof. $\mathcal{F}^+_0$ agrees with $\mathcal{F}^0_0$ mod null sets, but $\mathcal{F}^0_0$ is trivial. □

Corollary 6.2. *Events depending on $B(t)$ as $t \to \infty$ are trivial.*

Proof. Define $Y(t) = tB(1/t)$. Then $Y(t)$ is a Brownian motion because they are both have finite dimensional distributions which are Gaussian vectors that agree. But events with $t \to \infty$ are the same as events for $B(s)$ with $s \downarrow 0$. These are trivial. □

7 Equality of σ-Fields and Brownian Inversion

7.1 $\mathcal{F}_s^0$ and $\mathcal{F}_s^+$ are almost the same

Last time, we showed the Markov property for Brownian motion:

$$\mathbb{E}_x[Y \circ \theta_s \mid \mathcal{F}_s^+] = \mathbb{E}_{B(s)}[Y].$$

This is actually a bit stronger than a Markov property, since it uses $\mathcal{F}_s^+$, not $\mathcal{F}_s^0$.

Proposition 7.1. *$\mathcal{F}_s^+ = \mathcal{F}_s^0$ modulo null sets.*

Proof. We claim that $\mathbb{E}_x[Y \circ \theta_s \mid \mathcal{F}_s^0] = \mathbb{E}_{B(s)}[Y]$. The right hand side is $\mathcal{F}_s^0$-measurable, and the Markov property shows that it satisfies the definition of the conditional expectation. Then for any $\mathcal{F}$-measurable Z,

$$\mathbb{E}[Z \mid \mathcal{F}_s^+] = \mathbb{E}[Z \mid \mathcal{F}_s^0].$$

This follows from the monotone class argument, which tells us we only need to show it for $Z = \prod_{i=1}^k f(B(t_i))$. We can assume that $t_1 < t_2 < \cdots < r_m \leq s$ and $t_{m+1} > \cdots > t_k > s$. Then $Z = X \cdot (Y \circ \theta_s)$, where $X = \prod_{i=1}^m f(B(t_i))$ and $Y = \prod_{j=1}^{k-m} f(B(t_j - s))$. Then X is $\mathcal{F}_s^0$-measurable, so

$$\begin{aligned}
\mathbb{E}[Z \mid \mathcal{F}_s^0] &= \mathbb{E}[X(Y \circ \theta_s) \mid \mathcal{F}_s^0] \\
&= X\,\mathbb{E}[Y \circ \theta_s \mid \mathcal{F}_s^0] \\
&= X\,\mathbb{E}[Y \circ \theta_s \mid \mathcal{F}_s^+] \\
&= \mathbb{E}[X(Y \circ \theta_s) \mid \mathcal{F}_s^+] \\
&= \mathbb{E}[Z \mid \mathcal{F}_s^+].
\end{aligned}$$

□

7.2 $tB(1/t)$ is a Brownian motion

Last time, we mentioned the following property.

Proposition 7.2. *Let $Y(t) = tB(1/t)$. Then $Y(t)$ is a Brownian motion.*

Proof. $(Y(t_1), \ldots, Y(t_n))$ is a Gaussian random vector. So to prove that $Y(t_2) - Y(t_1) \perp Y_(t_4) - Y(t_3)$, for example, we only need to prove that they are uncorrelated. It now remains to show that we can define $Y(0) = 0$. □

We need to know that $\lim_{t\to\infty} \frac{B(t)}{t} = 0$ a.s.

Proposition 7.3. $\lim_{n\to\infty} \frac{B(n)}{n} = 0$. *a.s.*

Proof. $B(n) = \sum_{i=1}^n X_i$, where $X_i = B(n) - B(n-1)$. The X_i are iid with $N(0,1)$ distribution, so the strong law of large numbers gives the result. □

What if we want to find the following probability:

$$\mathbb{P}\left(\max_{m\in[n,n+1]}\frac{|B(m)-B(n)|}{n^{2/3}}\geq 1\right).$$

We can try looking at the following:

$$\mathbb{P}\left(\max_{m\in n+\mathbb{Q}_{[0,1]}^{(k)}}\frac{|B(m)-B(n)|}{n^{2/3}}\geq 1\right),$$

where $\mathbb{Q}_{[0,1]}^{(k)}=\{\ell/k\in[0,1]:k,\ell\in\mathbb{Z}\}$

You could try a union bound:

$$\leq\sum_{\ell=1}^{k}\mathbb{P}\left(\left|\frac{B(n+\ell/k)-B(n)}{n^{2/3}}\right|\geq 1\right).$$

However, this probability does not decay with k, and we have to add together k of them. So this will not work.

Let $X_\ell=B(n+\ell/k)-B(n+(\ell-1)/k)$, and let $Y_\ell=\sum_{\ell'<\ell}X_{\ell'}$. The X_ℓs are iid, so Y_ℓ is a Markov chain and a Martingale. We have the general inequality:

$$\mathbb{P}\left(\max_{1\leq\ell\leq k}|Y_\ell|\geq a\right)\leq\frac{\mathbb{E}[Y_k^2]}{a^2}$$

This gives us

$$\mathbb{P}\left(\max_{1\leq\ell\leq k}|Y_\ell|\geq n^{2/3}\right)\leq\frac{1}{n^{4/3}}.$$

Let $k=2^{\tilde{k}}$, and define the event $A_{\tilde{k}}=\{\max_{m\in n+\mathbb{Q}_k}|B(m)-B(n)|\leq n^{2/3}\}$. Then $A_{\tilde{k}}\supseteq A_{\tilde{k}+1}$. We also have that

$$\mathbb{P}(A_{\tilde{k}+1})\geq 1-n^{-4/3}.$$

So

$$\mathbb{P}\left(\bigcup_{\tilde{k}}A_{\tilde{k}}\right)\geq 1-n^{-4/3}.$$

This gives us

$$\mathbb{P}\left(\max_{m\in n+[0,1]}\left|\frac{B(m)-B(n)}{n^{2/3}}\right|\geq 1\right)\leq n^{-4/3}.$$

If we call this event C_n, we get that

$$\mathbb{P}(C_n\text{ i.o.})=0$$

by the first Borel-Cantelli lemma.

Together with the fact that $\lim_n\frac{B(n)}{n}\to 0$, we get:

Proposition 7.4. *With probability 1,*

$$\lim_{t\to\infty} \frac{B(t)}{t} \to 0.$$

Corollary 7.1. *The tail σ-field of Brownian motion is trivial.*

Proof. This follows from the fact that $\mathcal{F}_0^0 \equiv \mathcal{F}_0^+$, while $\mathcal{F}_0^0$ is trivial. $\square$

8 Tail Events, Limsups, and Stopping Times for Brownian Motion

8.1 Tail events do not depend on starting point

If $A \in \mathcal{T}$, we know that $\mathbb{P}_x(A) = 0$ or 1. On the other hand, this may depend on x (i.e. it is a function g of x). We want to show that for Brownian motion, $g(x) = g(0)$ for all x.

Example 8.1. Consider the event

$$A = \left\{ \limsup_{n \text{ is prime}} \frac{B(n)}{\sqrt{n}} \leq \frac{1}{2} \right\}.$$

Since A is in the tail σ-field $\mathcal{T}$, $\mathbb{1}_A = \mathbb{1}_D \circ \theta_1$ for a shifted event D. Then

$$\begin{aligned}
\mathbb{P}_0(A) &= \mathbb{E}[\mathbb{1}_A] \\
&= \mathbb{E}[\mathbb{E}[\mathbb{1}_A \mid \mathcal{F}_1]] \\
&= \mathbb{E}[\mathbb{E}[\mathbb{1}_A \circ \theta_1 \mid \mathcal{F}_1]] \\
&= \mathbb{E}[\underbrace{\mathbb{E}_{B(1)}[\mathbb{1}_D]}_{:=\varphi(B(1))}] \\
&= \int \varphi(a) p_1(0, a) \, da.
\end{aligned}$$

We know that $0 \leq \varphi(B(1)) \leq 1$. Since $p_1(0, a)$ is positive and we know that $\mathbb{P}_0(A) = 0$ or 1, we must have $\varphi(a) = \mathbb{P}_0(A)$ for a.e. a.

The same argument starting at x gives

$$\mathbb{P}_x(A) = \int \varphi(a) p_1(x, a) \, da = \mathbb{P}_0(A) \int p_1(x, a) \, da = \mathbb{P}_0(A).$$

8.2 Limsup of Brownian motion

Let's try to show that

$$\limsup_t \frac{B_t}{\sqrt{t}} = \infty \quad \text{a.s.}$$

By symmetry, this will also mean that

$$\liminf_t \frac{B_t}{\sqrt{t}} = -\infty \quad \text{a.s.}$$

Let

$$f(k) = \mathbb{P}\underbrace{\left(\frac{B_n}{\sqrt{n}} \geq k \right)}_{:=A_{n,k}} > 0.$$

This is independent of n because $B_n/\sqrt{n} \sim B(1)$ for all n. We then have

$$\mathbb{P}(A_{n,k} \text{ i.o.}) \geq \limsup \mathbb{P}(A_{n,k}) = f(k)$$

So

$$\mathbb{P}\left(\limsup_t \frac{B_t}{\sqrt{t}} \geq k\right) \geq f(k) > 0.$$

Since this is a tail event, it must then have probability 1.

8.3 Stopping times for Brownian motion

Define the σ-field $\mathcal{F}_s$ to be the smallest σ-field containing $\mathcal{F}_s^+$ and all the null sets. Let's now discuss the issue of stopping times. How should we define this?

We have

$$\{S < t\} = \bigcup_n \left\{S \leq t - \frac{1}{n}\right\} \in \mathcal{F}_t^0,$$
$$\{S \leq t\} = \bigcap_n \left\{S < t + \frac{1}{n}\right\} \in \mathcal{F}_t^+.$$

Since these only disagree on null sets, we are okay taking either definition.

Definition 8.1. A **stopping time** is a random variable $T : \Omega \to \mathbb{R}^+ \cup \{\infty\}$ such that $\{T < t\} \in \mathcal{F}_t$ for all $t \in \mathbb{R}_+$.

If $S^{(n)}$ are stopping times for all $n \in \mathbb{Z}$ and $S^{(n)} \searrow S$ a.s., then S is a stopping time. What does a.s. convergence mean in this context?

$$\mathbb{P}(\{\omega : \lim_{n\to\infty} S^{(n)}(\omega) = S(\omega)\}) = 1.$$

Proof. We can split up an event as

$$\{S < t\} = \bigcup_n \underbrace{\{S_n < t\}}_{\mathcal{F}_t}.$$

□

Remark 8.1. We have a similar result when $S^{(n)} \nearrow S$.

Proposition 8.1. *Let G be an open or closed set in $\mathbb{R}$. Then $S = \inf\{t : B_t \in G\}$ is a stopping time.*

Proof. If G is open,

$$\{S < t\} = \bigcup_{\substack{t' < t \\ t' \in \mathbb{Q}}} \underbrace{\{B_{t'} \in G\}}_{\in \mathcal{F}_t}.$$

Since this is a countable union, $\{S < t\} \in \mathcal{F}_t$.

If G is closed, we define $U_n = \bigcup_{x \in G} B(x, 1/n)$. Then we can define a stopping time S_n based on U_n, and $S_n \nearrow S$. So S is a stopping time. □

9 Strong Markov Property of Brownian Motion

9.1 σ-fields for stopping times

Recall, we said that a stopping time must satisfy $\{S < t\} \in \mathcal{F}_t$ for any t. We can define $S_n = \min\{\frac{k}{2^n} : \frac{k}{2^n} \geq S \geq \frac{k-1}{2^n}\}$. Then S_n is also a stopping time, and $S_n \searrow S$. Working with S_n is like working with stopping times in the discrete case. This is an example of a general technique: prove results for the discrete case and take a limit to transfer the result to the continuous case.

In the case of discrete time Markov chains, we define $\mathcal{F}_T = \{A : A \cap \{T \leq n\} \in \mathcal{F}_n \, \forall n\}$. Such events A can be expressed as $A = \bigcup_n A_n$, where $A_n = A \cap \{T = n\} \in \mathcal{F}_n$ for each n. The idea is that $\mathcal{F}_T$ is the information up to the stopping time T.

In the continuous case, we can define $\mathcal{F}_S = \{A : A \cap \{S \leq t\} \in \mathcal{F}_t \, \forall t\}$.

9.2 The strong Markov property of Brownian motion

Theorem 9.1 (strong Markov property of BM)**.** *Let $\{Y_a\}_{a \in \mathbb{R}}$ be a collection of functionals $C(\mathbb{R}) \to \mathbb{R}$, and let S be a stopping time. Then*

$$\mathbb{E}_0[Y_S \circ \theta_S \mid \mathcal{F}_S] = \mathbb{E}_{B(S)}[Y_S].$$

Remark 9.1. If we let $S = t$ be constant and set $Y_a = Y$, we get the Markov property we had before.

Example 9.1. Let $S := \inf(\{1\} \cup \{t : B_t \geq 1\})$. Let's find $\mathbb{E}[B_1 - B_S \mid \mathcal{F}_S]$. We can define $Y_S(f) = f(1 - S)$. Then $Y_S \circ \theta_S(f) = f(1)$. So

$$\mathbb{E}[B(1) \mid \mathcal{F}_S] = \mathbb{E}[Y_S \circ \theta_S \mid \mathcal{F}_S] = \mathbb{E}_{B(S)}[Y_S] = \mathbb{E}_{B(S)}[\tilde{B}(1 - S)],$$

where $\tilde{B}$ is an independent Brownian motion. If $\{S < 1\}$, then this is $\mathbb{E}_1[\tilde{B}(1 - S)] = 1$. The case for $\{S > 1\}$ will be discussed later.

Here is the idea of the proof:

Proof. We will prove that this is true for each $S_n \searrow S$. We need to show that for any $A \in \mathcal{F}_S$.

$$\mathbb{E}[\mathbb{1}_A \, \mathbb{E}_0[Y_S \circ \theta_S \mid \mathcal{F}_S]] = \mathbb{E}[\mathbb{1}_A \, \mathbb{E}_{B(S)}[\tilde{B}(1 - S)]].$$

Then we only need to show that this is true for the π-system of events $\{S < t\}$ and $\{B_u \leq x\}$. Then we simplify which Ys we want to prove this for using the monotone class argument. □

Remark 9.2. One subtlety to pay attention to is that $\mathcal{F}_S$ is not the same as $\mathcal{F}_S^+$; this is because S is random, so the situation is more complicated.

Example 9.2. We saw before that $\inf\{t > 0 : B_t = 0\} = 0$ a.s. If $B(0) = 1$, we will eventually hit 0 (at time $S = \inf\{t : B_t = 0\}$, say). Then we can ask the question of whether $\inf\{t > s : B_t = 0\} = 0$. The strong Markov property will let us answer this question. What should Y be?

10 Zeros of Brownian Motion, Time to Exceed a Value, and The Reflection Principle

10.1 Zeros of Brownian motion

Let $T_{t_0} = \inf\{t > t_0 : B(t) = 0\}$ be the first time to hit 0 after time t_0. Based on this, we can define $R_{t_0} = \inf\{t > T_{t_0} : B(t) = 0\}$, the next time to hit 0 after T_0.

Proposition 10.1. *For any t_0, with probability 1, $R_{t_0} = T_{t_0}$.*

Proof. We want to look at $F_{t_0} := \mathbb{1}_{R_{t_0} \neq T_{t_0}}$.

$$\begin{aligned}
\mathbb{E}[F_{t_0}] &= \mathbb{E}[\mathbb{E}[F_0 \circ \theta_{T_{t_0}} \mid \mathcal{F}_{T_{t_0}}]] \\
&= \mathbb{E}[\mathbb{E}[F_0 \mid \mathcal{F}_0]] \\
&= \mathbb{E}[F_0] \\
&= 0.
\end{aligned}$$

□

Define the event $A_q = \{R_q = T_q\}$, where $q \in \mathbb{Q}$. Then $\mathbb{P}(A_q) = 1$, so $\mathbb{P}(\bigcap_{q \in \mathbb{Q}} A_q) = 1$. This means that if $B(t) = 0$ for some t, then there exist either $s_n \nearrow t$ or $s_n \searrow t$ such that $B(s_n) = 0$ (with $n \in \mathbb{Z}$).

Corollary 10.1. *With probability 1, Brownian motion has no isolated zeros.*

10.2 First time to exceed a value

Let $x > 0$, and let $T_x = \min\{t > 0 : B(t) \geq x\}$ be the first time we exceed a. We can think of T_x as a function of x; this function is monotonically increasing. T_x has the property that its values are independent in separate intervals:

Proposition 10.2. *Let $a_1 \leq a_2 \leq b_1 \leq b_2$. Then $(T_{a_2} - T_{a_1}) \perp (T_{b_2} - T_{b_1})$.*

Proof. If $\mathbb{E}[F(A) \mid B]$ is constant for all functions F, then $A \perp B$. So want to show that

$$\mathbb{E}[F(T_{b_2} - T_{b_1}) \mid T_{a_2} - T_{a_1}]]$$

is constant for any function F. This will follow if we can prove that

$$\mathbb{E}[F(T_{b_2} - T_{b_1}) \mid \mathcal{F}_{T_{b_1}}]$$

is constant. We have, by the Strong Markov property, that

$$\mathbb{E}[F(T_{b_2} - T_{b_1}) \mid \mathcal{F}_{T_{b_2}}] = \mathbb{E}_0[F(T_{b_2 - b_1})].$$

□

What is the distribution of T_a for some fixed $a > 0$? This is difficult to prove on its own. The correct idea is to study T_x for varying x to find out things about T_a.

Define the **characteristic function** $\varphi_a(\lambda) = \mathbb{E}[e^{-\lambda T_a}]$. Then $\varphi_a(\lambda)\varphi_b(\lambda) = \varphi_{a+b}(\lambda)$ for all λ, as $T_a + T_b \stackrel{d}{=} T_{a+b}$. The unique solution to this kind of equation is $\varphi_a(\lambda) = e^{-ah_\lambda}$.

How do we find h_λ? Here is a trick: Define $Y_t = e^{\theta B_t - \theta^2 t}$; then Y_t is a Martingale, and $\mathbb{E}[Y_{t\wedge T}] = \mathbb{E}[T_{0\wedge T}] = \mathbb{E}[Y_0] = 1$ for any stopping time T. Then $\mathbb{E}[e^{\theta B_{T_a} - \theta^2 T_a}] = 1$, so we can find h_λ.

10.3 Reflection principle for Brownian motion

We can also ask the following question: What is

$$\mathbb{P}\left(\sup_{t\in[0,1]} B_t \geq a\right)?$$

Surprisingly, we get

$$\mathbb{P}\left(\sup_{t\in[0,1]} B_t \geq a\right) = 2\mathbb{P}(B_1 \geq a).$$

This is called the **reflection principle** for Brownian motion.[3] The idea is that if we hit a, we can reflect the rest of a path above and below the line $y = a$. These paths have the same probability of occurring. So we get

$$\begin{aligned}\mathbb{P}(B_1 \geq a) &= \mathbb{P}\left(B_1 \geq a \mid \sup_{t\in[0,1]} B_t \geq a\right)\mathbb{P}\left(\sup_{t\in[0,1]} B_t \geq a\right) \\ &= \frac{1}{2}\mathbb{P}\left(\sup_{t\in[0,1]} B_t \geq a\right).\end{aligned}$$

[3] Professor Yin has found this principle very useful in his research.

11 Distribution of The Last Zero in $[0,1]$ and Martingale Properties of Brownian Motion

11.1 First time to exceed a value

Proposition 11.1. *Let $T_a = \inf\{t > 0 : B_t \geq a\}$. Then*

$$\mathbb{P}(T_a < 1) = 2\mathbb{P}(B_1 > a).$$

Proof. Last time, we said

$$\mathbb{P}(B_1 > a) = \mathbb{P}(B_1 > a \mid T_a < 1) \cdot \mathbb{P}(T_a < 1).$$

So we want to show that $\mathbb{P}(B_1 > a \mid T_a < 1) = \frac{1}{2}$. We have

$$\begin{aligned}
\mathbb{E}[\mathbb{1}_{\{B_1 > B_{T_a}\}} \mid \mathcal{F}_a] &= \mathbb{E}[\mathbb{1}_{\{B_{1-T_a} > B_0\}} \circ \theta_{T_a} \mid \mathcal{F}_a] \\
&= \mathbb{E}_{B_{T_a}}[\mathbb{1}_{\{B_{1-T_a} > B_0\}}] \\
&= \mathbb{E}_a[\mathbb{1}_{\{B_{1-T_a} > a\}}].
\end{aligned}$$

This is $1/2$ if $T_a < 1$. □

Corollary 11.1. *Let Φ denote the CDF of the standard Gaussian distribution. Then*

$$\mathbb{P}(T_a < 1) = 2(1 - \Phi(a)).$$

11.2 Distribution of the last zero in $[0,1]$

We've shown that $\inf\{t > 0 : B_t = 0\} = 0$ a.s. What is the distribution of the last zero in $[0,1]$? Let $A = \sup\{0 \leq t \leq 1 : B_t = 0\}$. Then

$$\mathbb{P}(A \leq t) = \int p_t(0,y) \cdot \mathbb{P}(\text{no zeros between } t \text{ and } 1 \mid B(t) = y)\, dy.$$

By shifting the Brownian motion by t, the probability in the integrand is

$$\mathbb{P}_0(T_{-y} > 1 - t) = 2(1 - \Phi(y)).$$

After solving the integral, we get

$$\mathbb{P}(A \leq t) = \frac{2}{\pi} \arcsin(t).$$

Another related question: Let $a = \sup_{t \in [0,1]} B(t)$, and let $B(T_s) = a$. What is the distribution of T_s?

11.3 Martingale properties of Brownian motion

Definition 11.1. A random function $(X_t)_{t \geq 0}$ is a **martingale** (with respect to the filtration $\mathcal{F}_t$) if for all $t > s$, $\mathbb{E}[X_t \mid \mathcal{F}_s] = X_s$.

Equivalently, the condition is $\mathbb{E}[X_t - X_s \mid \mathcal{F}_s] = 0$.

Proposition 11.2. *$X_t = B_t^2 - t$ is a martingale.*

Proof. First, $X_t - X_s = B_t^2 - B_s^2 - (t-s)$. Then $B_t = B_s + Y$, where $Y \perp B_s$ and $Y \sim N(0, t-s)$. So

$$\mathbb{E}[X_t - X_s \mid \mathcal{F}_s] = \mathbb{E}[Y^2 + 2Y - (t-s) \mid \mathcal{F}_s] = \mathbb{E}[Y^2] + 2\,\mathbb{E}[Y] - (t-s) = 0. \qquad \square$$

Proposition 11.3. *Let $T = \inf\{t > 0 : B(t) \in \{a, b\}\}$. Then $\mathbb{E}[T] = -ab$.*

Proof. Since $X_t = B_t^2 - t$ is a martingale,

$$\mathbb{E}[B_T^2] - \mathbb{E}[T] = \mathbb{E}[X_T] = \mathbb{E}[X_0] = 0.$$

To find $\mathbb{E}[B_T^2]$, we have $\mathbb{E}[B_T^2] = \mathbb{P}(B_T = a)a^2 + \mathbb{P}(B_T = b)b^2$. Since B_t is a martingle, $\mathbb{E}[B_T] = 0$. So we can calculate

$$\mathbb{P}(B_T = a) = -\frac{b}{a}(1 - \mathbb{P}(B_T = a)) \implies \mathbb{P}(B_T = a) = \frac{b}{b-a}.$$

So we get

$$\mathbb{E}[T] = -\frac{a^2 b}{b-a} + \frac{b^2 a}{b-a} = -ab. \qquad \square$$

What if we want to find $\mathbb{E}[T^2]$? We can use another martingale with a B_t^4 term. Next time, we will talk about how to figure out such martingales involving Brownian motion.

12 Polynomial Brownian Motion Martingales and Arcsine Laws

12.1 Polynomial Brownian motion martingales

What kind of function of Brownian motion is a martingale? We want $\mathbb{E}[f(t, B_t) \mid \mathcal{F}_s] = f(s, B_s)$ for all $t > s$. We can also state this as $\mathbb{E}[f(t, B(t)) - f(s, B(s)) \mid \mathcal{F}_s]$.

Proposition 12.1. *If f is a polynomial, and*

$$\frac{\partial f}{\partial t} + \frac{1}{2} f_{xx} = 0,$$

then $f(t, B_t)$ is a martingale.

Remark 12.1. This is not the heat equation, but it is similar. The heat equation looks like $\frac{\partial f}{\partial t} - \frac{1}{2} f_{xx} = 0$. If we let $p_t(x, y) = f_{B_t \mid \{B_0 = x\}}(y)$, then $p_t(0, x)$ satisfies the heat equation.

Remark 12.2. For high-dimensional Brownian motion, the formula should be

$$\frac{\partial f}{\partial t} + \frac{1}{2} \Delta f = 0.$$

How do we think of $p_t(x, y)$. Certainly, $\int p_t(0, y)\, dy = 1$. Here is how physicists think about it. If we have 1 pound of sand at $t = 0$, we can move the sand around randomly according to Brownian motion. Then at time $t = t_0$, $p_{t_0}(0, x)$ is the density of sand at x. The fact that $p_t(x, y)$ satisfies the heat equation explains why the variance of $p_t(0, y)$ spreads out as t grows (the probability spreads out like heat).

If f is a martingale, we get $\mathbb{E}[f(t, B_t) - f(0, B_0)] = 0$. What does this mean in physics? This is like

$$\int_{\text{sand}} f(t, Bt) = f(0, B_0).$$

Let $f_n = f(t, \text{position of sand particle } n)$. Then

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} f_n(t) = f(0) = f(0, B_0).$$

What fs satisfy this condition? If f is constant or linear with respect to position, this condition holds. If you want a 2nd derivative condition, then you need $\frac{\partial f}{dt} + \Delta f = 0$.

Proof. We have $\mathbb{E}[f(t, B_t) - f(0, B_0)] = 0$. This is

$$\int f(t, y) p_t(x, y)\, dy - f(0, x) = 0.$$

If we take the derivative with respect to t, we get

$$\int f_1(t,y)p_t(x,y)\,dt + \int f(t,y)\frac{\partial p_t}{\partial t}(x,y) = 0.$$

By integration by parts, we get

$$\int (f_1 + \frac{1}{2}f_{2,2})p_t(x,y) = 0.$$

So $f_1 + \frac{1}{2}f_{2,2} = 0$. $\square$

Remark 12.3. It is not necessary for f to be a polynomial. $f = e^{\theta B_t - \frac{1}{2}\theta^2 t}$ is also a martingale.

12.2 Arcsine laws and time of the maximum in Brownian motion

Let T be the first time such that $B_T = \sup_{t\in[0,1]} B_t$. Last time, we learned that the last zero of Brownian motion in $[0,1]$ is distributed like arcsine. There are two other Brownian motion arcsine laws:

1. $|\{t : B(t) > 0, t \in [0,1]\}|$,
2. T as defined above.

The way to calculate T is to first find the joint density of (T, M), where $M = \sup_{t\in[0,1]} B_t$. Let $M(t) = \sup_{s\in[0,t]} B_s$, and let $X_t = B(t) - M(t) \leq 0$. We can also consider $Y_t = -|\tilde{B}(t)|$, which is a different Brownian motion. We claim that $X_t \overset{d}{=} Y_t$. Here is a heuristic argument

1. First, we have $|\{t : B(t) = M(t))\}| = 0$.
2. Next, if $B(t_0) \neq M(t_0)$, then there are an interval I and $t_0 \in I$ such that $B(t) - M(t)$ looks like a Brownian motion in I.
3. Now T for $B(t)$ is the last zero for $\tilde{B}(t)$. This is because the last zero of $\tilde{B}(t)$ and the last zero of Y_t have the same distribution. And Y_t and X_t have the same distribution.

The idea to prove this is to use a random walk. If we take a limit of scaled random walks, we will eventually get Brownian motion. We will go over this next time, in a result called Donsker's theorem.

If S_n is the result of a simple random walk on $\mathbb{Z}$ at time n, then let $X_n = S_n - M_n$. If $X_{n-1} \neq 0$, then $X_n = X_{n-1} \pm 1$ with probability 1/2 each. If $X_{n-1} = 0$,

$$X_n = \begin{cases} -1 & \text{with probability } 1/2, \\ 0 & \text{with probability } 1/2. \end{cases}$$

Then $Y_n = -|\tilde{S}_n|$ has the same distribution as X_n. This result will extend to Brownian motion.

13 Skorokhod's Representation Theorem

13.1 Skorokhod's representation theorem

We will study two new ways of understanding Brownian motion:

1. Brownian motion as a limit of simple random walks
2. Orthogonal polynomial method ot construct Brownian motion

The first one lets us prove results about Brownian motion using combinatorial arguments. The OPM is useful for computer simulation of Brownian motion and other applications.

Theorem 13.1 (Skorokhod's representation theorem)**.** *Let X be a real-valued random variable with $\mathbb{E}[X] = 0$. There exists a family of stopping times T_α with respect to $B(t)$ (where α is a random label) such that $B(T_\alpha) \overset{d}{=} X$ and $\mathbb{E}[T_\alpha] = \mathbb{E}[X^2]$.*

Example 13.1. Let $X = \pm 1$ with probability $1/2$ each. Let $T = \inf\{t : |B(t)| \geq 1\}$. Then $B(T) \sim X$ and $\mathbb{E}[T] = \mathbb{E}[X^2]$.

Example 13.2. Let $X = \pm 2$ with probability $1/2$ each. Then we can take $T = \inf\{t : |B(t)| \geq 2\}$.

Example 13.3. Let $X = \pm 1, \pm 2$ with probability $1/4$ each. Let $T_k = \inf\{t : |B(t)| \geq k\}$. Let $\alpha = 1$ or 2 with probability $1/2$ each. Then $B(T_\alpha) \overset{d}{=} X$.

Here is the outline of the proof.

Proof. Step 1: If $\mathbb{P}(X = a \text{ or } b) = 1$, then let $T_{a,b} = \inf\{t : B(t) = a \text{ or } b\}$.

Step 2: We want a random variable $\alpha : \Omega \to \mathbb{R}^2$ with a distribution such that $B(T_\alpha) \overset{d}{=} X$ and $\mathbb{E}[T_\alpha] = \mathbb{E}[X^2]$. In the discrete case, we have

$$\mathbb{P}(B(T_\alpha) = u) = \mathbb{E}_\alpha[\mathbb{P}_{\mathrm{BM}}(B(T_{u,v}) = u)] = \mathbb{E}_\alpha\left[\frac{v}{|u - v|}\right].$$

$\square$

13.2 Proof of CLT using Skorohod's representation theorem

If we have the SLLN and this representation theorem, we can actually produce a proof of the central limit theorem.

Corollary 13.1 (CLT). *Suppose that X_n are iid random variables with $\mathbb{E}[X_i] = 0$. Then*

$$\frac{\sum_{n=1}^N X_n}{\left(\sum_{n=1}^N X_n^2\right)^{1/2}} \xrightarrow{d} \mathcal{N}(0,1).$$

Proof. By Skorokhod's representation theorem, $X_1 \overset{d}{=} B(T_\alpha)$ and $X_2 \overset{d}{=} \tilde{B}(T_{\tilde{\alpha}})$, where $\alpha, \tilde{\alpha}$ are iid and $B \perp \tilde{B}$. Then $X_1 + X_2 \overset{d}{=} B(T_\alpha + \tilde{T}_{\tilde{\alpha}})$, where $\tilde{T}_{\tilde{\alpha}} = \inf\{t - T_\alpha : t > T_\alpha, B(t) - B(T_\alpha) \in \tilde{\alpha}\}$. (Recall $T_\alpha = \inf\{t > 0 : B(t) \in \alpha\}$.) The reason we can do this is that $B(T_\alpha + \tilde{T}_{\tilde{\alpha}}) - B(T_\alpha) \overset{d}{=} B(T_{\tilde{\alpha}})$. In fact, $T_\alpha, \tilde{T}_{\tilde{\alpha}}$ are iid.

We can extend this to $X_1 + X_2 + \cdots + X_n \overset{d}{=} B(T^1_{\alpha_1} + T^2_{\alpha_2} + \cdots + T^n_{\alpha_n})$. By the SLLN, $\sum_n T^n_{\alpha_n} \to N\,\mathbb{E}[T^1_{\alpha_1}] = N\,\mathbb{E}[X^2]$. So we have $X_1 + X_n \overset{d}{=} B(Y_N)$, where $Y_n \to N\,\mathbb{E}[X_1^2]$ a.s. So

$$\frac{X_1 + \cdots + X_N}{\sqrt{N}} \xrightarrow{d} \mathcal{N}(0, \mathbb{E}[X^2]).$$ □

13.3 Brownian motion as a limit of simple random walks

Let $X_i \sim$ iid Ber(1/2), and let

$$S^N_X = \begin{cases} \sum_{k=1}^m X_k & \text{if } X = m, x \in \mathbb{N} \\ \text{linear combination of } S^n_{[x]}, S^n_{[x]+1} & x \notin \mathbb{N}. \end{cases}$$

In other words, we linearly interpolate between the values of a random walk. This gives us a graph (i.e. a random continuous function $\mathbb{R}_+ \to \mathbb{R}$). Then let

$$f^N(t) = \frac{S^N_{tN}}{\sqrt{N}}.$$

Then f^N converges in distribution to Brownian motion on $[0,1]$.

Usually convergence in distribution is not so strong. Next time, we will talk about how to improve this for our Brownian motion.

14 Donsker's Theorem

14.1 Donsker's theorem

Let S_m^n be a simple random walk ($1 \leq m \leq n$). $f^{(n)}(t)$ is defined with S_m^n by linearly interpolating between the values and rescaling. Then $f^{(n)} : [0,1] \to \mathbb{R}$ is a random function.

Theorem 14.1 (Donsker)**.** $f^{(n)} \xrightarrow{d} B(\cdot)$.

Remark 14.1. That is, we can consider $\mathbb{P}_{f^{(n)}}$, which is a measure on $C([0,1])$. This converges weakly to $\mathbb{P}_{B(\cdot)}$ as a measure on $C([0,1])$. When we talk about weak convergence of measures, we mean $\mu_n(A) \to \mu(A)$ for all open, measurable A. Say Brownain motion is associated to the space $(\Omega, \mathcal{F}, \mathbb{P})$. Then the limiting measure here is $\mathbb{P}|_{\mathcal{F}(C([0,1]))}$.

This only talks about convergence of the distributions as measures on $C([0,1])$. We can see that $|\{x_0 : \frac{d}{dx} f^{(n)}(x_0) \text{ exists}\}| = 1$, but $|\{x : B'(x) \text{ exists}\}| = 0$. The issue is that $\{\omega : [g'(t_0)](\omega) \text{ exists}\}$ is not an open set in $C([0,1])$.

14.2 Applications of Donsker's theorem

Corollary 14.1. *If $\psi : C([0,1]) \to \mathbb{R}$ is continuous $\mathbb{P}_0$-a.s. (which is the $\mathbb{P}$ for Brownian motion), then $\psi(f_n) \xrightarrow{d} \psi(B)$.*

Corollary 14.2. *Let $M(f) = \max_t f(t)$. Then $M(f_n) \xrightarrow{d} M(B)$.*

Proof. M is continuous on $C([0,1])$. □

Corollary 14.3. *Let $P_0(f) = |\{x : f(x) > 0\}|$. Then $P_0(f_n) \xrightarrow{d} P_0(B)$.*

Remark 14.2. P_0 is not continuous in $C([0,1])$.

Proof. The set of points where P_0 is continuous has probability 1 for $\mathbb{P}_{\text{BM}}$. Indeed, if $|\{x : g(x) = 0\}| = 0$, then P_0 is continuous at g. □

Similarly, we can use Donsker's theorem to find the last zero of Brownian motion in $[0,1]$.

14.3 Proof of Donsker's theorem

Here is the proof of Donsker's theorem. The idea is to "grow a simple random walk on Brownian motion."

Proof. Let $S_m^n = \sum_{k=1}^m X_k$. We know that $f_n(t) \overset{d}{\approx} B(t)$ for $t = k/n$, but this is hard to deal with. The correct idea is to look at $n^{-1/2} S_m^n = n^{-1/2} \sum_{k=1}^m X_k = B(\tau_1^n + \cdots + \tau_m^n)$, where $\tau_{i+1} = \inf\{t - \tau_i > 0 : |\tilde{B}(t) - \tilde{B}(\tau_i)| \geq 1\}$. Depending on whether the Brownian motion goes up or down, we can tell the random walk to go up or down.

Here is the idea of how the convergence works: Now all $f^{(n)}$ grow on the same B. So the $f^{(n)}$ share good properties of B. If $\mathcal{A}_{\varepsilon,\delta} = \{B : |x-y| \leq \delta \implies |B(x) - B(y)| \leq \varepsilon\}$. In $\mathcal{A}_{\varepsilon,\delta}$, for all large enough n, f_n is ε, δ-continuous. $\square$

15 Orthogonal Polynomial Method for Constructing Brownian Motion

15.1 Overview

This lecture, we will talk about a method used to construct Brownian motion which is good for computers to simulate. Previously, we constructed Brownian motion on dyadic rational numbers and extended it continuously to $\mathbb{R}$. We also constructed Brownian motion as a limit in distribution of scaled simple random walks using Donsker's theorem.

Here are the basic properties of Brownian motion:

1. $B(0) = 0$.
2. $B(t) \sim N(0, t)$.
3. $B(I_1) \perp B(I_2)$ if $I_1 \cap I_2 = \varnothing$, where $B(I) = B(b) - B(a)$ if $I = [a, b]$.
4. $B(\cdot) \in C([0, 1])$.

We will construct a sequence of random functions $f^{(n)}$ that converges to Brownian motion in the $\|\cdot\|_\infty$ sense. The limiting random function will satisfy the above 4 properties.

15.2 Orthogonal functions

Given functions ψ_n, we have the inner product

$$\langle \psi_n, \psi_m \rangle := \int_0^1 \psi_n(x)\psi_m(x)\, dx.$$

We will construct an orthonormal set of functions ψ_n (i.e. $\langle \psi_n, \psi_m \rangle = 0$, $\langle \psi_n, \psi_n \rangle = 1$).[4]

Let ψ_n be an orthonormal basis of $L^2([0, 1])$. Then if $\psi \in L^2$, we have

$$\phi = \sum_k \langle \phi, \psi_k \rangle \, \psi_k.$$

Now define

$$W_n(t) = \sum_{k=1}^n X_k \cdot \int_0^t \psi_k(s)\, ds, \qquad X_k \sim \text{iid } \mathcal{N}(0, 1).$$

Proposition 15.1. *For every $t \in [0, 1]$, $W_n(t)$ is Gaussian with $\mathbb{E}[W_n(t)] = 0$ and variance $\mathbb{E}[W_n^2(t)] \xrightarrow{n\to\infty} t$.*

[4] Orthogonal polynomials are very useful in random matrix theory.

Proof. For the variance,

$$\begin{aligned}
\mathbb{E}[W_n(t)^2] &= \sum_{k=1}^{n} \left| \int_0^t \psi_k(s)\, ds \right|^2 \\
&= \sum_{k=1}^{n} \left| \left\langle \mathbb{1}_{[0,t]}, \psi_k \right\rangle \right| \\
&\xrightarrow{n\to\infty} \left\langle \mathbb{1}_{[0,t]}, \mathbb{1}_{[0,t]} \right\rangle \\
&= t.
\end{aligned}$$

□

Furthermore, $W_n(t) - W_n(s) \xrightarrow{d} \mathcal{N}(0, t-s)$, and $(W_n(I_1), W_n(I_2)) \xrightarrow{d} (B(I_1), B(I_2))$, where B is a Brownian motion.

15.3 Haar functions

We want to say that $W_n(\cdot) \xrightarrow{d} B(\cdot)$. But the limit of the left hand side will be an L^2 funciton, not necessarily a continuous function. Can we find some orthonormal basis of polynomials such that $W_n(\cdot)$ has a $\|\cdot\|_\infty$ limit in $C([0,1])$? Recall the following result:

Proposition 15.2. *Let $f_n \in C([0,1])$ for each n. If $f_n \to f$ uniformly, then $f \in C([0,1])$.*

Now we only need to find an orthonormal basis such that $W_n(\cdot)$ are a Cauchy sequence in $C([0,1])$.

Definition 15.1. The **Haar functions** are the functions

$$f_k^n(x) = \begin{cases} 2^{n/2} & x \in [2^{-n}k, 2^{-n}k + 2^{-(n+1)}] \\ -2^{n/2} & x \in (2^{-n}k + 2^{-(n+1)}, 2^{-n}(k+1)]. \end{cases}$$

Now define $\psi_{2^n+k} = f_k^n$, and let $\widetilde{W}_n = W_{2^n}$. Then

$$\|\widetilde{W}_{n+1} - \widetilde{W}_n\| \le 2^{-n/2} \max\{|X_{2n}|, \dots, |X_{2^{n+1}}|\}.$$

We have that $\mathbb{P}(X_i > n) \sim e^{-n^2/2}$, so

$$\mathbb{P}(\max\{x_1, \dots, X_{2^n}\}) \lesssim e^{-n^2/4}$$

for large n. So

$$\|\widetilde{W}_{n+1} - \widetilde{W}\|_\infty \le 2^{-n+1} n$$

with probability $1 - e^{-n^2/4}$. Using the Borel-Cantelli lemma, we get that the probability these events don't hold infinitely often equals 0.

This method is good for calculating things such as the distribution of the last zero of Brownian motion in $[0,1]$.

16 Itô's Formula

16.1 Integrating with respect to Brownian motion

Let $f \in C^2([0,1])$. We want to say something like

$$f(B_t) = f(0) + \int_0^t f'(B_s)\, dB_s.$$

Itô's formula tells us that there is actually an extra term:

$$f(B_t) = f(0) + \int_0^t f'(B_s)\, dB_t + \frac{1}{2}\int_0^t f''(B_s)\, ds.$$

If we change a $f(B_t)$ a little bit, how can we measure this change? The answer is

$$df(B_t) = f'(B_t)\, dB_t + \frac{1}{2} f''(B_t)\, dt.$$

Some people also write

$$\Delta f(B_t) = (d \to \Delta)$$

for this in physics and finanical contexts.

There is something that is not clear in the above equation. What does the integral $\int_0^t f'(B_s)\, dB_s$ mean?

Definition 16.1. We can define an integral with respect to Brownian motion by

$$\int_0^s g(B_t)\, dB_t := \lim_{0=t_1<\cdots<t_n=s} \sum_{k=1}^n g(B(t_k))(B(t_{k+1}) - B(t_k))$$

This is very similar to a Riemann sum, except a Riemann sum has $g(t_k^*)$, where $t_k^* \in [t_k, t_{k+1}]$. In our definition, we explicitly pick t_k, instead. In fact, the limit will change if we replace t_k with $st_k + (1-s)t_{k+1}$ for some $s < 1$.

Example 16.1. Here is an example to show that the limit can be different if we change t_k to t_{k+1}.

$$\sum_k [g(B(t_{k+1})) - g(B(t_k))](B(t_{k+1}) - B(t_k)) \approx \sum_k g'(B_{t_k})(B(t_{k+1}) - B(t_k))^2$$

Suppose $g' \approx 2$. Then this is

$$\begin{aligned} &\approx 2 \sum_k (B(t_{k+1}) - B(t_k))^2 \\ &\approx 2 \sum_k t_{k+1} - t_k \\ &= 2s. \end{aligned}$$

16.2 Examples and applications

Example 16.2. Say you buy stocks every day. Let A_n be the number of stocks you have on day n, and let $\Delta B_n = (B_{n+1} - B_n)$ be the change of stock price on day n. Then we want to calculate

$$\sum_{n=1}^{365} A_n \cdot \Delta B_n$$

We have that $A_n \in \mathcal{F}_n$, where $\mathcal{F}_n = \sigma(B_1, \ldots, B_n)$. We also assume that as $n \to \infty$, B_n looks like a Brownian motion. In the limit, we get

$$\int A_t \, dB_t,$$

where A_t is $\mathcal{F}_t$-measurable ($\mathcal{F}_t$ is our σ-field for Brownian motion).

Example 16.3. Itô's formula implies

$$\sin(B_t) = \int \cos(B_s) \, dB_t - \frac{1}{2} \int \sin(B_s) \, ds.$$

Often, we use Itô's formula backwards, to find the value of the integral.

Example 16.4. We have

$$B_t^2 = \int_0^t B_s \, dB_s + t,$$

so we can solve to get

$$\int_0^t B_s \, dB_s = B_t^2 - t.$$

In fact, we have the following theorem:

Theorem 16.1. *Let $g \in C^2$. Then*

$$F(t) = \int_0^t g(B_s) \, dB_s$$

is a martingale.

We will prove this later.

16.3 Proof of Itô's formula

Theorem 16.2 (Itô's formula)**.** *Let* $f \in C^2([0,1])$. *Then*

$$f(B_t) = f(0) + \int_0^t f'(B_s)\, dB_t + \frac{1}{2}\int_0^t f''(B_s)\, ds.$$

Here is a heuristic argument:

Proof. We have

$$f(B_t) = f(B_{t_0}) + f'(B_{t_0}) \underbrace{(B_t - B_{t_0})}_{\approx (\Delta t)^{1/2}} + \frac{1}{2} f''(B_{t_0}) \underbrace{(B_t - B_{t_0})}_{\approx \Delta t} + \cdots,$$

where $\Delta t = t - t_0$. We can get rid of the terms with order > 2, since they will be small when Δt is small. Then

$$f(B_t) - f(B_0) = \sum_k f'(t_k) B(\Delta_k) + \frac{1}{2}\sum_k f''(B_{t_k})(B(\Delta_k))^2,$$

$$B(\Delta_k) := B(t_{k+1}) - B(t_k).$$

When we take the limit, since the $B(\Delta_k)$ are independent of each other, so we get

$$\int f(B_s)\, dB_s + \lim \frac{1}{2}\sum_k f''(B_{t_k}) A_k,$$

where the A_k are independent $\mathcal{N}(0, \Delta_k)^2$. The law of large numbers makes the right hand side approximately $\frac{1}{2}\sum_k f''(B_{t_k})\, \mathbb{E}[A_k]$, so the right hand side converges to $\int f''(B_s)\, ds$. □

What if we want to prove this for $f \in C^2(\mathbb{R})$? We prove it for when $\|f'\|_\infty, \|f''\|_\infty$ are bounded. Then with high probability, $\|B_t\|_{L^\infty(0,s)} < \infty$, so we can extend to the general case.

17 Strengthening Itô's Formula, and Generalizing the Itô Integral

17.1 Strengthened Itô's formula

Here is the stronger version of Itô's formula.

Theorem 17.1 (Itô's formula). *Let $f \in C^2(\mathbb{R})$ with $\|f'\|_\infty, \|f''\|_\infty \leq M$. Then*

$$f(B_t) = f(0) + \int_0^t f'(B_s)\, dB_t + \frac{1}{2}\int_0^t f''(B_s)\, ds.$$

We want to show that

$$\sum_{\{t_k\}} f'(B_{t_k})B(\Delta_k) \to \int f(B_s)\, dB_s$$

$$\sum_{\{t_k\}} f''(B_{t_k})B(\Delta_k)^2 \to \int f(B_s)\, dB_s$$

where $B(\Delta_k) := B(t_{k+1}) - B(t_k)$.

For the first statement, we need to show that $\sum_{\{t_k\}} f'(B_{t_k})B(\Delta_k)$ has a limit when the mesh size of $\{t_k\} \to 0$. What does this convergence mean? This is not just a sequence; it is a net.

Definition 17.1. A **net** is a partially ordered collection $\{X_\alpha\}_{\alpha \in I}$ such that for any α, β, there exists a γ such that $\gamma > \alpha$ and $\gamma > \beta$.

For partitions, we have a net: for $T = \{t_k\}$ and $\hat{T} = \{\hat{t} - k\}$, $T \leq \hat{T} \iff \{t_k\} \subseteq \{\hat{t}_k\}$.

Proof. We want to show that this net is Cauchy. Write $B(\Delta_k) = \sum_{\hat{\ell}} B(\Delta_{\hat{\ell}})$, where the $\hat{t}_\ell \in [t_k, t_{k+1}]$. We have

$$\mathbb{E}\left[\sum_k f'(B_k)B(\Delta_k) - \sum_{\hat{k}} f'(B_{\hat{k}})B(\Delta_{\hat{k}})\right]^2 = \mathbb{E}\left[\sum_k \sum_{\hat{t}_\ell \in [t_k, t_{k+1}]} [f'(B_k) - f'(B_{\hat{\ell}})]B(\Delta_{\hat{\ell}})\right]^2$$

$B(\Delta_{\hat{\ell}})$ is independent of random variables measurable with respect to $\mathcal{F}_{\hat{t}_\ell}$, so a lot of the terms cancel (because they have zero expectation).

$$= \mathbb{E}\left[\sum_k \sum_{\hat{t}_\ell \in [t_k, t_{k+1}]} [f'(B_k) - f'(B_{\hat{\ell}})]^2 B(\Delta_{\hat{\ell}})^2\right]$$

$$= \mathbb{E}\left[\sum_k \sum_{\hat{t}_\ell \in [t_k, t_{k+1}]} |f'(B_k) - f'(B_{\hat{\ell}})|^2 \Delta_{\hat{\ell}}\right]$$

$$\leq \sup_{\substack{s,t \in [0,T] \\ |s-t| \leq \max t_{k+1} - t_k}} \mathbb{E}[|f'(B_s) - f'(B_t)|]^2 \cdot T$$

This goes to 0 if $\max_k (t_{k+1} - t_k) \to 0$. So this is a Cauchy net, and thus it has a limit.

For the second statement we want to prove, compare

$$\sum_k f''(B_{t_k}) B^2(\Delta_k), \sum_k f''(B_{t_k}) \Delta_k.$$

The right term has the limit $\int f''(B_s)\, ds$. Subtracting these two gives

$$\sum_k f''(B_{t_k})[B(\Delta_k)^2 - \Delta_k]$$

Then $f''(B_{t_k})$ is $\mathcal{F}_{t_k}$-measurable, and $B(\Delta_k)^2 - \Delta_k$ is independent of $\mathcal{F}_{t_k}$-measurable random variables and has zero expectation.

In general, for these kinds of sums, we have

$$\mathbb{E}\left[\sum_k g_k h_k\right] = \sum_{k,k'} \mathbb{E}[g_k g_{k'} h_k h_{k'}]$$

If $k' > k$, then $g_k g_{k'} h_{k'} \in \mathcal{F}_{t_{k'}}$. But $h_k \perp \mathcal{F}_{t'_k}$.

$$\begin{aligned}
&= \sum_k \mathbb{E}[g_k^2 h_k^2] \\
&\leq M^2\, \mathbb{E}\left[\sum_k h_k^2\right] \\
&\leq M^2 \cdot T \cdot C \cdot \max_k (\Delta_k).
\end{aligned}$$

So we get that $\sum_k f''(B_{t_k})[B(\Delta_k)^2 - \Delta_k] \to 0$ in L^2, which implies convergence in probability. $\square$

17.2 The Itô integral for more general functions

We have defined $\int f(B_s)\, dB_s$ when $f \in C^2$. What if f depends on the entire path of Bronian motion until time t? The general case is

$$\int_0^T f(\omega, s)\, dB_s, \qquad \omega \in \Omega_B.$$

Example 17.1. Let $f(\omega, t) = \max_{s \leq t} |B_s(\omega)|$. Then we are looking at

$$\int_0^T \max_{s \leq t} |B_s| \, dB_t.$$

What kind of function should f be?

1. If Brownian motion has the measure space $(\Omega_B, \mathcal{F}_B, \mathbb{P}_B)$, then f should be $\mathcal{F}_B \otimes \mathcal{T}$-measurable.
2. For fixed t, $f(\omega, t) \in \mathcal{F}_t$.
3. $f \in L^2(\Omega_B \times [0, T])$: that is, $\mathbb{E}[\int f^2(\omega, t) \, dt] < \infty$.

We let $\mathcal{H}$ be the collection of f satisfying these 3 properties. We start from

$$f(\omega, t) = a(\omega) \cdot \mathbb{1}_{[t_1, t_2]}(t).$$

What property should $a(\omega)$ satisfy to satisfy property (2) above? We want $a \in \mathcal{F}_{t_1}$.

18 The Itô Integral for L^2 Functions

18.1 The Itô integral for simple functions

Recall: If $f \in L^2(\Omega_B \otimes [0,T])$, where $f(\omega,t)$ is $\mathcal{F}_t$ measurable for each t, we want to understand the integral

$$I_T(f) = \int_0^T f(\omega,t)\,dB_t.$$

If $f(\omega,t)$ is the number of stocks we have at time t, then the integral gives the profit we get between 0 and T. We start by analyzing the function

$$f = a(\omega)\mathbb{1}_{[t_1,t_2]}(t), \qquad 0 \le t_1 \le t_2 \le T$$

This is a step function on a fixed interval with a random height. In this case,

$$I_T(f) = a(\omega) \cdot B(t_2) - B(t_1).$$

This is intuitive: this function says we buy a stocks at time t_1 and sell them at t_2; so the profit is the change in price of stocks from t_1 to t_2 times the number of shares I have.

The integral is linear, so for simple functions,

$$f = \sum_{k=1}^n a_k(\omega)\mathbb{1}_{[t_k^1,t_k^2]}(t) \quad a_k \in \mathcal{F}_{t_k^1} \implies I_T(f) = \sum_{k=1}^n a_k(\omega)(B(t_k^2) - B(t_k^1)).$$

18.2 Extending the Itô integral to general L^2 functions

If $f \in L^2$, we want to find a sequence of simple functions $f_k \xrightarrow{L^2} f$ so we can let $I_T(f) := \lim_k I_T(f_k)$.

Lemma 18.1. *If f is a simple function,*

$$\|f\|_{L^2(\Omega\times[0,T])} = \|I_T(f)\|_{L^2(\Omega)}.$$

Proof. Suppose $f = a(\omega)\mathbb{1}_{[s_1,s_2]}(t) + b(\omega)\mathbb{1}_{[t_1,t_2]}$ with $s_1 < t_1 < s_2 < t_2$. First,

$$\mathbb{E}\left[\int_0^T f^2\,dt\right] = \mathbb{E}[(t_1 - s_1)a^2 + (s_2 - t_1)(a+b)^2 + (t_2 - s_2)b^2].$$

On the other hand,

$$\mathbb{E}[I_T^2(f)] = \mathbb{E}[[(B(t_1) - B(s_1))a + (B(s_2) - B(t_1))(a+b) + (B(t_2) - B(s_2))b]^2]$$

Say the intervals are $J_1 = [s_1,t_1]$, $J_2 = [t_1,s_2]$ and $J_3 = [s_2,t_2]$. Then if we look at $\mathbb{E}[B(J_1)aB(J_2)(a+b)]$ for example, $B(J_2)$ is independent of the rest. So the crossing terms cancel.

$$\begin{aligned} &= \mathbb{E}[(B(t_1) - B(s_1))^2a^2 + (B(s_2) - B(t_1))^2(a+b)^2 + (B(t_2) - B(s_2))^2b^2] \\ &= (t_1 - s_1)\,\mathbb{E}[a^2] + (s_2 - t_1)\,\mathbb{E}[(a+b)^2] + (t_2 - s_2)\,\mathbb{E}[b^2]. \end{aligned}$$ □

So we get that $I_T : L^2(\Omega \times [0,1]) \to L^2(\Omega)$ is isometric on simple functions. So if $\|f_m - f_n\|_2 \to 0$, we get $\|I_T(f_m) - I_T(f_n)\|_2 \to 0$. So we can convert Cauchy sequences in $L^2(\Omega \times [0,1])$ to Cauchy sequences in $L^2(\Omega)$, find the limit, and use it to define $I_T(f)$.

Remark 18.1. $I_T(f)$ is "only" L^2-unique. So if $h = I_T(f)$ except in a probability 0 set, h is also $\lim I_T(f_k)$. This is the same as with the definition of conditional expectation.

18.3 The Itô integral as a random function

If $t \leq T$, let

$$F(\omega, t) = \int_0^t f(\omega, s)\, dB_s$$

F is a random function. We should believe that $F \in C([0,1])$ a.s. Here is the issue: for any t, we know $F(\omega, t)$ in a probability 1 set. But we want to have a random variable for all t at once. We can have $\mathbb{P}(\widetilde{F}(\omega, t) = F(\omega, t)) = 1$ for fixed t, but it is still possible that $\mathbb{P}(\widetilde{F} \neq F) = 1$ because $\{\widetilde{F} = F\} = \bigcap_t \{\widetilde{F}(\omega, t) = F(\omega, t)\}$; this is an uncountable intersection.

So if we want to define the random function $F(\omega, t)$, it is not a "simple" extension of $I_T(f)$. To make the construction work out, we need to make sure F is continuous. But this is hard in general; in general, if I have X_t for each t, it's not easy to find $F(t) \in C([0,1])$ with $F(t) \stackrel{d}{=} X_t$ for each t. We will need to find a sequence of continuous functions that converge uniformly to $F(\omega, t)$.

19 The Itô Integral as a Continuous Random Function

19.1 Constructing the integral as an a.s. limit of continuous function

Let $\mathcal{H} \subseteq L^2(\Omega \times [0,1])$ be the collection of **adapted** functions; that is, $f \in \mathcal{H}$ if $f(\omega, t) \in \mathcal{F}_t$ for all t. Last time we defined the Itô integral so that

$$I_T(f) = \lim_n I_T(f_n),$$

where f_n are simple functions with $f_n \xrightarrow{L^2} f$.

We wanted to construct a random function $F(\omega) \in C([0,T])$ such that $F_t(\omega) \overset{d}{=} I_t(f)$. Note that in $\int f(B_s)\, dB_s$, we had a Riemann sum:

$$\sum_k f(B(t_k))B(\Delta t_k)$$

We fixed ω first, cut the path B_t into small pieces, and took the limit of the sum. For $f \in \mathcal{H}$ we did not define $I_T(f)$ in this way. We defined $I_T(f)$ first, based on $f_n \to f$. So for each ω, we do not know what the path F_t looks like.

Theorem 19.1. *With probability 1, $t \mapsto F_t$ is continuous.*

Proof. Suppose

$$f_n = \sum_k a_k^{(n)} \mathbb{1}_{[t_k^{(n)}, t_{k+1}^{(n)}]}(t).$$

Then define

$$F_n(t) = \sum_{k \le k_0} a_k^{(n)}(\omega)[B(t_{k+1}^{(n)}) - B(t_k^{(n)})] + a_{k_0+1}^{(n)}(\omega)B(t), \qquad k_0 = \max\{k : t_{k+1} \le t\}.$$

This $F_n(t)$ is continuous a.s. We want to take the limit to get $F(t)$; that is, we want a limit point of $(F_n(t))$ in $C([0,1])$. So we want to find a Cauchy sequence in $C([0,T])$. We we want to look at

$$\max_t |F_n(t) - F_m(t)|$$

We defined F_n by $f(m)$, so if we denote $F_n = F(f_n)$, we see that $F_n - F_m = F(f_n - f_m)$. So we want to find

$$X_{m,n} := \max_t |F(f_n - f_m)(t)|$$

This difference, which we can call $F_{n,m}(t)$, is is a martingale with respect to t. Doob's inequality says

$$\mathbb{E}[X_{m,n}^2] \le C\, \mathbb{E}[F_{n,m}(T)^2]$$

So Chebyshev's inequality gives

$$\begin{aligned}\mathbb{P}(X_{m,n} \geq \varepsilon) &\leq \frac{C}{\varepsilon^2}\,\mathbb{E}[F_{n,m}(T)^2] \\ &= \frac{C}{\varepsilon^2}\|I_T(f_n) - I_T(f_m)\|_{L^2}^2 \\ &= \frac{C}{\varepsilon^2}\|f_n - f_m\|_2^2.\end{aligned}$$

Pick a subsequence such that $\|f_{k_n} - f_{k_m}\|_2^2 \leq 2^{-3n}$ for $k_m \geq k_n$. Then we get

$$\mathbb{P}(X_{n,m} \geq 2^{-n}) \leq 2^{-n}.$$

By the Borel-Cantelli lemma, with probability 1, there is an $n_0(\omega)$ such that when $n \geq n_0$,

$$\|F_{k_n} - F_{k_{n+1}}\|_\infty \leq 2^{-n}.$$

So F_{k_n} is Cauchy a.s. So f_{k_n} has a limit in $C([0,T])$. $\square$

19.2 Further considerations about this construction

1. For fixed t, does $F(t) \overset{d}{=} I_t(f)$? Yes, because $F_{k_n}(t) \overset{d}{=} I_t(f_{k_n})$ for each n.

2. $F(t)$ is also a martingale with respect to t: $F(s) = \mathbb{E}[F)t_{|}\mathcal{F}_s]$, so then $F_{k_n}(s) = \mathbb{E}[F_{k_n}(t) \mid F(s)]$. F is a mar

Based on how we defined this version of $F(t)$, it is unclear how much the path $F(\omega, t)$ depends on $B(\omega)$ for fixed ω. If we picked ω_0 so $B(\omega_0)$ is the 0 function, can we even say that $F(\omega_0)$ is the 0 function?

20 When is the Itô Integral Zero?

20.1 The Itô integral with respect to a stopping time

Let ω_0 be such that $B(\omega_0, t) = 0$ for all t. Then do we get $(\int_0^t f\, dB_s)(\omega)$ for all t? The issue is that $\int_0^T f\, dB_s$ is only L^2-unique.

When discussing the Itô integral, we constructed it as a limit of $I_T(f_n)$, where $f_n \to f$ are simple functions:

$$f_n = \sum a_k^{(n)}(\omega) \mathbb{1}_{[t_k^{(n)}, t_{k+1}^{(n)}]}(t), \qquad \int_0^t f_n\, dB_s = \sum a_k^{(n)} B(\Delta_k^{(n)}).$$

Then if $(\int_0^t f_n\, dB_s)(\omega_0) = 0$ for all t, then $(\int_0^t f\, dB_s)(\omega_0) = 0$.

Theorem 20.1. *Let $f(\omega, t)$ be $\Omega \times [0, T]$ measurable and adapted to $\mathcal{F}_t$. Let*

$$X(\omega, t) = \int_0^t f(\omega, s)\, dB_s,$$

and let ν be a stopping time such that $f(\omega, t) = 0$ if $t < \nu(\omega)$. Then $X(\omega, t) = 0$ if $t < \nu(\omega)$.

Example 20.1. Let

$$f(\omega, t) = \begin{cases} 1 & |B(\omega, t)| \geq 2 \\ 0 & \text{otherwise,} \end{cases} \qquad \nu(\omega) = \inf\{t : |B(\omega, t)| \geq 1\}$$

Then $X(\omega, t) = 0$ if $t < \nu(\omega)$.

We want to say something like

$$\int_0^t f\, dB_s = \int_0^{t \wedge \nu} f\, dB_s = \int_0^{t \wedge \nu} 0\, dB_s = 0,$$

But we do not have any definition for integrating with bounds determined by a random variable.

Lemma 20.1. *Let $f, g \in L^2$ be adapted, and let ν be a stopping time. If $f = g$ for $t \leq \nu$, then $X(t) = Y(t)$ for $t \leq \nu$. Here, $X(t) = \int_0^t f\, dB_s$, $Y(t) = \int_0^t g\, dB_s$.*

Proof. We know $f_n \to f$, where $f_n = \sum a_k^{(n)}(\omega) \mathbb{1}_{[t_k^{(n)}, t_{k+1}^{(n)}]}(t)$ are simple functions. We can choose these to be $t_k^{(n)} = k/2^n$. Define

$$\widetilde{f}_n = \sum a_k^{(n)}(\omega) \mathbb{1}_{[t_k^{(n)}, t_{k+1}^{(n)}]}(t) \mathbb{1}_{\{\nu \leq t_k^{(n)}\}}.$$

Then $\widetilde{f}_n(\omega, t) = 0$ if $t \le \nu$. These are still simple functions, so we can compute the integrals of them; we have

$$\left(\int_0^t \widetilde{f}_n \, dB_s\right)(\omega, t) = 0 \qquad \text{if } t < \nu.$$

We need to show $\widetilde{f}_n \to f$ in $L^2(\Omega \times [0, T])$. Check this.

If we have $(\int \widetilde{f}_n \, dB_s)(\omega, t) = 0$ if $t < \nu$, then for fixed t, $(\int_0^t f \, dB_s)(\omega, t) = 0$ in the event $t < \nu$ (with probability 1). But this is not enough to say that $\int_0^t f \, dB_s = 0$ for all t a.s. But it holds at all rational t with probability 1. Since $I_t(f)$ is continuous in t, we get that this is 0 for all t with probability 1. $\square$

20.2 Extending the integral to L^2_{loc}

Recall $\mathcal{H}^2 = \{f \in L^2(\Omega \times [0, T]) : f \text{ is adapted}\}$. Even some basic functions do not make it into this collection: $e^{B(s)^4} \notin L^2$. We must find some way to extend the integral to a more general class of functions. Define $L^2_{\text{loc}} = \{f : \mathbb{P}(\int_0^T |f(\omega, t)|^2 \, dt < \infty) = 1, f \text{ is adapted}\}$. This class contains $\mathcal{H}^2$: if $\mathbb{E}[X] < \infty$, then $X < \infty$ a.s. This class even encompasses functions such as $e^{e^{e^{B(s)}}}$.

Why stop at 2? We can define L^p_{loc} similarly; then $L^p_{\text{loc}} \subseteq L^q_{\text{loc}}$ if $p > q$.

Example 20.2. Let $f(s) = (1-s)^{-1/2}$; this is not random. Then what is $\int_0^1 f \, dB_s$? Define

$$Z_n = \int_{1-2^{-n}}^{1-2^{-n-1}} f \, dB.$$

Then $\mathbb{E}[Z_n] = Z_n$, and $\text{Var}(Z_n) = \|f\|_{L^2(1-2^{-n}, 1-2^{n-1})} = \log(2)$. The Z_n are independent, but they all have the same variance. So if we break the integral down into more and more pieces, there is no way we can make sense of this.

21 Itô Integration of L^2_{loc} Functions and Local Martingales

21.1 Why only L^2_{loc}?

Why should we not try to integrate functions in L^p_{loc} for $p \neq 2$?

Proposition 21.1. *Let $f(t) := \int_0^t (1-s)^{-1/2}\, dB_s$. Then $\lim_{s\to 1} f(s)$ does not exist.*

Proof. The idea is that $f(t_2) - f(t_1) \perp f(t_1) - f(t_3)$ if $(t_1, t_2) \cap (t_3, t_4) = \varnothing$. If we look at $\|f(t)\|_{L^2}$ for fixed t, this goes to ∞ as $t \to 1$.

$$\|f(t)\|_{L^2} = \int_0^t (1-s)^{-1}\, ds \xrightarrow{t\to 1} \infty.$$

Then let t_k be such that $\|f(t_k) - f(t_{k-1})\|_{L^2} \geq 2\|f(t_{k-1})\|_{L^2}$. $\square$

21.2 Defining the Itô integral for L^2_{loc} functions

If $f \in L^2_{\mathrm{loc}}$, we want to define the Itô integral

$$F_t = \int_0^t f\, dB_s.$$

We know that

$$\mathbb{P}\left(\int_0^T f^2(t,\omega)\, dt < \infty\right) = 1.$$

The idea is to define $f^{(n)}(t) = f(t \wedge \tau_n)$, where $\tau_n := \inf\{r : \int_0^r f^2\, ds \geq n\}$. Then

$$\mathbb{E}\left[\int_0^T (f^{(n)})^2\, dt\right] < \infty.$$

Now, we can let

$$F_t^{(n)} = \int_0^t f^{(n)}\, dB_s, \qquad F_t = \lim_{n\to\infty} F_t^{(n)}.$$

We get that $f(t)^{(n+1)} = f(t)^{(n)}$ for $t \leq \tau_n$. From our considerations last time, this gives $F^{(n+1)}(t) = F^{(n)}(T)$ for $t \leq \tau_n$. And since $f \in L^2_{\mathrm{loc}}$, $T \wedge \tau_n \to T$.

Moreover, F_t is a continuous function (which we get from the sequential consistency $F^{(n+1)}(t) = F^{(n)}(T)$ for $t \leq \tau_n$). In general we have the following, for any stopping times τ_n.

Proposition 21.2. *Let $f \in L^2_{\mathrm{loc}}$, and let $\tau = (\tau_n)_n$ be stopping times with $\tau_n < \tau_m$ for $m > n$ and $T \wedge \tau_n \to T$. Then $f_n(t) := f_{t\wedge\tau_n}$ are $\mathcal{H}^2$ functions. With these stopping times,*

$$F_t^{(\tau)} := \lim_n F_t^{(n)},$$

where the convergence is uniform convergence on compact sets.

We want to show that this definition is independent of τ.

Proposition 21.3. *If τ and μ are families of stopping times satisfying these properties, then $F^{(\tau)} = F^{(\nu)}$. a.s.*

Proof. Define $\mathcal{W}_n = \min\{\tau_n, \nu_n\}$. Then $F_t^{(n),\tau} = F_t^{(n),\nu}$ for all $t \leq \mathcal{W}_n$. $\square$

21.3 Local martingales

If $f \in \mathcal{H}$, then $\int f \, dB_s$ is a martingale. How about for $f \in L^2_{\text{loc}}$? It is not, but it is a local martingale.

Definition 21.1. Let X_t be a continuous random process with stopping times $\tau_n \nearrow +\infty$ and let $X_t^{(n)} = X_{t \wedge \tau_n}$. (X_t, τ) is a **local martingale** if $X_t^{(n)}$ is a martingale for each n.

Here is a natural example of a local martingale which is not a martingale.

Example 21.1. A simple random walk is a martingale, and we can linearly interpolate to get a continuous random process X_t. Look at $\tau := \inf\{n : S_n = -1\}$. Then $\tau < \infty$ a.s., but

$$\mathbb{E}[\lim_{t \to \infty} X_{t \wedge \tau}] = -1 \neq \lim_{t \to \infty} \mathbb{E}[X_{t \wedge \tau}].$$

If $f(t)$ is increasing and continuous and X_t is a martingale, then $Z_t := X(f(t))$ is still a martingale. We are going to use this change the scale of the times. Let $Y_t = B_{t \wedge \tau}$, where $\tau = \inf\{t : B_t \leq -1\}$. Now define

$$X_t = Y_{\frac{t}{1-t}}, \qquad 0 < t < 1.$$

Then X_t is a martingale on $[0, 1)$. Now extend

$$X_t^+ := \begin{cases} X_t & 0 \leq t < 1 \\ -1 & t \geq 1. \end{cases}$$

We claim that X_t^+ is a local martingale. Define $\tau_n = \min(\inf\{t : X_t \geq n\}, n)$. Then $\tau_n \nearrow \infty$. Then $X_{t \wedge \tau_n}^+$ is a bounded martingale.

22 Local Martingales for Stochastic Integrals

22.1 Stochastic integrals as local martingales

Example 22.1. Let $X_t = B_{t\wedge\tau}$, where $\tau = \inf\{t : B_t \le -1\}$. Then define

$$Y_t = \begin{cases} X_{\frac{t}{1-t}} & t \le 1 \\ -1 & t \ge 1. \end{cases}$$

Then Y_t is a local martingale with

$$\tau_n = \begin{cases} \inf\{t : Y_t \ge n\} & \text{if inf} < \infty \\ n & \text{otherwise.} \end{cases}$$

We need to check that $Y_t^{(n)} := Y_{t\wedge\tau_n}$ is a martingale. First, $Y_t^{(n)} \in \mathbb{R}$, for each t, so this is well-defined. Then $Y_t^{(n)} = B_{\widetilde{t}\wedge\widetilde{\tau}_n} = B_{\frac{t}{1-t}\wedge\widetilde{\tau}_n}$, where $\widetilde{\tau}_n = \frac{\tau_n}{1-\tau_n}$ and $\widetilde{\tau}_n = \inf\{\widetilde{t} : B_{\widetilde{t}} = n \text{ or } -1\}$.

If $f \in L^2_{\text{loc}}([0,T])$, why is $\int_0^t f\,dB_s$ a local martingale? We know that $\int_0^T f^2\,dt < \infty$ a.s., and $F_t = F_T$ if $t > T$. Let

$$\tau_n = \begin{cases} \inf\{t : \int_0^t f^2 \ge n\} & \text{if inf} < \infty \\ nT & \text{otherwise.} \end{cases}$$

22.2 Recovering martingales from local martingales

What properties do local martingales have?

Example 22.2. Let X_t be a local martingale with $X_0 = 0$, and let $\tau = \inf\{t : X_t = -a \text{ or } b\}$. We want to calculate $\mathbb{P}(X_\tau = a)$. When this is a martingale, we use the fact that $\mathbb{E}[X_\tau] = 0$.

Define $X_t^{(n)} := X_{t\wedge\tau_n}$, which is a martingale. Then $X_{t\wedge\tau}^{(n)}$ is a (bounded) martingale. Then $\mathbb{E}[X_\tau^{(n)}] = 0$. This gives us $\mathbb{E}[X_{\tau\wedge\tau_n}] = 0$. But $X_\tau = \lim_{n\to\infty} X_{\tau\wedge\tau_n}$ as $n \to \infty$. Since this is pointwise bounded convergence, we get L^1 convergence: $\mathbb{E}[X_{\tau\wedge\tau_n}] = \mathbb{E}[X_t]$.

Theorem 22.1. *Let X_t be a local martingale with a sequence of stopping times τ_n, and let τ be a stopping time. Then $Y_t := X_{t\wedge\tau}$ is a local martingale. Furthermore, if X_t is bounded ($\sup_{\omega,t} |X_t| \le M$), then X_t is a martingale.*

Proof. The martingale property is $\mathbb{E}[X_t \mid \mathcal{F}_s] = X_s$. For local martingales, we have $\mathbb{E}[X_{t\wedge\tau_n} \mid \mathcal{F}_s] = X_{s\wedge\tau_n}$. Letting $n \to \infty$ so $\tau_n \to \infty$, the right hand side becomes $X_{s\wedge\tau_n}$. However, this does not necessarily mean that the left hand side goes to $\mathbb{E}[X_t \mid \mathcal{F}_s]$; we can have $f_n \to f$ but $\int f_n \not\to \int f$. But Fatou's lemma tells us that if $X_t \ge 0$,

$$\mathbb{E}[X_t \mid \mathcal{F}_s] \le \liminf_n \mathbb{E}[X_{t\wedge\tau} \mid \mathcal{F}_s] \le \lim_n X_{s\wedge\tau_n} \le X_s.$$

So if $X_t \ge 0$, then X_t is a supermartingale. So if $|X_t| \le M$, then X_t is a martingale. $\square$

22.3 Local martingale form of Itô's formula

We have learned one form of Itô's formula:

$$f(B_t) - f(0) = \int_0^t f'(B_s)\,dB_s + \frac{1}{2}\int_0^t f''(B_s)\,ds$$

for $f \in C^2$.

People like to use the formula in another way:

$$\int_0^t f(B_s)\,dB_s = \int_0^{B_t} f(s)\,ds - \frac{1}{2}\int_0^t f'(B_s)\,ds$$

for $f \in C^1$. Here, the right hand side depends on ω, so the left hand side should, as well. So the left hand side is determined "ω by ω" or "pathwise."

Let's extend this notion to $f(t, B_t)$ for $f \in C^1 \times C^2$. We have

$$f(t, B_t) = f(0, B_0) + \int_0^t f^{(1,0)}(s, B_s)\,ds + \int_0^t f^{(0,1)}(s, B_s)\,dB_s + \frac{1}{2}\int_0^t f^{(0,2)}(s, B_s)\,dB_s,$$

where the superscripts denote partial derivatives. Here, if $f^{(1,0)} = f^{(0,2)}$, then $f(t, B_t)$ equals a constant plus the 3rd term, which is a local martingale. This will give us that f is a local martingale.

23 Constructing Martingales and Itô Integration in $\mathbb{R}^d$

23.1 Constructing martingales using Itô integration

If we have $f(t, B_t)$, then Itô's formula gives us

$$df(t, B_t) = f_t\, dt + f_x dB_s + \frac{1}{2} f_{xx}\, dt$$

If $f_t = -\frac{1}{2} f_{xx}$, we get that

$$f(t, B_t) = \int f_x\, dB_s.$$

This implies that $f(t, B_t)$ is a local martingale. If this is bounded, it must be a martingale.

Example 23.1. Suppose we have the line $\mu t + a$. What is $\mathbb{P}(\exists t \text{ s.t. } B_t = \mu t + a)$? Let's try looking at $X_t = B_t - \mu t$. Then we want $\mathbb{P}(\exists t \text{ s.t. } X_t = a)$.

In general, let's look at $X_t = \mu t + \sigma B_t$, where $\mu, \sigma \in \mathbb{R}$. If we let $\tau = \inf\{t : X_t = a \text{ or } b\}$, then the probability we want is the limit $\lim_{b \to \infty} \mathbb{P}(X_\tau = a)$. If $\mu = 0$, then X_t is a martingale. Then $Y_t := \frac{X_t - b}{a - b}$ is also a martingale, and

$$Y_\tau = \begin{cases} 1 & X_\tau = a \\ 0 & X_\tau = b. \end{cases}$$

Then

$$\mathbb{P}(Y_\tau = 1) = \mathbb{E}[Y_\tau] = \mathbb{E}[Y_0] = -\frac{b}{a - b}.$$

The middle step comes from the fact that $Y_{t \wedge \tau}$ is bounded.

Example 23.2. Let's find some function with the form $f(t, x) = \widetilde{f}(\mu t + \sigma z)$ and $f_t + \frac{1}{2} f_{xx} = 0$; such a function will make $\widetilde{f}(X_t)$ a local martingale. The derivative condition is $\mu \widetilde{f}' + \frac{1}{2}\sigma^2 \widetilde{f}'' = 0$. Notice that if $\mu = 0$, then $\widetilde{f}(s) = cs + d$ for some c, d; so $cX_t + d$ is a martingale.

Let's try the function

$$\widetilde{f}_{A,B}(s) = \frac{e^{-2\mu s/\sigma^2} - e^{-2\mu B/\sigma^2}}{e^{-2\mu A/\sigma} - e^{-2\mu B/\sigma^2}}.$$

Then $\widetilde{f}$ satisfies this differential equation, $\widetilde{f} = 1$ if $s = A$, and $\widetilde{f} = 0$ if $s = B$. So

$$\mathbb{E}[\widetilde{f}_{A,B}(X_t)] = \mathbb{P}(X_\tau = A) = \mathbb{E}[\widetilde{f}_{A,B}(X_0)].$$

23.2 Brownian motion and Itô integration in $\mathbb{R}^d$

We can construct Brownian motion in $\mathbb{R}^d$ by constructing a length d vector of independent Brownian motions $B(t) = (B_1(t), B_2(t), \ldots, B_d(t))$. Here is what Itô's formula looks like in $\mathbb{R}^d$:

Theorem 23.1. *For suitable $f(t, B(t))$.*

$$\begin{aligned} df(t, B(t)) &= f_t\,dt + \sum_{k=1}^{d} f_{x_k}\,dB_k + \frac{1}{2}\sum_{k=1}^{d} f_{x_k x_k}\,dt \\ &= f_t\,dt + \nabla f \cdot dB + \frac{1}{2}\Delta f\,dt. \end{aligned}$$

Equivalently,

$$f(t, B(t)) = \int_0^t f_t(s, B(s))\,ds + \sum_{k=1}^{d} f_{x_k}(s, B(s))\,dB_k(s) + \frac{1}{2}\int_0^t f_{x_k x_k}(s, B(s))\,dt.$$

Corollary 23.1. *If $f(t,x)$ satisfies $f_t + \frac{1}{2}\Delta f = 0$, then $f(t, B_t)$ is a local martingale.*

Look at 2-dimensional brownian motion. Then for $a \in \mathbb{R}^2$,

$$\mathbb{P}(\exists t \text{ s.t. } B_t = a) = 0.$$

We can also find the following.

Proposition 23.1.

$$\mathbb{P}(\exists \{t_k\} \textit{ s.t. } \lim_k B_{t_k} = a) = 1.$$

This is a nontrivial question to ask. However, this is a well-known result nowadays. Interestingly, this probability is not equal to 1 when $d > 2$.

Proof. We want to find $\mathbb{P}_a(\exists t \text{ s.t. } |B_t| \le r)$; if this probability is 1 for $r > 0$, then the result holds by taking a sequence of r going to 0. Let $\tau := \inf\{t : |B_t| = r \text{ or } r_L\}$, where $r_L > R$. Then we want $\mathbb{P}(|B_\tau| = r)$. Let's find some $f(t, B_t)$ such that $f_t + \frac{1}{2}f_{xx} = 0$. We also need a kind of property like $f(t,x) = 1$ if $|B_t| = r$ and $f(t, B_t) = 0$ if $|B_t| = r_L$. We can do this by requiring $f(t,x) = h(\|x\|)$, i.e. it only depends on h. This gives $\Delta f = 0$. So we want a function like $h(x)\log|x|$. The actual choice is

$$f(x) = \frac{\log|x| - \log(r_L)}{\log(r) - \log(r_L)}.$$

This gives

$$\mathbb{P}(|B_\tau| = r) = \mathbb{E}[f(B(t))] = \mathbb{E}[f(B(0))] = \frac{\log(R) - \log(r_L)}{\log(r) - \log(r_L)} \xrightarrow{r_L \to \infty} 1.$$

This completes the proof. ∎

Remark 23.1. If $d = 3$, we get that $f(x)$ is like $1/|x|$. In particular,

$$f(x) = \frac{1/|x| - 1/|r_L|}{1/|r| - 1/|r_L|}.$$

The same calculation gives

$$\mathbb{P}(|B_\tau| = r) = \frac{1/R - 1/r_L}{1/r - 1/r_L} \xrightarrow{r_L \to \infty} \frac{r}{R}.$$

24 Integrating With Respect to Random Processes

24.1 Integrating with respect to random processes

The 1-dimensional version of Itô's formula says that

$$df(t, B_t) = f_t\,dt + f_x\,dB_t + \frac{1}{2} f_{xx}\,dt.$$

For a Brownian motion $B_t = (B_1, \ldots, B_d)$ in $\mathbb{R}^d$, we have

$$df(t, B_t) = f_t\,dt + \nabla f \cdot dB + \frac{1}{2}\Delta f\,dt.$$

Let's cover something more general.

Suppose we have a process

$$X_t = X_0 + \int_0^t a(\omega, s)\,ds + \int_0^t b(\omega, s)\,dB_s.$$

Here, $a(t), b(t) \in \mathcal{F}_t$ for any t, where $\mathcal{F}_t$ is the filtration with respect to Brownian motion. We can write this as $dX_t = a(t)\,dt + b(t)\,dB_t$. How can we define $\int_0^t f(s)\,dX_s$? And in $\mathbb{R}^d$, what if we have $dX_t = a(t)\,dt + \widehat{b}(t)\,dB(t)$, where X_t, a are vectors and $\widehat{b}$ is a matrix?

To figure out the value of $\int_0^t f(s)\,dX_s$, we can write

$$df(t, X_t) = f_t\,dt + f_x\,dX_t + \frac{1}{2} f_{xx} b^2(t)\,dt,$$

where

$$f_x\,dX_t = (f_x a(t))\,dt + b(t)\,dB_t.$$

If we replace X_t with B_t, we get $a(t) = 0$ and $b(t) = 1$; but this does not really help us understand this generalization.

Here is why the formula looks like this: we have $f_t \Delta t + f_x \Delta X_t + \frac{1}{2} f_{xx} (\Delta t)^2$. How do we understand the last term? Look at $\mathbb{E}\left[\frac{(\Delta X_t)^2}{\Delta t}\right]$. As $\Delta t \to 0$, we have $\Delta X_t \to a(t)\Delta t + b(t)\Delta B_t$. So

$$\frac{(\Delta X_t)^2}{\Delta t} \to b^2(t) \cdot \frac{(\Delta B_t)^2}{\Delta t}.$$

So the last term in our formula is actually like $\frac{1}{2} f_{xx} (dX_t)^2$.

Example 24.1. Let $X_t = \int \cos(t)\,dB_t$. What is $\sin(X_t)$? This is a question from an interview book.

Returning to the vector version $dX_t = a(t)\,dt + \widehat{b}(t) dB_t$, we have

$$df(t, X_t) = f_t\,dt + \nabla \cdot dx_t + \frac{1}{2}\,dX_t \cdot \mathrm{Hess}(f)\,dX_t.$$

24.2 Quadratic variation

If we know X_t, can we find $\widehat{b}(t)$? The idea is that X_t is what you actually observe, and we are modeling it as Brownian motion. Then we want to recover some information about the model. If X_t is our process, then we want to find the **quadratic variation** $\langle X \rangle_t$:

$$\lim_{\Delta_t \downarrow 0} \sum_k (X_{t_k} - X_{t_{k-1}})^2.$$

We find that

$$\langle X \rangle_t = \int_0^t b^2(s)\, ds.$$

Then

$$\frac{d\langle X \rangle_t}{dt} = b^2(t).$$

In applications, we can assume that b is a.s. non-random. This gives us a lot of information about how the sample paths of our process vary. This is used in financial mathematics to "observe" the variation of stock prices.

24.3 Dyson spheres

We will discuss Dyson Brownian motion next time. Dyson is famous in science fiction circles for a different idea. He thought that in any developed culture, there will be a race towards more efficient forms of energy production. Solar energy if one of the most efficient and long-lasting forms of energy, so he thought that people would build solar panels close to their sun, where they can get the most energy. So eventually, the sun would be covered by solar panels, in what is called a **Dyson sphere**. Dyson proposed that to look for alien life, we should look for stars where the brightness has been reduced (to indicate that it has been covered by solar panels.

25 Dyson Brownian Motion

25.1 The drifting term in the rate of change of the eigenvalues

Let's say we have Hermitian matrices $H^t = H^0 + H^t_G$, where $H^t = (H^t)^\dagger$ and $H^0 = (H^0)^\dagger$. Let λ^r_k be the k-th largest eigenvalue of H^t and u^t_k be such that $\|U^t_k\|_2 = 1$ and $H^t u^t_k = \lambda^t_k u^t_k$ (eigenvector).

We let $[H^t_G]_{i,j} \sim B(t)$ for $i < j$ and $[H^t_G]_{i,i} \sim B(2t)$. The entries of H^t are independent. with $H^t_G = (H^t_G)^\dagger$. Then $H^t_{i,j} = H^0_{i,j} + B_{i,j}(t)$ is a process for each $i \le j$. So we can consider eigenvalues as a function of these processes: $\lambda^t_k(\{H_{i,j}\}_{i\le j})$. What if we change the eigenvalues a little bit?

$$\begin{aligned} d\lambda^t_k &= \sum_{i,j} \frac{\partial \lambda_k}{\partial H^t_{i,j}}\, dB^t_{i,j} + \frac{1}{2} \sum_{i,j,i',j'} \frac{\partial^2 \lambda_k}{\partial H_{i,j} \partial H_{i',j'}}\, dB_{i,j}\, dB_{i',j'} \\ &= \nabla \lambda_k \cdot dH^t + dH^t \cdot \mathrm{Hess}(\lambda_k)\, dH^t. \end{aligned}$$

Since these $N \times N$ matrices are Hermitian, we have N^2 processes which are not all independent.

Since $dB_{i,j} dB_{i',j'} = dt$ if $\{i,j\} = \{i',j'\}$ and 0 otherwise, we get

$$d\lambda^t_k = \sum_{i,j} \frac{\partial \lambda_k}{\partial H^t_{i,j}}\, dB^t_{i,j} + \frac{1}{2} \sum_{i,j} \left[\frac{\partial \lambda_k}{\partial H^2_{i,j}} + \frac{\partial^2 \lambda_k}{\partial H_{i,j} \partial H_{j,i}} \right] dt$$

There are results that say

$$\frac{\partial \lambda_k}{\partial_{H_{i,j}}} = u_k(i) u_k(j), \qquad \frac{\partial u_k}{\partial H_{i,j}} = \sum_{\ell \ne k} \frac{u_k(i) u_\ell(j)}{\lambda_\ell - \lambda_k} u_\ell.$$

So the right term, the **"drifting term"**, is

$$\sum_\ell \frac{-1}{\lambda_\ell - \lambda_k}\, dt.$$

So if λ_k is very close to λ_{k+1}, we get a very large negative term. So there is some large force to pull λ_k down, away from λ_{k+1}. This gives us a surprising property: the order of the eigenvalues never changes! That is, if we plot all the eigenvalues with t, the curves never intersect.

25.2 The diffusion term

Let's do some calculations with the first term, the **"diffusion term"**. This is

$$\sum_{i,j} u_k(i) u_k(j)\, dB_{i,j}(t)$$

Now define

$$B_k(t) = \int \sum_{i,j} u_k^t(i) u_k(j)\, dB_{i,j}(t).$$

This is a process, and we can check that for any fixed t, $B_k(t) \sim \mathcal{N}(0,t)$: We can calculate

$$\mathbb{E}[B_k^2(t)] = \sum_{i,j} u_k^t(i) u_k^t(j) u_k^t(i) u_k^t(j) \cdot t = t.$$

We can also calculate for $\ell \neq k$,

$$\mathbb{E}[B_k(t) B_\ell(t)] = \sum_{i,j} i_k^t(i) u_k^t(j) u_\ell(i) u_\ell(j) = 0$$

because the u_k, u_ℓ are orthogonal to each other. In fact, we can prove that

$$\mathbb{E}[B_k^m(t)] = \mathbb{E}[\mathcal{N}(0,t)^m].$$

Moreover, $B_k(\cdot)$ is a Brownian motion, and $\{B_k\}_{k=1}^N$ are independent.

So we have that

$$d\lambda_k = dB_k + \sum_\ell \frac{-1}{\lambda_\ell - \lambda_k}.$$

What if I define some other process

$$d\widetilde{\lambda}_k = d\widetilde{B}_k + \sum_\ell \frac{-1}{\widetilde{\lambda}_\ell - \widetilde{\lambda}_k}$$

with the same distribution? This doesn't come from a matrix, but we can still use it to study the distribution of λ_k.

26 Stochastic Differential Equations

26.1 Examples of SDEs

If we have a process satisfying a stochastic differential equation, we may want to recover a concrete description of it. We want to solve **stochastic differential equations** of the form

$$dX_t = f(t, X_t)\, dt + g(t, X_t)\, dB_t.$$

Example 26.1. Consider the equation

$$dX_t = \mu X_t\, dt + \sigma X_t\, dB_t,$$

where μ, δ are fixed.

To solve this, we hope that $X_t = f(t, B_t)$. If this is correct, we get that

$$f_t = \frac{1}{2} f_{xx} = \mu f, \qquad f_x = \sigma f.$$

We can then solve this equation to get a solution of the form

$$f = e^{\sigma x + g(t)}.$$

Plugging this back into the equation, we get

$$f = e^{\sigma x + (\mu - \sigma/2)t}.$$

So the solution to the SDE is

$$X_t = e^{\sigma B_t + (\mu - \sigma/2)t} \cdot X_0.$$

These do not always have this kind of solution. Here is an example:

Example 26.2. Consider the **Ornstein-Uhlenbeck process** satisfying the equation

$$dX_t = -aX_t\, dt + \sigma\, dB_t.$$

If we try to solve it the same way, we get

$$f_t + \frac{1}{2} f_{xx} = -af, \qquad f_x = \sigma.$$

Then we get

$$f = \sigma x + g(t), \qquad g' = -a(\sigma x + g(t)).$$

So there are no solutions of this type to this differential equation.

To solve this, recall the trick to solving the ODE $f' = -af + g(x)$. We introduce a factor like $e^{\int a}$. The idea is to introduce a factor

$$Y = g(t) \cdot X(t).$$

Then we get

$$\begin{aligned} dY &= g' X(t)\,dt + g(t) \cdot dX_t \\ &= g' X_t\,dt + g(t)(-aX_t)\,dt + g(t)\sigma\,dt. \end{aligned}$$

We want something of the form $g' = ag$, so $g = ce^{at}$. With this g, we get

$$dY = \sigma\,dB_t \implies Y = Y_0 + \sigma B_t$$

We then get

$$X = ce^{-at}(Y_0 + \sigma B_t).$$

Remark 26.1. This process has $\mathbb{E}[X_t^2] < \infty$, so it has a limit.

Example 26.3. Let's say we have a **Brownian bridge**, a Brownian motion with $B_1 = 0$. This event has probability 0, so we can't condition on $B_1 = 0$. We can set up the equation

$$X_t = -\frac{X_t}{1-t}\,dt + \,dB_t.$$

We can solve this as before with $a(t) = -\frac{1}{1-t}$. We have $g' = a(t)g$, so $e^{\int_0^t a(s)\,ds} \cdot C$. And for $Y = g \cdot X$, we get

$$Y = g\,dB_t = \frac{1}{1-t}\,dB_t.$$

So

$$X = (1-t)\int_0^t \frac{1}{1-s}\,dB_s.$$

gives us the formula for a Brownian bridge.

26.2 Existence and uniqueness of solutions to SDEs

Theorem 26.1. *Consider the equation*

$$dX_t = \mu(t, X_t)\,dt + \sigma(t, X_t)\,dB_t.$$

Suppose μ, σ are smooth, Lipschitz:

$$|\mu(t,x) - \mu(t,y)| \le K|x-y|, \qquad |\sigma(t,x) - \sigma(t,y)| \le L|x-y|,$$

and $\mu\sigma \le (1+|x|)C$. Then there is a unique solution to this equation.

Proof. To prove uniqueness, suppose

$$X_t = \int_0^t \mu(s, X_s)\, ds + \int_0^t \sigma(s, X_s)\, dB_s + X_0,$$

$$Y_t = \int_0^t \mu(s, Y_s)\, ds + \int_0^t \sigma(s, Y_s)\, dB_s + Y_0.$$

By the Lipschitz condition, we get that $X_t - Y_t$ is bounded by the integral of something like $X_s - Y_s$. That is, the L^∞ norm is bounded by the L^1 norm. Such a function must be 0. In fact, we actually bound $\mathbb{E}[(X_t - Y_t)^2]$ like this.

For existence, take a sequence of iterates

$$X_t^{(n+1)} = \int_0^t \mu(s, X_s^{(n)})\, ds + \int_0^t \sigma(s, X_s^{(n)})\, dB_s + C_0,$$

and show that this converges. We do this by using the contraction mapping theorem: we show that

$$\|X^{(n+1)} - X^{(n)} \leq C\|X^{(n)} - X^{(n-1)}\|$$

for some $C < 1$. To get $C < 1$, we make the time interval small enough. □

27 The Martingale Central Limit Theorem

27.1 Motivation

The central limit theorem says something like

$$\frac{S_n}{\sqrt{n}} \xrightarrow{d} \mathcal{N}(0,1).$$

We have also learned about Donsker's theorem, which extends this idea to Brownian motion.

S_n is not necessarily small. If X_i are iid, then $\mathbb{E}[S_n] = n\,\mathbb{E}[X_i]$. The central limit theorem tells us that the fluctuation of S_n is much less than $n\,\mathbb{E}[X]$. We know that $\mathbb{E}[S_n] \sim n$ and $\mathrm{Var}(S_n) \sim n$. This is because

$$\mathbb{E}\left[\left(\sum_{i=1}^n X_i - \mathbb{E}[X_i]\right)^2\right] = \sum_{i,j} \mathbb{E}[(X_i - \mathbb{E}[X_i])(X_j - \mathbb{E}[X_j]])],$$

where these terms are 0 if $i \neq j$ by independence. So the property comes from the increments $X_k = S_{k+1} - S_k$.

But the situation is more complicated for martingales. Let $S_1, S_2, \dots$ be a martingale. Then if

$$\frac{1}{n}\sum_{k=1}^n \mathbb{E}[(S_{k+1} - S_k)^2 \mid \mathcal{F}_k] \to c,$$

then

$$\frac{S_n}{\sqrt{n}} \sim \mathcal{N}(0,1).$$

In the iid case, $c = 1$. If we define $X_k = S_{k+1} - S_k$, then X_k and X_ℓ are not independent. So this result is nontrivial (and maybe even unintuitive). The idea is that under certain conditions, the X_k are independent.

27.2 The Markov chain CLT and martingale CLT

Theorem 27.1 (Martingale CLT). *Let $\{S_k\}_k$ be a martingale, an dlet $X_k = S_{k+1} - S_k$. Suppose that*

$$\frac{\sum_{k=1}^{\lceil n \cdot t \rceil} \mathbb{E}[X_k^2 \mid \mathcal{F}_k]}{tn} \to 1 \qquad \forall t.$$

Then

$$\frac{S_{(n\cdot)}}{\sqrt{n}} \to B(\cdot).$$

Compare this to the Markov chain central limit theorem. Let $X_1, X_2, \ldots$ be a Markov chain with a stationary distribution π. If we start the chain at π, then $X_1, X_2, \ldots \overset{d}{=} X_2, X_3, \ldots$; i.e. the sequence is **stationary**. The ergodic theorem says that $\frac{1}{n}\sum_{k=1}^n f(X_k) \to \mathbb{E}_\pi[f(X_1)]$; i.e. the sequence is **ergodic**.

Theorem 27.2 (Markov chain CLT)**.** *Let $\{X_n\}_n$ be an ergodic, stationary sequence with $\mathbb{E}[X_n \mid \mathcal{F}_{n-1}] = 0$ and $\mathbb{E}[X_i^2] = 1$. Then*

$$\frac{S_n}{\sqrt{n}} \xrightarrow{d} \mathcal{N}(0,1).$$

The central limit theorem for Markov chains is a special case of the theorem for martingales. Let's prove this assuming the martingale CLT.

Proof. We want to show that

$$\frac{1}{n}\sum_k \mathbb{E}[X_k^2 \mid \mathcal{F}_{k-1}] \to 1.$$

Define $u_k := \mathbb{E}[X_k^2 \mid \mathcal{F}_{k-1}]$; this is also an ergodic, stationary sequence. So the property in the martingale CLT is satisfied. $\square$

Now let's prove the martingale CLT.

Proof. If $\{S_n\}_n$ is a martingale with $S_0 = 0$ (and $\mathbb{E}[S_n^2] < \infty$), then we can define stopping times $T_1, \ldots, T_n$ such that $(S_1, \ldots, S_n) \overset{d}{=} (B_{T_1}, \ldots, B_{T_n})$. This is a repeated use of Skorokhod's representation theorem. We then find that

$$\mathbb{E}[X_k^2 \mid \mathcal{F}_{k-1}] = \mathbb{E}[T_k - T_{k-1} \mid \mathcal{F}^B_{T_{k-1}}].$$

If $T_n \approx n$, we are done. We have $T_n = \sum_k T_k - T_{k-1}$ and $\sum_k \mathbb{E}[T_k - T_{k-1} \mid \mathcal{F}_{T_{k-1}}] = n$. If we can show that both are close, we will be done.

Let $\tau_k = T_k - T_{k-1}$, and let $V_k = \mathbb{E}[T_k - T_{k-1} \mid \mathcal{F}_{T_{k-1}}]$. We want to show that $\mathbb{E}[(\sum_k \tau_k - V_k)^2] = O(n)$. If $k < \ell$,

$$\begin{aligned}\mathbb{E}[(\tau_k - \mathbb{E}[\tau_k \mid \mathcal{F}_{T_{k-1}})(\tau_\ell - \mathbb{E}[\tau_\ell \mid \mathcal{F}_{T_{\ell-1}}])] &= \mathbb{E}[\tau_k\tau_\ell] - \mathbb{E}[\tau_\ell\,\mathbb{E}[\tau_k \mid \mathcal{F}_{T_{k-1}}]] - \mathbb{E}[\tau_l\,\mathbb{E}[\tau_\ell \mid \mathcal{F}_{T_{\ell-1}}] \\ &\quad + \mathbb{E}[\mathbb{E}[\tau_\ell \mid \mathcal{F}_{T_{\ell-1}}]\,\mathbb{E}[\tau_k \mid \mathcal{F}_{T_{k-1}}]]\end{aligned}$$

The third term becomes $-\mathbb{E}[\tau_k\tau_\ell]$. We can calculate the other terms similarly.

$$= 0.$$

$\square$