MA4F7 Brownian Motion

March 16, 2013

Contents

These notes are based on the 2013 `MA4F7 Brownian Motion` course, taught by Roger Tribe, typeset by Matthew Egginton.

No guarantee is given that they are accurate or applicable, but hopefully they will assist your study.

Please report any errors, factual or typographical, to m.egginton@warwick.ac.uk

The key aim is to show that scaled random walks converge to a limit called Brownian motion. In 1D, $\mathbb{P}\{t \mapsto B_t \text{ nowhere differentiable }\} = 1$

$\mathbb{E}(B_t) = 0, \mathbb{E}(B_t^2) = t$ and so $t \mapsto B_t$ is not differentiable at 0. By shifting gives it at any t.

We also have that $P(\int_0^1 \chi(B_s > 0)ds \in dx) = \frac{1}{\sqrt{\pi x(1-x)}}dx$

$P(x + B \text{ exits } D \text{ in } A) = U(x)$ where $\Delta U(x) = 0$ and $U(x) = \begin{cases} 1 & \text{on } A \\ 0 & \text{on} \partial D \setminus A \end{cases}$. For a disc with inner radius a and outer radius b, $U(x) = \frac{\log b - \log |x|}{\log b - \log a}$ and this converges to 1 as $b \to \infty$. Thus the probability that Brownian motion hits any ball is 1.

For random walks, $P(x+ \text{ r.v. exits at y}) = U(x)$ where $U(x) = \frac{1}{4}(U(x+e_1) + U(x - e_1) + U(x+e_2) = U(x-e_2))$ which can be thought of as a discrete Laplacian. Thus we have a nice equation for Brownian motion, but a not so nice one for random walks.

1 Brownian Sample Paths

Our standard space is a probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Definition 1.1 *A stochastic process* $(B_t, t \geq 0)$ *is called a* **Brownian Motion** *on* $\mathbb{R}$ *if*

1. $t \mapsto B_t$ *is continuous for a.s.* ω
2. *For* $0 \leq t_1 < t_2 < ... < t_n$ *we have* $B_{t_2} - B_{t_1}, ..., B_{t_n} - B_{t_{n-1}}$ *are independent*
3. *For* $0 \leq s < t$ *we have* $B_t - B_s$ *is Gaussian with distribution* $N(0, t-s)$.

But does this even exist, and if it does, do the above properties characterise B. The answer to both is yes, and we will show these later.

We now define the terms used in the above definition, to avoid any confusion.

Definition 1.2 *A* **random variable** Z *is a measurable function* $Z : \Omega \to \mathbb{R}$. *In full,* $\mathbb{R}$ *has the Borel* σ*-algebra* $\mathcal{B}(\mathbb{R})$ *and measurable means if* $A \in \mathcal{B}(\mathbb{R})$ *then* $Z^{-1}(A) \in \mathcal{F}$.

Definition 1.3 *A stochastic process is a family of random variables* $(X_t, t \geq 0)$ *all defined on* Ω.

We do not worry what Ω is, we are only interested in the law/distribution of Z, i.e. $\mathbb{P}(Z \in A)$ or $\mathbb{E}(f(Z))$ where $\mathbb{P}(Z \in A) = \mathbb{P}\{\omega : Z(\omega) \in A\}$

If we fix ω, the function $t \mapsto B_t(\omega)$ is called the **sample path** for ω.

The first property above means that the evaluation of B_t at ω is continuous, for almost all ω. Sadly some books say that $\mathbb{P}\{\omega : t \mapsto B_t(\omega) \text{ is continuous }\} = 1$ but how do we know this set is measurable.

Definition 1.4 Z *a real variable is Gaussian* $N(\mu, \sigma^2)$ *if it has density*

$$\mathbb{P}(Z \in dz) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dz$$

for $\sigma^2 > 0$*, meaning integrate both sides over a set* A *to get the probability over* A. *If* $\sigma = 0$ *then* $\mathbb{P}(Z = \mu) = 1$.

1.0.1 Related Animals

The Brownian Bridge is $X_t = B_t - tB_1$ for $t \in [0, 1]$

An Ornstein-Uhlenbeck process is one, for $C > 0$, of the form $X_t = e^{-Ct} B_{e^{2Ct}}$ and is defined for $t \in \mathbb{R}$. We will check that X here is stationary. Also $(X_{t+T} : t \geq 0)$ is still an O-U process. This arises as the solution of the simplest SDE $\frac{dX}{dt} = -CX_t + \sqrt{2C}\frac{d\hat{B}}{dt}$, or in other form, $X_t = X_0 - C\int_0^t X_s ds + \int_0^t \sqrt{2C} dB_t$.

A Brownian motion on $\mathbb{R}^d$ is a process $(B_t : t \geq 0)$ such that $B = (B_t^1, ..., B_t^d)$ where each $t \mapsto B_t^k$ is a Brownian motion on $\mathbb{R}$ and they are independent.

1.1 Brownian Motion as a Gaussian Process

Proposition 1.5 (Facts about Gaussians) *1. $Z \overset{D}{\sim} N(\mu, \sigma^2)$ then for $c \geq 0$ we have $cZ \overset{D}{\sim} N(c\mu, c^2\sigma^2)$*

2. $Z_1 \overset{D}{\sim} N(\mu_1, \sigma_1^2)$, $Z_2 \overset{D}{\sim} N(\mu_2, \sigma_2^2)$ and are independent then $Z_1 + Z_2 \overset{D}{\sim} N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$

3. $Z_k \overset{D}{\sim} N(\mu_k, \sigma_k^2)$ and if $Z_k \to Z$ then $\lim_{k\to\infty} \mu_k = \mu, \lim_{k\to\infty} \sigma_k^2 = \sigma^2$ and $Z \overset{D}{\sim} N(\mu, \sigma^2)$.

The convergence above can be any one of the following.

1. Almost surely convergence: $Z_k \overset{a.s.}{\to} Z$ means $P(\{\omega : Z_k(\omega) \to Z(\omega)\}) = 1$
2. In probability: $Z_k \overset{prob}{\to} Z$ means $P(|Z_k - Z| > \varepsilon) \underset{k\to\infty}{\to} 0$ for all $\varepsilon > 0$
3. In distribution: $Z_k \overset{D}{\to} Z$ means $\mathbb{E}(f(Z_k)) \to \mathbb{E}(f(Z))$ for any continuous and bounded f.

Example 1.1 *$I = \int_0^1 B_t dt$ is a Gaussian variable.*

$$\begin{aligned} I &= \lim_{k\to\infty} \frac{1}{N}(B_{1/N} + B_{2/N} + ... + B_{N/N}) \\ &= \lim_{k\to\infty} \frac{1}{N}((B_{N/N} - B_{(N-1)/N}) + 2(B_{(N-1)/N} - B_{(N-2)/N}) + ... + N(B_1 - B_0)) \end{aligned}$$

and all these are independent and so Gaussian.

1.1.1 Transforms

Definition 1.6 *We define the* **Fourier transform***, or the* **characteristic function** *to be*

$$\phi_Z(\theta) = \mathbb{E}(e^{i\theta Z})$$

For example, if $Z \overset{D}{\sim} N(\mu, \sigma^2)$ then $\phi_Z(\theta) = e^{i\theta\mu}e^{-\sigma^2/2}$

Proposition 1.7 (More facts about Gaussians) *4. $\phi_Z(\theta)$ determines the law of Z, i.e. if $\phi_Z(\theta) = \phi_Y(\theta)$ then $P(Z \in A) = P(Y \in A)$.*

5. Z_1, Z_2 independent if and only if $\mathbb{E}((e^{i\theta_1 Z_1} e^{i\theta_2 Z_2}) = \mathbb{E}(e^{i\theta_1 Z_1})\mathbb{E}(e^{i\theta_2 Z_2})$ for all θ_1, θ_2.

6. $\phi_{Z_k}(\theta) \to \phi_Z(\theta)$ *if and only if* $Z_k \stackrel{D}{\to} Z$.

These all hold true for $Z = (Z_1, ..., Z_d)$ with $\phi_Z(\theta_1, ..., \theta_d) = \mathbb{E}(e^{i\theta_1 Z_1 + ... + i\theta_d Z_d})$

Definition 1.8 $Z = (Z_1, ..., Z_d) \in \mathbb{R}^d$ *is* **Gaussian** *if* $\sum_{i=1}^d \lambda_k Z_k$ *is Gaussian in* $\mathbb{R}$ *for all* $\lambda_1, ..., \lambda_d$. $(X_t, t \geq 0)$ *is a* **Gaussian process** *if* $(X_{t_1}, ..., X_{t_N})$ *is a Gaussian vector on* $\mathbb{R}^N$ *for any* $t_1, ..., t_N$ *and* $N \geq 1$.

Check that Brownian motion is a Gaussian process, i.e. is $(B_{t_1}, ..., B_{t_N})$ a Gaussian vector, or is $\sum \lambda_k B_{t_k}$ Gaussian on $\mathbb{R}$. We can massage this into $\mu_1(B_{t_1} - B_0) + ... + \mu_N(B_{t_N} - B_{t_{N-1}})$ and so is Gaussian. As an exercise, check this for Brownian bridges and O-U processes.

Proposition 1.9 (Even more facts about Gaussians) *7. The Law of the Gaussian* $Z = (Z_1, ..., Z_d)$ *is determined by* $\mathbb{E}(Z_k)$ *and* $\mathbb{E}(Z_j, Z_k)$ *for* $j, k = 1, ..., d$

8. Suppose $Z = (Z_1, ..., Z_d)$ *is Gaussian. then* $Z_1, ..., Z_d$ *are independent if and only if*

$$\mathbb{E}(Z_j Z_k) = \mathbb{E}(Z_j)\mathbb{E}(Z_k)$$

for all $j \neq k$.

For 7, it is enough to calculate $\phi_Z(\theta)$ and see that it is determined by them. For 8, noe need only check that the transforms factor.

Example 1.2 (B_t) *a Brownian motion on* $\mathbb{R}$. *Then* $\mathbb{E}(B_t) = 0$ *and, for* $0 \leq s < t$,

$$\mathbb{E}(B_s B_t) = \mathbb{E}((B_t - B_s)(B_s - B_0) + (B_s - B_0)^2) = \mathbb{E}(B_t - B_s)\mathbb{E}(B_s - B_0) + \mathbb{E}(B_s - B_0)^2 = s$$

and similarly equals t *if* $0 \leq t < s$.

Do the same for Brownian bridges and O-U processes.

Theorem 1.10 (Gaussian characterisation of Brownian motion) *If* $(X_t, t \geq 0)$ *is a Gaussian process with continuous paths and* $\mathbb{E}(X_t) = 0$ *and* $\mathbb{E}(X_s X_t) = s \wedge t$ *then* (X_t) *is a Brownian motion on* $\mathbb{R}$.

Proof We simply check properties 1,2,3 in the definition of Brownian motion. 1 is immediate. For 2, we need only check that $\mathbb{E}((X_{t_{j+1}} - X_{t_j})(X_{t_{k+1}} - X_{t_k}))$ splits. Suppose $t_j \leq t_{j+1} \leq t_k \leq t_{k+1}$ and then $\mathbb{E}((X_{t_{j+1}} - X_{t_j})(X_{t_{k+1}} - X_{t_k})) = t_{j+1} - t_{j+1} - t_j + t_j = 0$ as required. For 3, $X_t - X_s$ is a linear combination of Gaussians and so is Gaussian. It has mean zero and

$$\mathbb{E}(X_t - X_s)^2 = \mathbb{E}(X_s^2 - 2X_s X_t + X_t^2) = s - 2s \wedge t + t = t - s$$

Q.E.D.

Suppose $I = \int_0^1 B_s ds$ and $\mathbb{E}(I) = \int_0^1 \mathbb{E}(B_s) ds = 0$ and also

$$\mathbb{E}(I^2) = \mathbb{E}(\int_0^1 B_s ds \int_0^1 B_r dr) = \int_0^1 \int_0^1 \mathbb{E}(B_s B_r) ds dr = \int_0^1 \int_0^1 s \wedge r ds dr = \frac{1}{3}$$

but we need to check that we can use Fubini, so we need to check that $K = \mathbb{E}(\int_0^1 \int_0^1 |B_r||B_s| dr ds) < \infty$. Now

$$K = \int_0^1 \int_0^1 \mathbb{E}(|B_r||B_s|) dr ds \leq \int_0^1 \int_0^1 \mathbb{E}(B_r^2)\mathbb{E}(B_s^2) dr ds \leq \sqrt{rs} < 1$$

as we wanted.

Lemma 1.11 (Scaling Lemma) *Suppose that B is a Brownian motion on $\mathbb{R}$ and $c > 0$. Define $X_t = \frac{1}{c}B_{c^2t}$ for $t \geq 0$. Then X is a Brownian motion on $\mathbb{R}$*

Proof Clearly it has continuous paths and $\mathbb{E}(X_t) = 0$. Now

$$\mathbb{E}(X_s X_t) = \mathbb{E}(\frac{1}{c}B_{c^2s}\frac{1}{c}B_{c^2t}) = s \wedge t$$

and also

$$\sum_1^N \lambda_k X_{t_k} = \sum_1^N \frac{\lambda_k}{c} B_{c^2 t_k}$$

and this is Gaussian since B_t is Gaussian. *Q.E.D.*

Lemma 1.12 (Inversion lemma) *Suppose that B is a Brownian motion on $\mathbb{R}$. Define $X_t = \begin{cases} tB_{\frac{1}{t}} & t > 0 \\ 0 & t = 0 \end{cases}$. Then X is a Brownian motion.*

Proof

$$\sum_1^N \lambda_k X_{t_k} = \sum_1^N \lambda_k t_k B_{\frac{1}{t_k}}$$

which is still Gaussian for $t_k > 0$. If any of the $t_k = 0$ then the addition to the above sum of this term is zero, so we are fine. Clearly $\mathbb{E}(X_t) = 0$ and

$$\mathbb{E}(X_s X_t) = \mathbb{E}(stB_{\frac{1}{s}}B_{\frac{1}{t}}) = ts\frac{1}{t} = s$$

for $s < t$. We also have no problem for $t > 0$ with the continuity of paths. However we need to check that it is continuous at $t = 0$, i.e. that $tB_{\frac{1}{t}} \to 0$ as $t \to 0$, or that $\frac{1}{s}B_s \to 0$ as $s \to \infty$. We expect that $B_t \approx \pm\sqrt{t}$ and so $B_t/t \to 0$ should be clear.

However, we know that $(X_{t_1}, ..., X_{t_N}) = (\hat{B}_{t_1}, ..., \hat{B}_{t_N})$ providing $t_i > 0$ for a Brownian motion $\hat{B}$ and since $\hat{B}_t \to 0$ as $t \to 0$ surely $X_t \to 0$ as well. We pin this down precisely:

$$\begin{aligned}[X_t \to 0 \text{ as } t \to 0] &= \{X_q \to 0, \text{ as } q \to 0, q \in \mathbb{Q}\} \\ &= \{\forall \varepsilon > 0 : \exists \delta > 0, q \in \mathbb{Q} \cap (0, \delta] \implies |X_q| < \varepsilon\} \\ &= \bigcap_{N=1}^\infty \bigcup_{M=1}^\infty \bigcap_{q \in \mathbb{Q} \cap (0, \frac{1}{M}]} \{|X_q| < \frac{1}{N}\}\end{aligned}$$

and so

$$\mathbb{P}[X_t \to 0 \text{ as } t \to 0] = \lim_{N\to\infty} \lim_{M\to\infty} \lim_{k\to\infty} \mathbb{P}\{|X_q| < \frac{1}{N}\} = \mathbb{P}[\hat{B}_t \to 0]$$

where $q_1, q_2, ...$ lists $\mathbb{Q} \cap (0, \frac{1}{M}]$. *Q.E.D.*

We used in the above that

$$A_1 \supseteq A_2 \supseteq ... \text{ then } \mathbb{P}(\cap A_N) = \lim_{N\to\infty} \mathbb{P}(A_N)$$

$$A_1 \subseteq A_2 \subseteq ... \text{ then } \mathbb{P}(\cup A_N) = \lim_{N\to\infty} \mathbb{P}(A_N)$$

Corollary 1.13 $B_t/t \to 0$ *as* $t \to \infty$.

In fact, $B_t/t^\alpha \to 0$ for $\alpha > \frac{1}{2}$ but $\limsup_{t\to\infty} \frac{B_t}{\sqrt{t}} = \infty$ and $\liminf_{t\to\infty} \frac{B_t}{\sqrt{t}} = -\infty$ and so B_t visits every $x \in \mathbb{R}$ infinitely many times.

This brings us nicely into the next subsection.

1.2 Growth rate of paths

Theorem 1.14 (Law of the Iterated Logarithm) *Suppose that B_t is a Brownian motion on $\mathbb{R}$. Then*

$$\limsup_{t\to\infty} \frac{B_t}{\psi(t)} = +1$$

and

$$\liminf_{t\to\infty} \frac{B_t}{\psi(t)} = -1$$

where $\psi(t) = \sqrt{2t\ln(\ln t)}$

$\limsup_{t\to\infty} X_t = \lim_{t\to\infty}\sup_{s\geq t} X_s$ and $\limsup X_t \leq 1$ means that $\forall\varepsilon > 0$ then $\sup_{s\geq t} X_s \leq 1+\varepsilon$ for large t, which is the same as for all $\varepsilon > 0$ X_t is eventually less than $1+\varepsilon$.

$\limsup X_t \geq 1$ if and only if $\forall\varepsilon > 0$ we have $\sup_{s\geq t} X_s \geq 1-\varepsilon$ for large t. which is the same as for all $\varepsilon > 0$ there exists a sequence $s_N \to \infty$ with $X_{s_N} \geq 1-\varepsilon$.

It is on an example sheet that $X_t = e^{-t}B_{e^{2t}}$ then the Law of the Iterated logarithm can be converted to get that $\limsup \frac{X_t}{\sqrt{2\ln t}} = 1$.

We can also compare this to Z_N an iid $N(0,1)$ and then $\frac{\limsup Z_N}{\sqrt{2\ln N}} = 1$

Proof We first show that

$$\mathbb{P}(\limsup \frac{B_t}{\psi(t)} \leq 1) = 1$$

and this is the case if and only if

$$\mathbb{P}(B_t \leq (1+\varepsilon)\psi(t) \text{ for large } t) = 1$$

We first perform a calculation:

$$\begin{aligned}\mathbb{P}(B_t > (1+\varepsilon)\psi(t) \text{ for large } t) &= \mathbb{P}(N(0,t) > (1+\varepsilon)\sqrt{2t\ln(\ln t)}) \\ &= \mathbb{P}(N(0,1) > (1+\varepsilon)\sqrt{2\ln(\ln t)}) \\ &= \int\limits_{(1+\varepsilon)\sqrt{2\ln(\ln t)}}^{\infty} \frac{e^{-\frac{z^2}{2}}}{\sqrt{2t}}dz\end{aligned}$$

Lemma 1.15 (Gaussian Tails)

$$\frac{1}{a}\left(1-\frac{2}{a^2}\right)e^{-a^2/2} \leq \int_a^\infty e^{-z^2/2}dz \leq \frac{1}{a}e^{-a^2/2}$$

Then we get that

$$\begin{aligned}\int\limits_{(1+\varepsilon)\sqrt{2\ln(\ln t)}}^{\infty} \frac{e^{-\frac{z^2}{2}}}{\sqrt{2t}}dz &\leq \frac{1}{(1+\varepsilon)\sqrt{2\ln(\ln t)}\sqrt{2\pi}}e^{-2((1+\varepsilon)\sqrt{2\ln(\ln t)})^2/2} \\ &= \frac{1}{(1+\varepsilon)\sqrt{2\ln(\ln t)}\sqrt{2\pi}}e^{-(1+\varepsilon)^2\ln(\ln t)}\end{aligned}$$

The strategy now is to control B along a grid of times $t_N = \theta^N$ for $\theta > 1$. Then

$$\mathbb{P}(B_{\theta^N} > (1+\varepsilon)\psi(\theta^N)) \leq \frac{1}{\sqrt{2\pi}}\frac{1}{1+\varepsilon}\frac{1}{\sqrt{2\ln(N\ln\theta)}}e^{-(1+\varepsilon)^2\ln(N\ln\theta)} \leq C(\theta,\varepsilon)N^{-(1+\varepsilon)^2}$$

Lemma 1.16 (Borel-Cantelli part 1) *If* $\sum_1^\infty \mathbb{P}(A_N) < \infty$ *then*

$$\mathbb{P}(\text{ only finitely many } A_N \text{ happen }) = 1$$

Proof Let $\chi_{A_N} = \begin{cases} 1 & A_N \\ 0 & A_N^c \end{cases}$ and then the number of A_Ns that occur is given by $\sum_1^\infty \chi_{A_N}$ and so

$$\mathbb{E}[\sum_1^\infty \chi_{A_N}] = \sum_1^\infty \mathbb{E}[\chi_{A_N}] = \sum_1^\infty \mathbb{P}(A_N) < \infty$$

and so $\sum_1^\infty \chi_{A_N}$ is finite a.s. *Q.E.D.* Then by BCI we have that $B_{\theta^N} \leq (1+\varepsilon)\psi(\theta^N)$ for all large N. We now need to control B over (θ^N, θ^{N+1}).

Lemma 1.17 (Reflection trick)

$$\mathbb{P}(\sup_{s\leq t} B_s \geq a) = 2\mathbb{P}(B_t \geq a)$$

for $a \geq 0$

Proof Define $\Omega_0 = \{\sup_{s\leq t} B_s \geq a\}$ and then

$$\begin{aligned}\mathbb{P}(\Omega_0) &= \mathbb{P}(\Omega_0 \cap \{B_t > a\}) + \mathbb{P}(\Omega_0 \cap \{B_t = a\}) + \mathbb{P}(\Omega_0 \cap \{B_t < a\}) \\ &= 2\mathbb{P}(\Omega_0 \cap \{B_t > a\}) \\ &= 2\mathbb{P}\{B_t > a\}\end{aligned}$$

We will carefully justify this later by examining the hitting time $T_a = \inf\{t : B_t = a\}$. We consider $(B_{T_a+t} - a, t \geq 0$ and check that this is still a Brownian motion.

$$\mathbb{P}(T_a \leq t) = \mathbb{P}(\sup_{s\leq t} B_s \geq a) = 2\mathbb{P}(\frac{B_t}{\sqrt{t}} \geq \frac{a}{\sqrt{t}}) = 2\int_a^\infty \frac{1}{\sqrt{2\pi}} e^{z^2/2} dz$$

and also

$$\mathbb{P}(T_a \in dt) = \frac{d}{dt}\mathbb{P}(T_a \leq t) = 2\frac{1}{\sqrt{2\pi}} e^{-a^2/2t} \frac{a}{2} t^{-\frac{3}{2}} = \frac{1}{\sqrt{2\pi t^3}} e^{-a^2/2t} =: \phi(t)$$

and so $\mathbb{E}(T_a) = \infty$. *Q.E.D.*

Thus from this we get that

$$\mathbb{P}(\sup_{s\leq\theta^N} B_s) \geq (1+\varepsilon)\psi(\theta^N)) = 2\mathbb{P}(B_{\theta^N} > (1+\varepsilon)\psi(\theta^N)) \leq 2C(\varepsilon,\theta)N^{-(1+\varepsilon)^2}$$

Borel-Cantelli part 1 still applies and so for large N we have

$$\sup_{s\leq\theta^N} B_s \leq (1+\varepsilon)\psi(\theta^N)$$

This if $t \in [\theta^N, \theta^{N+1}]$ we have

$$\frac{B_t}{\psi(t)} \leq \frac{(1+\varepsilon)\psi(\theta^{N+1})}{\psi(\theta^N)} = (1+\varepsilon)\sqrt{\theta}\frac{\sqrt{\ln((N+1)\ln\theta)}}{\sqrt{\ln(N\ln\theta)}} \to (1+\varepsilon)\sqrt{\theta}$$

and thus we have that

$$\limsup_{t\to\infty} \frac{B_t}{\psi(t)} \leq (1+\varepsilon)\sqrt{\theta}$$

for all $\varepsilon > 0$ and $\theta > 1$ and so

$$\limsup_{t\to\infty} \frac{B_t}{\psi(t)} \leq 1$$

We now show

$$\limsup_{t\to\infty} \frac{B_t}{\psi(t)} \geq (1-\varepsilon)$$

If we choose $t_N = \theta^N$ for $\theta > 1$ then

$$\mathbb{P}(B_{\theta^N} > (1-\varepsilon)\psi(\theta^N)) \geq C(\theta,\varepsilon)N^{-(1-\varepsilon)^2}$$

Lemma 1.18 (Borel-Cantelli part 2) *If* $\sum_1^\infty \mathbb{P}(A_N) = \infty$ *and* $A_N s$ *are independent, then*

$$\mathbb{P}(\text{ infinitely many } A_N \text{ occur}) = 1$$

Proof $Z = \sum_1^\infty \chi_{A_N}$ is the total number of A_Ns that occur. From BCI we get that

$$\mathbb{E}(Z) < \infty \implies \mathbb{P}(Z < \infty) = 1$$

or $\mathbb{E}(e^{-Z}) = 0 \iff \mathbb{P}(Z = \infty) = 1$. Then

$$\begin{aligned}
\mathbb{E}(e^{-\sum \chi_{A_N}}) &= \mathbb{E}(\prod_1^\infty e^{-\chi_{A_N}}) \\
&= \prod_{1^\infty} \mathbb{E}(e^{-\chi_{A_N}}) \\
&= \prod_1^\infty (1-\alpha\mathbb{P}(A_N)) \\
&\leq \prod_1^\infty e^{-\alpha\mathbb{P}(A_N)} \\
&= e^{-\alpha\sum\mathbb{P}(A_N)} \\
&= 0
\end{aligned}$$

Q.E.D.

We use this on $A_N = \{B_{\theta^N} > (1-\varepsilon)\psi(\theta^N)\}$ but these are not independent, but nearly so for large N. We finalise by correcting this. We define $\hat{A}_N = \{B_{\theta^N} - B_{\theta^{N-1}} > (1-\varepsilon)\sqrt{1-\theta^{-1}}\psi(\theta^N)\}$ and these are independent. Then

$$\mathbb{P}(\hat{A}_N) = \mathbb{P}(A_N) \geq C(\theta,\varepsilon)N^{-(1-\varepsilon)^2}$$

BC2 tells us that infinitely many $\hat{A}_N$ do occur a.s., i.e.

$$B_{\theta^N} \geq (1-\varepsilon)\sqrt{1-\theta^{-1}}\psi(\theta^N) - B_{\theta^{N-1}} \geq (1-\varepsilon)\sqrt{1-\theta^{-1}}\psi(\theta^N) - (1+\varepsilon)\psi(\theta^{N-1})$$

and so

$$\frac{B_{\theta^N}}{\psi(\theta^N)} \geq (1-\varepsilon)\sqrt{1-\theta^{-1}} - (1+\varepsilon)\frac{\psi(\theta^{N-1})}{\psi(\theta^N)}$$

and

$$\frac{\psi(\theta^{N-1})}{\psi(\theta^N)} = \frac{\sqrt{\ln((N-1)\ln\theta)}}{\sqrt{\theta}\sqrt{\ln(N\ln\theta)}} \to \sqrt{\theta^{-1}}$$

and so

$$\limsup_{N\to\infty} \frac{B_{\theta^N}}{\psi(\theta^N)} \geq (1-\varepsilon)\sqrt{1-\theta^{-1}} - (1+\varepsilon)\sqrt{\theta^{-1}}$$

and taking θ large and ε small gives the result *Q.E.D.*

We make some observations:

1. Can we do better? $\mathbb{P}(B_t \leq h_t \text{ for large } t) = \begin{cases} 1 \\ 0 \end{cases}$ for h_t deterministic. This is called the 0-1 law, and we see it in week 4. For $h_t = \sqrt{Ct\ln\ln t}$ if $C < 0$ then we get 0 and if $C > 2$ then we get 1. This uses an integral test for h_t.

2. Random walk analogue. Suppose that $X_1, X_2, ...$ are iid with $\mathbb{E}(X_k) = 0$, $\mathbb{E}(X_k^2) = 1$ and $S_N = X_1 + ... + X_N$. then

$$\limsup_{N\to\infty} \frac{S_N}{\sqrt{2N\ln\ln N}} = 1$$

This was proved in 1913 but the proof was long. Was proved in a shorter manner using the Brownian motion result in 1941.

3. $X_t = tB_{\frac{1}{t}}$ is still a Brownian motion and so $\limsup_{t\to\infty} \frac{tB_{\frac{1}{t}}}{\psi(t)}$ or alternatively

$$\lim_{s\to 0} \frac{B_s}{\sqrt{2s\ln\ln(\frac{1}{s})}} = 1$$

and so we have a result about small t behaviour.

4. $\mathbb{P}(B \text{ diff at } 0) = 0$ and if we fix $T_0 > 0$ then define $X_t = B_{T_0+t}$ then this is still a Brownian motion. Thus

Corollary 1.19 $\mathbb{P}(B \textit{ diff at } t_0) = 0$ *for all* t_0.

5. Suppose $U \sim U[0,1]$ is uniform r.v. Then define X_t by the value up until U and then monotone increasing up until 1. Then $\mathbb{P}(X \text{ diff at } t_0) = 1$ but it is not differentiable at all t. and so we cannot easily conclude that Brownian motion id differentiable everywhere

6. **Corollary 1.20**

$$\text{Leb}\{t : B \textit{ is diff at } t\} = 0$$

Proof

$$\mathbb{E}(\int_0^\infty \chi(B \text{ diff at } t)dt) = \int_0^\infty \mathbb{E}\chi(B \text{ diff at } t)dt = 0$$

Q.E.D.

The points where it is differentiable are examples of random exceptional points.

1.3 Regularity

Definition 1.21 *A function $f : [0, \infty) \to \mathbb{R}$ is* α-**Holder continuous**, *for $\alpha \in (0, 1]$, at t if there exists $M, \delta > 0$ such that*

$$|f_{t+s} - f_t| \leq M|s|^\alpha$$

for $|s| \leq \delta$.

The case of $\alpha = 1$ is called Lipschitz.

The aim of the next part is to show that $\mathbb{P}(B \text{ is } \alpha \text{ Holder continuous at all } t \geq 0) = 1$ provided $\alpha < 1/2$ and that $\mathbb{P}(B \text{ is } \alpha \text{ Holder continuous at any } t \geq 0) = 0$ provided $\alpha > 1/2$.

Corollary 1.22 $\mathbb{P}(B$ *differentiable at any* $t) = 0$

The reasons for this are as follows. A differentiable function must lie in some cone as

$$\frac{f(t+s) - f(s)}{s} \to a$$

and so $(f(t+s)-f(s))/s \in (a-\varepsilon, a+\varepsilon)$ for small s and thus $|f(t+s)-f(s)| \leq (|a|+\varepsilon)|s|$ for small s, and so Lipschitz holds with $M = |a| + |eps$.

Proposition 1.23 *Define* $\Omega_{M,\delta} = \{$*for some* $t \in [0, 1], |B_{t+s} - B_t| \leq M|s|$ *for all* $|s| \leq \delta\}$ *and then* $\mathbb{P}(\Omega_{M,\delta}) = 0$ *and thus* $\mathbb{P}(\cup_{M=1}^{\infty} \cup_{N=1}^{\infty} \Omega_{M,\frac{1}{N}}) = 0$

Proof This hasn't been bettered since 1931. Suppose that there exists a $t \in [K/N, (K+1)/N]$ where it is Lipschitz, i.e. $|B_{t+s} - B_t| \leq M|s|$ for $|s| < \delta$. Then if $(K+1)/N, (K+2)N \in [t, t+\delta]$ then

$$|B_{(K+1)N} - B_t| \leq \frac{M}{N} \qquad |B_{(K+2)N} - B_t| \leq \frac{2M}{N}$$

and so by the triangle inequality we get

$$|B_{(K+1)N} - B_{(K+2)/N}| \leq \frac{3M}{N}$$

and then we have

$$\mathbb{P}(\Omega_{M,\delta}) \leq \mathbb{P}\left[\text{ for some } K = 1, ..., N-1 : |B_{(K+1)N} - B_{(K+2)/N}| \leq \frac{3M}{N} \right] \tag{1.1}$$

We first calculate the probability of the event on the right hand side.

$$\mathbb{P}[|N(0, 1/N)| \leq \frac{3M}{N}] = \mathbb{P}[|N(0, 1)| \leq \frac{3M}{\sqrt{N}}] = \int_{-\frac{3M}{\sqrt{N}}}^{\frac{3M}{\sqrt{N}}} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz \leq \frac{6M}{\sqrt{2\pi N}}$$

and then the last part of equation (1.1) is equal to

$$= \mathbb{P}(\cup_1^{N-1} |B_{(K+1)N} - B_{(K+2)/N}| \leq \frac{3M}{N}) \leq \sum_1^{N-1} (\frac{6M}{\sqrt{2\pi N}}) = \frac{6M}{\sqrt{2\pi}} \sqrt{N}$$

but this is not useful, because it does not tend to zero. We modify this by taking more points. We already know that

$$|B_{(K+2)N} - B_{(K+1)/N}| \leq \frac{3M}{N}$$

but we also have that

$$|B_{(K+3)N} - B_{(K+2)/N}| \leq \frac{5M}{N} \qquad |B_{(K+4)N} - B_{(K+3)/N}| \leq \frac{7M}{N}$$

and then we say that

$$\mathbb{P}(\Omega_{M,\delta}) \leq \mathbb{P}\left[\text{ for some } K = 1, ..., N-1 : \left[\begin{array}{ll} & |B_{(K+1)N} - B_{(K+2)/N}| \leq \frac{3M}{N} \\ \text{and} & |B_{(K+3)N} - B_{(K+2)/N}| \leq \frac{5M}{N} \\ \text{and} & |B_{(K+4)N} - B_{(K+3)/N}| \leq \frac{7M}{N} \end{array}\right]\right]$$

with $4/N \leq \delta$. and this is equal to

$$\mathbb{P}(\cup_1^{N-1}\{A \text{ and } B \text{ and } C\}) \leq \sum_1^{N-1} \mathbb{P}(\{A \text{ and } B \text{ and } C\}) \leq \sum_1^{N-1} \frac{D}{n^{3/2}} = \frac{D}{\sqrt{N}} \to 0$$

and so $\mathbb{P}(\Omega_{M,\delta}) = 0$. *Q.E.D.*

We now turn to the other statement, that B is α-Holder continuous if $\alpha \in (0, 1/2)$.

Theorem 1.24 (Kolmogorov's continuity) *Suppose $(X_t, t \in [0,1])$ has continuous paths and satisfies*

$$\mathbb{E}[|X_t - X_s|^p] \leq C|t-s|^{1+\gamma}$$

for some $\gamma, p > 0$. Then X has α-Holder paths for $\alpha \in 0, \frac{\gamma}{p})$.

We use this for X a Brownian motion. In this case

$$\mathbb{E}[|B_t - B_s|^p] = \mathbb{E}[|N(0, t-s)|^p] = \sqrt{t-s}^p \mathbb{E}[|N(0,1)|^p] = c_p|t-s|^{p/2}$$

We then have that Brownian motion has α-Holder continuous paths of orders

$$\alpha < \frac{p/2 - 1}{p} = \frac{1}{2} - \frac{1}{p}$$

and so Holder with any $\alpha < \frac{1}{2}$ but not $\alpha = \frac{1}{2}$

Lemma 1.25 (Markov Inequality) *Suppose that $Z \geq 0$ is a non negative random variable. Then*

$$\mathbb{P}(Z \geq a) \leq \frac{1}{a}\mathbb{E}(Z)$$

Proof

$$\mathbb{E}(Z) = \mathbb{E}(Z\chi_{\{Z<a\}}) + \mathbb{E}(Z\chi_{\{Z\geq a\}}) \geq \mathbb{E}(Z\chi_{\{Z\geq a\}}) \geq a\mathbb{P}(Z \geq a)$$

Q.E.D.

We use this as follows

$$\mathbb{P}(|X_t - X_s| \geq a) = \mathbb{P}(|X_t - X_s|^p \geq a^p) \leq \frac{1}{a^p}\mathbb{E}(|X_t - X_s|^p) \leq \frac{C|s-t|^{1+\gamma}}{a^p}$$

Suppose we have a grid with size 2^{-N}. Then we define $A_N = \cup_{K=1}^{2^N}\{|X_{k/2^N} - X_{(K-1)/2^N}| > \frac{1}{2^{N\alpha}}\}$
We then estimate this

$$\mathbb{P}(A_N) \leq \sum_1^{2^N} \mathbb{P}\{|X_{k/2^N} - X_{(K-1)/2^N}| > \frac{1}{2^{N\alpha}}\} \leq C2^N 2^{-N(1+\gamma)} 2^{N\alpha p} = 2^{-N(\gamma-\alpha p)}$$

and thus we need $\gamma > \alpha p$ or $\alpha < \gamma/p$. This is the key idea of the proof. We now have $\sum_1^\infty \mathbb{P}(A_N) < \infty$ and so by Borel Cantelli I we have that only finitely many A_Ns occur, i.e. there exists $N_0(\omega)$ such that

$$|X_{K/2^N} - X_{(K-1)/2^N}| < \frac{1}{2^{N\alpha}} \text{ for all } K = 1, ..., 2^N \text{ and for } N > N_0$$

We also know that $\mathbb{P}(N_0 < \infty) = 1$. We now fix ω such that $N_0(\omega) < \infty$. We can control X on $\mathcal{D} = \{$ dyadic rationals i.e. of form $K/2^M\}$. Fix $t, s \in \mathcal{D}$ with

$$1/2^M \leq |t - s| \leq 1/2^{M-1}$$

where $M \geq N_0$. We have two cases. Either we straddle two points $K/2^M$ and $(K+1)/2^M$ or we straddle only one point $K/2^M$. We consider the first case, and leave the second case to the reader. We have

$$t = \frac{K+1}{2^M} + \frac{1 \text{ or } 0}{2^{M+1}} + ... + \frac{1 \text{ or } 0}{2^{M+L}}$$

and then we have

$$|X_t - X_{(K+1)/2^M}| \leq \frac{K+1}{2^M} + \frac{1}{2^{M+1}} + ... + \frac{1}{2^{M+L}} \leq \frac{1}{2^{(M+1)\alpha}} \frac{1}{1-2^{-\alpha}}$$

and similarly

$$|X_s - X_{(K)/2^M}| \leq \frac{1}{2^{(M+1)\alpha}} \frac{1}{1-2^{-\alpha}}$$

and thus

$$\begin{aligned} |X_s - X_t| &\leq |X_s - X_{(K)/2^M}| + |X_{(K)/2^M} - X_{(K+1)/2^M}| + |X_t - X_{(K+1)/2^M}| \\ &\leq \frac{1}{2^{(M+1)\alpha}} \frac{2}{1-2^{-\alpha}} + \frac{1}{2^{M\alpha}} \\ &\leq \frac{C}{2^{M\alpha}} \end{aligned}$$

and thus using the assumptions of Kolmogorov we have this less than $C|t-s|^\alpha$. Invoking continuity gives the result.

2 Brownian motion as a Markov Process

Roughly, the future $(X_s; s \geq t)$ is independent of the past $(X_s : s \leq t)$ conditional on the present X_t.

Definition 2.1 *Suppose that $\mathcal{F}, \mathcal{G}$ two σ-fields on Ω are called* **independent** *if $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$ for all $A \in \mathcal{F}$ and $B \in \mathcal{G}$.*

Two variables X and Y are called **independent** *if*

$$\mathbb{P}(X \in C_1, Y \in C_2) = \mathbb{P}(X \in C_1)\mathbb{P}(Y \in C_2)$$

or equivalently

$$\mathbb{E}(h(X)g(Y)) = \mathbb{E}(h(X))\mathbb{E}(g(Y))$$

for all measurable bounded functions h, g.

They are equivalent as follows. We can take $h = \chi_{C_1}$ and $g = \chi_{C_2}$ and then the two statements are the same. We then take simple functions and limits of simple functions to get any functions, using the standard machine.

Definition 2.2 *$\sigma(X)$, called the* **σ-field generated by** *X is defined as*

$$\{X^{-1}(C) : C \text{ measurable }\}$$

Note that now X is independent of Y if and only if $\sigma(X)$ is independent of $\sigma(Y)$.

Definition 2.3 *$\sigma(X_t, t \in I)$ called the* **σ-field generated by** *X_t is defined to be*

$$\sigma\{X_t^{-1}(C) : t \in I, C \text{ measurable }\}$$

i.e. the smallest σ field containing the above sets.

Theorem 2.4 (Markov property of Brownian motion 1) *Suppose that B is a Brownian motion, and fix $t_0 > 0$. Define $X_t = B_{t+t_0} - B_{t_0}$ and then $\sigma(X_t : t \geq 0)$ is independent of $\sigma(B_t : t \leq t_0)$.*

Note that this is very close to independent increments.

For example $\int_0^{t_0} B_s ds$ is independent of $\int_0^{t_1} X_s ds$. Need to check that $\int_0^{t_0} B_s ds$ is measurable with respect to $\sigma(B_t)$.

$$\int_0^{t_0} B_s ds = \lim_{N\to\infty} \sum_1^N \frac{1}{N} B_{K/N}$$

which is measurable since the limit and sum of measurable functions is measurable.

A second example is $\sup_{t\leq s} B_t$ is independent of $\sup_{t\leq t_1} X_t$. We can write this as $\sup_{t\leq s} B_t = \sup_{q\in[0,1]\cap\mathbb{Q}} B_q = \lim_{N\to\infty} \max\{B_{q_1}, ..., B_{q_N}\}$ where q_i lists $[0,1] \cap \mathbb{Q}$. Now

$$w \mapsto (B_{q_1}(w), ..., B_{q_N}(w)) \mapsto \max B_{q_i}(w)$$

and the latter is continuous, and the former is measurable, since if $C = C_1 \times ... \times C_N$ then $\{w : (B_{q_1}(w), ..., B_{q_N}(w)) \in C\}$ is measurable.

Definition 2.5 *A collection of subsets A_0 is called a π-system if it is closed under finite intersections.*

Lemma 2.6 *If $\mathcal{F}_0$ and $\mathcal{G}_0$ are two different π-systems and $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$ for all $A \in \mathcal{F}_0$ and $B \in G_0$ then $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$ holds for $\sigma(\mathcal{F}_0)$ and $\sigma(G_0)$, i.e. independence of π-systems gives independence of generated σ-fields.*

For us, $\sigma(X_t, t \geq 0) = \sigma\{X_t^{-1}(C) : t \geq 0, C \in \mathcal{B}(\mathbb{R})\}$ and this is generated by the π-system $\{X_{t_1}^{-1}(C_1) \cap ... \cap X_{t_N}^{-1}(C_N) : t_1, ..., t_N \geq 0, C_K \in \mathcal{B}(\mathbb{R})\}$
Proof (of theorem 2.4) By the above lemma, we need only check that $(X_{t_1}, ..., X_{t_N})$ is independent of $(B_{s_1}, ..., B_{s_N})$. We have

$$\begin{aligned}\mathbb{E}\left[e^{i\sum \lambda_k B_{t_k}} e^{i\sum \mu_k (B_{T+t_k} - B_{t_k})}\right] &= \mathbb{E}\left[e^{-\sum \hat{\lambda}_k (B_{s_k} - B_{s_{k-1}})} e^{i\sum \hat{\mu}_k (B_{T+t_k} - B_{T+t_{k-1}})}\right] \\ &= \mathbb{E}\left[e^{i\sum \lambda_k B_{t_k}}\right] \mathbb{E}\left[e^{i\sum \mu_k (B_{T+t_k} - B_{t_k})}\right]\end{aligned}$$

by independent increments. *Q.E.D.*

Proposition 2.7 *The following are true*

1. *$a \leq b \leq c \leq d$ then $\mathbb{P}(\max_{[a,b]} B_t \neq \max_{[c,d]} B_t) = 1$*
2. *$\mathbb{P}(\exists! t^\star \in [a, b] : B_{t^\star} = \max_{[a,b]} B_t) = 1$*
3. *Local maxima are dense in $[0, \infty)$.*

Proof We can get 2 from one by using that $\mathbb{P}(\max_{[a,b]} B_t \neq \max_{[c,d]} B_t : \forall a \leq b \leq c \leq d$ rational $) = 1$ We can get 3 from 2 by doing this for all rational (a, b) simultaneously. We now prove 1. Take $X_t = B_{c+t} - B_c$, and let $Y = \max_{[c,d]}(B_t - B_c) = \max_{[0,d-c]} X_t$ which is independent of $\sigma(B_s, s \leq c)$. Take $Z = \max_{[a,b]}(B_t - B_c)$ and we aim to show that $\mathbb{P}(Z \neq Y) = 1$. Y and Z are independent by the Markov property but Y has a density. By the reflection principle, this implies that $\mathbb{P}(Y = Z) = 0$. *Q.E.D.*

At the last part of this, we have used that if $\mathbb{E}(g(Y)h(Z)) = \mathbb{E}(g(Y))\mathbb{E}(h(Z))$ and if $\nu(dy) = \mathbb{P}(Y \in dy)$ and $\mu(dz) = \mathbb{P}(Z \in dz)$ then

$$\mathbb{E}(\phi(Y, Z)) = \int\int \phi(y, z)\nu(dy)\mu(dz)$$

This can be shown by the standard machine, with first taking $\phi(Y, Z) = \begin{cases} 1 & Y = Z \\ 0 & Y \neq Z \end{cases}$

Definition 2.8 *We define the* **Brownian filtration** *by $\mathcal{F}_t^B := \sigma(B_s : s \leq t)$, and we also define $\mathcal{F}_{t+}^B = \cap_{s>t}\mathcal{F}_s^B$ and the* **germ field** *is $\mathcal{F}_{0+}^B$.*

Note that if $\frac{dB}{dt}(t)$ existed then it would be in $\mathcal{F}_{t+}^B$. However, $\limsup_{t\to 0} \frac{B_t}{h(t)}$ is in $\mathcal{F}_{0+}^B$, since this limsup is $\mathcal{F}_\delta^B$ measurable for any $\delta > 0$.

Theorem 2.9 (Markov property version 2) *Fix $T \in \mathbb{R}$ and define $X_t = B_{T+t} - B_T$. Then $\sigma(X_t, t \geq 0)$ is independent of $\mathcal{F}_{t+}^B$.*

Corollary 2.10 *$T = 0$ so $\sigma(B_t, t \geq 0)$ is independent of $\mathcal{F}_{0+}^B$ so is independent of itself.*

Corollary 2.11 *$\mathcal{F}_{0+}^B$ is a 0/1 field.*

Thus if X is $\mathcal{F}^B_{0+}$ measurable then $\{\omega : Z(\omega) \leq c\}$ has probability 0 or 1. Thus Z is a constant almost surely.

Example 2.1 (Lebesgue's Thorn) *Suppose that B is a d dimensional Brownian motion and F is an open set in $\mathbb{R}^d$. Take $\tau = \inf\{t : B_t \in F\}$. We claim that $\{\tau = 0\} \in \mathcal{F}^B_{0+}$. Indeed $\{\tau = 0\} = \cap_{N=N_0}^\infty \cup_{q\in\mathbb{Q}\cap[0,1/N]} \{B_q \in F\} \in \mathcal{F}^B_{1/N_0}$ for all N_0. Lebesgue's thorn is an example of F, e.g. $F = \{(x,y,z) : \sqrt{y^2+z^2} \leq f(x)\}$ a volume of rotation. Here $\mathbb{P}(\tau > 0) = 1$ for thin thorns, or is 0 for thick thorns.*

Proof (of theorem 2.9) Take $A \in \mathcal{F}^B_{t+}$ and $B \in \sigma(X_s, s \geq 0)$. The π-system lemma says we can choose $B = \{X_{s_i} : i = 1, ..., m\}$, i.e. enough to check that

$$\mathbb{E}[e^{-\sum \lambda_k X_{s_k}} e^{i\theta\chi_A}]$$

splits as a product of the two expectations.

Let $X^\varepsilon_s = B_{t+\varepsilon+s} - B_{t+\varepsilon}$ and note that this is independent of $\mathcal{F}^B_{t+\varepsilon}$. Then

$$\mathbb{E}[e^{-\sum \lambda_k X^\varepsilon_{s_k}} e^{i\theta\chi_A}] = \mathbb{E}[e^{-\sum \lambda_k X^\varepsilon_{s_k}}]\mathbb{E}[e^{i\theta\chi_A}]$$

by the Markov property version 1. Letting $\varepsilon \to 0$ and then using the DCT we get the required result. *Q.E.D.*

Example 2.2 (Shakespeare problem) *Suppose we have 2 dimensional Brownian motion.*

Then assume that

$$\mathbb{P}(B_t, t \in [0,1] \text{ traverses a tube } A) = p_0 > 0$$

and then

$$\mathbb{P}(\frac{1}{\sqrt{2}}B_t, t \in [0,1/2] \text{ traverses a tube } \frac{1}{\sqrt{2}}A) = p_0$$

then let

$$A_N = \{\frac{1}{2^{N/2}}B_t : t \in [0, 2^{-N}] \text{ traverses } 2^{-N/2}A\}$$

and then $\mathbb{P}(A_N) = p_0$ by scaling. Then note that $A_N \in \mathcal{F}^B_{2^{-N}}$ and let $\Omega_0 = \cap_{M=M_0}^\infty \cup_{N=M}^\infty A_N = [A_N$ i.o. $] \in \mathcal{F}^B_{2^{-M_0}}$ and so $\Omega_0 \in \mathcal{F}^B_{0+}$ and now

$$\mathbb{P}(\Omega_0) = \lim_{M\to\infty} \mathbb{P}(\cup_{N=M}^\infty A_N) \geq p_0$$

and so $\mathbb{P}(\Omega_0) = 1$.

2.1 Markov transition functions

You may know that

$$\mathbb{P}(X_0 = x_0, ..., X_N = x_N)p(x_0, x_1)...p(x_N, x_{N-1}$$

and this is equivalent to

$$\mathbb{P}(X_N = x_N | X_0 = x_0, ..., X_{N-1} = x_{N-1}) = p(x_n, x_{N-1})$$

Our aim is to come up with something similar in the continuous case. We intuitively think of the following as the probability that we start at x and end up in dy in time t.

Definition 2.12 *A Markov transition kernel is a function $p : (0, \infty) \times \mathbb{R}^d \times \mathcal{B}(\mathbb{R}^d) \to [0, 1]$ such that*

1. *$A \mapsto p_t(x, A)$ is a probability measure on $\mathcal{B}(\mathbb{R}^d)$*
2. *$x \mapsto p_t(x, A)$ is measurable.*

Definition 2.13 *A Markov process $(X_t, t \geq 0)$ on $\mathbb{R}^d$ with transition kernel stared at x means that*

$$\mathbb{P}(X_{t_1} \in A_1, ..., X_{t_N} \in A_N) = \int_{A_1,...,A_N} p_{t_N - t_{N-1}}(x_{N-1}, dx_N)...p_{t_1}(x, dx_1)$$

Observe that we have that

$$\mathbb{E}(f(X_{t_1}, ..., X_{t_N})) = \int_{\mathbb{R}^N} f(x_1, ..., x_N) p_{t_N - t_{N-1}}(x_{N-1}, dx_N)...p_{t_1}(x, dx_1)$$

Example 2.3

1. *Brownian motion with $p_t(x, dy) = \frac{1}{\sqrt{2\pi t}} e^{-(x-y)^2/2t} dy$*
2. *$X_t = B_t + x + ct$*
3. *Reflected Brownian motion $X_t = |x + B_t|$*
4. *Absorbed Brownian motion $X_t = \begin{cases} x + B_t & t < \tau \\ 0 & t \geq \tau \end{cases}$ where $\tau = \inf\{t : B_t = 0\}$*
5. *Ornstein-Uhlenbeck processes*
6. *Radial Brownian motion. Suppose that B_t is Brownian motion on $\mathbb{R}^d$ and define $X_t = \sqrt{\sum_{i=1}^d (B_t^i)^2}$*
7. *$dX = \mu(X)dt + \sigma(X)dB_t$ an SDE.*
8. *A Brownian Bridge is NOT an example. However it is a time inhomogeneous Markov process, where p is a function of two times.*

Definition 2.14 *We define $B_t^x = x + B_t$ i.e. Brownian motion starting at x.*

We check that $(B_t^x, t \geq 0)$ is a Markov process started at x with kernel $p_t(x, dy) = q_t(y - x)dy$ where $q_t(z) = \frac{1}{\sqrt{2\pi t}} e^{-z^2/2t}$ The short proof is as follows:

$$\begin{aligned}\mathbb{P}(B_{t_1}^x \in dx_1, ..., B_{t_N}^x \in dx_N) &= \mathbb{P}(B_{t_1}^x \in dx_1, B_{t_2}^x - B_{t_1}^x \in d(x_1 - x_2), ..., B_{t_N}^x - B_{t_{N-1}}^x \in d(x_N - x_{N-1})) \\ &= q_{t_1}(x_1 - x)dx_1 q_{t_2 - t_1}(x_2 - x_1)dx_2 ... q_{t_N - t_{N-1}}(x_N - x_{N-1})dx_N\end{aligned}$$

and then we should integrate both sides over $A_1, ..., A_N$.

The longer version is as follows:

$$\mathbb{E}(f(B_{t_1}^x, ..., B_{t_N}^x)) = \mathbb{E}(g(B_{t_1}^x - x, B_{t_2}^x - B_{t_1}^x, ..., B_{t_N}^x - B_{t_{N-1}}^x))$$

if

$$
\begin{aligned}
g(y_1, ..., y_N) &= f(x + y_1, x + y_1 + y_2, ...) \\
&= \int_{\mathbb{R}^N} g(y_1, ..., y_N) q_{t_1}(y_1)...q_{t_N - t_{N-1}}(y_N) dy_1...dy_N \\
&= \int_{\mathbb{R}^N} f(x_1, ..., x_N) q_{t_1}(x_1 - x)...q_{t_N - t_{N-1}}(x_N - x_{N-1}) J dx_1...dx_N
\end{aligned}
$$

where J is the jacobian of the change of variables $y_1 = x_1 - x$,...,$y_N = x_N - x_{N-1}$, and equals 1.

Definition 2.15 *$X_t, t \geq 0$ is a Markov process with transition kernel p means*

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

for all $s < t$.

This is shorter than the previous definition, and implies it by an induction argument, which we will do.

2.1.1 Conditional Expectation

Suppose that we have a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a random variable $X : \Omega \to \mathbb{R}$. Suppose $\mathcal{G} \subset \mathcal{F}$ is a sub σ field.

Our aim is to take $\mathbb{E}(X|\mathcal{G})$ to be the closest $\mathcal{G}$ measurable function to X.

The natural place for this is $L^2(\Omega, \mathcal{F}, \mathbb{P}) = \{X : \Omega \to \mathbb{R} | \mathbb{E}(X^2) < \infty\}$ and this is an inner product space with inner product $(X, Y) = \mathbb{E}(XY)$ and a corresponding norm $||X||_2 = \sqrt{\mathbb{E}(X^2)}$.

For $X \in L^2(\mathcal{F})$ there exists a unique $Y \in L^2(\mathcal{G})$ such that

$$(X - Y, Z) = 0$$

for all $Z \in L^2(\mathcal{G})$. This means

$$\mathbb{E}((X - Y)Z) = 0 \iff \mathbb{E}(XZ) = \mathbb{E}(YZ)$$

for all $Z \in L^2(\mathcal{G})$. It is enough though to check for $Z = \chi_A$ for $A \in \mathcal{G}$, by the standard machine of measure theory. This expectation is then

$$\int_A X d\mathbb{P} = \int_A d\mathbb{P} \tag{2.1}$$

and we write $Y = \mathbb{E}(X|\mathcal{G})$.

Proposition 2.16 *For $X \in L^2(\mathcal{F})$ there is a unique $Y \in L^2(\mathcal{G})$ satisfying* (2.1)*. This can be improve to $X \in L^1(\mathcal{F})$ or to $X \geq 0$.*

There are several special cases

1. $\mathcal{G} = \{\emptyset, \Omega\}$ and then $\mathbb{E}(X|\mathcal{G}) = \mathbb{E}(X)$
2. $\mathcal{G} = \{\emptyset, A, A^c, \Omega\}$ and then $\mathbb{E}(X|\mathcal{G}) = \begin{cases} y_1 & A \\ y_2 & A^c \end{cases}$. If $X = \chi_B$ then $y_1 = \mathbb{P}(B|A)$ and $y_2 = \mathbb{P}(B|A^c)$ and this is like an extension to Bayes formula

3. $\mathcal{G} = \mathcal{F}$ and then $\mathbb{E}(X|\mathcal{G}) = X$.

Lemma 2.17 *Suppose X is $\mathcal{G}$ measurable and Z is independent of $\mathcal{G}$, and $\phi : \mathbb{R}^2 \to \mathbb{R}$ is bounded and measurable. Then*

$$\mathbb{E}(\phi(X,Z)|\mathcal{G}) = \mathbb{E}(\phi(x,Z))|_{x=X}$$

This can be proved using the measure theory machine.

Example 2.4 *1.* $\mathbb{E}(B_t|\mathcal{F}_s^B) = \mathbb{E}(B_s + (B_t - B_s)|\mathcal{F}_s^B) = B_s$

2. $\mathbb{E}(e^{B_t}|\mathcal{F}_s^B) = \mathbb{E}(e^{B_t - B_s}e^{B_s}|\mathcal{F}_s^B) = e^{B_s}e^{(t-s)/2}$

3. $\mathbb{E}(B_t^2|\mathcal{F}_s^B) = \mathbb{E}(B_s^2 + 2B_s(B_t - B_s) + (B_t - B_s)^2|\mathcal{F}_s^B) = B_s^2 + 2B_s\mathbb{E}(B_t - B_s) + \mathbb{E}(B_t - B_s)^2 = B_s^2 + t - s$

4. (X, Y) a Gaussian vector on $\mathbb{R}^2$ with mean 0 and $Y \neq 0$. We then postulate that

$$\mathbb{E}(X|\sigma(Y)) = \alpha Y$$

but we need to find α. We do this as follows. Set $X = \alpha Y + (X - \alpha)Y$ and we need to check that $\mathbb{E}(Y(X - \alpha Y)) = 0$. Thus we choose $\alpha = \frac{\mathbb{E}(XY)}{\mathbb{E}(Y^2)}$, and so then $\mathbb{E}(X|\sigma(Y)) = \mathbb{E}(\alpha Y + (X - \alpha)Y|\sigma(Y)) = \alpha Y + 0$.

5. We can do a similar thing for $\mathbb{E}(\int_0 1 B_s ds|\sigma(B_1)$ and set it equal to αB_1. We then get $\alpha = 1/2$.

Lemma 2.18 *$X \geq 0$ implies that $\mathbb{E}[X|\mathcal{G}] \geq 0$ a.s.*

Proof $Y = \mathbb{E}[X|\mathcal{G}]$ means that $\mathbb{E}(YZ) = \mathbb{E}(XZ)$ for measurable Z. Let $A = \{\omega : Y(\omega) < 0\}$. Then

$$0 \leq \int_A X d\mathbb{P} = \int_A Y d\mathbb{P} \leq 0$$

and A is $\mathcal{G}$ measurable, so $\mathbb{P}(A) = 0$. *Q.E.D.*

We check that Brownian motion is Markov with this new definition. Let $X_t = x + B_t$. Note that $\mathcal{F}_s^X = \mathcal{F}_s^B$ and so

$$\begin{aligned}\mathbb{P}(X_t \in dy|\mathcal{F}_s^X) &= \mathbb{P}(x + B_s + B_t - B_s \in dy|\mathcal{F}_s^B) \\ &= \mathbb{P}(Z + N(0, t-s) \in dy)|_{z=x+B_s} = q_{t-s}(y - z)dy|_{z=x+B_s} \\ &= q_{t-s}(y - X_s)dy\end{aligned}$$

Suppose now that $X_t = |B_t^x|$ and then $\mathcal{F}_s^X \subset \mathcal{F}_s^B$ and we guess that the transition kernel is given by

$$\hat{q}_{t-s}(y - x)dy = q_{t-s}(y - x)dy + q_{t-s}(-y - x)dy$$

since we can either reach dy or $-dy$ at time t. We consider $\mathbb{P}(X_t \in dy|\mathcal{F}_s^X) = \mathbb{P}(|B_t^x| \in dy|\mathcal{F}_s^X)$. Then using the tower property we get

$$\mathbb{P}(|x + B_t| \in dy|\mathcal{F}_s^B) = \hat{q}_{t-s}(y - (x + B_s))dy = \hat{q}_{t-s}(y - |x + B_s|)dy = \hat{q}_{t-s}(y - X_s)dy$$

and since this is already X_s measurable, conditioning on X_s does nothing.

However there is a warning. Is $X_t = f(B_t^x)$ still Markov? The answer is usually not. In the above case, were were just lucky to get this. For example, consider radial Brownian motion, $X_t = \sqrt{B_1(t)^2 + ... + B_d(t)^2}$ and we want to show that

$$\mathbb{E}(f(X_{t_1}, ..., X_{t_N})) = \int f(x_1, ..., x_N) p_{t_N - t_{N-1}}(dx_N - x_{N-1}) ... p_{t_1}(dx_1 - x)$$

The measure theory machine says that it is enough to check for $f(x_1, ..., x_n) = \prod_1^n \phi_k(x_k)$. We show this by induction:

$$\begin{aligned} \mathbb{E}(\prod_1^n \phi_k(x_k)) &= \mathbb{E}(\mathbb{E}(\prod_1^n \phi_k(x_k) | \mathcal{F}_{t_{N-1}^X})) \\ &= \mathbb{E}(\prod_1^{n-1} \phi_k(x_k) \mathbb{E}(\phi_n(X_{t_n} | \mathcal{F}_{t_{n-1}^X})) \\ &= \mathbb{E}(\prod_1^{n-1} \phi_k(x_k)) \int \phi_n(x_n) p_{t_n - t_{n-1}}(x_{t_{n-1}}, dx_n) \end{aligned}$$

and then we use induction hypothesis.

2.2 Strong Markov Process.

Define $X_s(w) = B_{T+s} - B_T = B_{T(w)+s}(w) - B_{T(w)}(w)$ and is it still a Brownian motion? This is the case if $T_a = \inf\{t : B_t = a\}$, but not if $T = \sup\{t \leq 1 : B_t = 0\}$. In other words, T must not look into the future.

Definition 2.19 *$(\mathcal{F}_t, t \geq 0)$ is called a* **filtration** *on $(\Omega, \mathcal{F}, \mathbb{P})$ if $\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}$ for $s \leq t$.*

The key example is $(X_t, t \geq 0)$ and $\mathcal{F}_t^X = \sigma(X_s : s \leq t)$.

Definition 2.20 *$T : \Omega \to [0, \infty]$ is called a* **stopping time** *for a filtration $(\mathcal{F}_t, t \geq 0)$ if $\{T \leq t\} \in \mathcal{F}_t$ for all $t \geq 0$.*

The key example is $(\mathcal{F}_t^B, t \geq 0)$ and $T_a = \inf\{t : b_t = a\}$. We give two justifications as to why this is a stopping time below.

$$\{T_a \leq t\} = \{\sup_{s \leq t} B_s \geq a\} = \{\sup_{s \in \mathbb{Q} \cap [0,t]} B_s \geq a\} \in \mathcal{F}_t^B$$

since $\sup B_s$ is a measurable object.

$$\{T_a \leq t\} = \cap_{N=1}^{\infty} \cup_{q \in \mathbb{Q} \cap [0,t]} \{|B_q - a| \leq \frac{1}{N}\}$$

On the example sheet, we have $T_K = \inf\{t : B_t = K\}$ for $K \subset \mathbb{R}^d$ closed is a stopping time, but if K is open then it is not a stopping time, as you cannot write it in $\mathcal{F}_t^B$. However, you can write it in $\mathcal{F}_{t+}^B$.

Theorem 2.21 *B a Brownian motion, $\mathcal{F}_t^B$, T is an $\mathcal{F}_t^B$ stopping time with $T < \infty$. Then define $X_s = B_{T+s} - B_T$ and it is a Brownian motion, and independent of $\mathcal{F}_T^B$*

Definition 2.22 (Information up to a stopping time) *Suppose that T is a stopping time. Then we define*

$$\mathcal{F}_T = \{A : A \cap \{T \leq t\} \in \mathcal{F}_t \text{ for all } t \geq 0\}$$

We need to check that $\mathcal{F}_T$ is indeed a σ-field, that if $S \leq T$ then $\mathcal{F}_S \subseteq \mathcal{F}_T$ for two stopping times and that T is $\mathcal{F}_T$ measurable.

To show the first of these note that $\emptyset, \Omega$ are clearly in $\mathcal{F}_T$. Then if $A \in \mathcal{F}_T$

$$A^c \cap \{T \leq t\} = \{T \leq t\} \setminus (A \cap \{T \leq t\}) \in \mathcal{F}_t$$

Now if $A_1, A_2, ... \in \mathcal{F}_T$ then

$$\bigcap_1^\infty A_N \cap \{T \leq t\} = \bigcap_1^\infty (A_N \cap \{T \leq t\})$$

To show the last note that it is enough to check that $\{T \leq s\} \in \mathcal{F}_T$ for all $s \in \mathbb{R}$. Then

$$\{T \leq s\} \cap \{T \leq t\} = \{T \leq \min(s,t)\} \in \mathcal{F}_{\min(s,t)} \subseteq \mathcal{F}_t$$

We now justify the part of the Reflection principle that before was somewhat hand wavy. This here is rigorous though.

$$\mathbb{P}(T_a \leq t) = \mathbb{P}(T_a \leq t, B_t \geq a) + \mathbb{P}(T_a \leq t, B_t \leq a)$$

and the former is $\mathbb{P}(B)t \geq a)$ and the latter we consider below:

$$\begin{aligned}
\mathbb{P}(T_a \leq t, B_t \leq a) &= \mathbb{P}(T_a \leq t, X_{t-T_a} \leq 0) \\
&= \mathbb{E}(\mathbb{P}(T_a \leq t, X_{t-T_a} \leq 0 | \mathcal{F}^B_{T_a})) \\
&= \mathbb{E}(\chi_{T_a \leq t} \mathbb{P}(X_{t-s} \leq 0)|_{s=T_a}) \\
&= \frac{1}{2}\mathbb{P}(T_a \leq t)
\end{aligned}$$

Proof (of theorem 2.21) We first assume that T is discrete, i.e. $\Omega = \cup_k \{T = t_k\}$, i.e. T only takes the values $t_1 < t_2 < ....$ Pick $A \in \mathcal{F}^B_T$ and look at

$$\begin{aligned}
\mathbb{E}(F(B_{t+T} - B_T)\chi_A) &= \sum_k \mathbb{E}(\chi_A \chi_{\{T=t_k\}} F(B_{t_k+t} - B_{t_k}, t \geq 0)) \\
&= \sum_k \mathbb{E}(\chi_{A \cap \{T=t_k\}} F(B_{t_k+t} - B_{t_k}, t \geq 0))
\end{aligned}$$

but $A \cap \{T = t_k\} = A \cap \{T \leq t_k\} \setminus A \cap \{T \leq t_{k-1}\} \in \mathcal{F}^B_{t_k}$ and so by the first Markov property we get the above equal to

$$\sum_k \mathbb{E}(\chi_A \chi_{\{T=t_k\}}) \mathbb{E}(F(B_{t_k+t} - B_{t_k})) = \mathbb{E}(F(B_t, t \geq 0))\mathbb{P}(A)$$

We now consider general $T < \infty$. We approximate it by discrete stopping times. Define

$$T_N = \frac{j}{N} \qquad\qquad \text{if } T \in [\frac{j-1}{N}, \frac{j}{N})$$

Then $T_N \to T$ as $N \to \infty$ and $B_{T_N} \to B_T$ due to the continuity of B. We now must verify that T_N is a stopping time

$$\{T_N \leq t\} = \{T < \frac{j-1}{N}\} \in \mathcal{F}^B_{\frac{j-1}{N}} \subseteq \mathcal{F}^B_t$$

if $t \in [\frac{j-1}{N}, \frac{j}{N})$. Then the above implies that $B_{T_N+t} - B_{T_N}$ is a Brownian motion and is independent of $\mathcal{F}^B_{T_N} \supseteq \mathcal{F}^B_T$ If we let $N \to \infty$ then considering characteristic functions gives

$$\mathbb{E}(e^{i\sum\theta_k(B_{T_N+s_k}-B_{T_N})}\chi_A) = \mathbb{E}(e^{i\sum\theta_k(B_{T_N+s_k}-B_{T_N})} = \mathbb{E}(e^{i\sum\theta_k(B_{T+s_k}-B_T)}\mathbb{P}(A)$$

where we have used the DCT in the ultimate equality. *Q.E.D.*

We consider the hitting time process $(T_a : a \geq 0)$. Let $S_t = \max_{s\leq t} B_s$ and then $a \mapsto T_a$ is trying to be the inverse function to S_t but T_a has jumps in any interval (b, c) so this isnt a true inverse.

Proposition 2.23 *$T_b - T_a$ is independent of $(T_c, c \leq a)$ and $T_b - T_a = T_{b-a}$*

This is an example of an independent increments process.
Proof Define $X_t = B_{T_a+t} - a$ and this is a new Brownian motion and is independent of $\mathcal{F}^B_{T_a}$ Then $T_b - T_a = \inf\{t : X_t = b - a\} \stackrel{D}{=} T_{b-a}$ and T_c is measurable with respect to $\mathcal{F}^B_{T_c} \subseteq \mathcal{F}^B_{T_a}$. *Q.E.D.*

Proposition 2.24 *Let $Z = \{t : B_t = 0\}$. Then*

1. *$Leb(Z) = 0$*
2. *There are no isolated points in Z.*
3. *Z is uncountable*
4. *The Hausdorff dimension of Z is $1/2$.*

Recall that $t \in Z$ is **isolated** if there exists $\varepsilon > 0$ such that $(t - \varepsilon, t + \varepsilon) \cap Z = \{t\}$.
Proof

$$\text{leb}(Z) = \int_0^\infty \chi(t \in Z)dt = \int_0^\infty \chi(B_t = 0)dt$$

and $\mathbb{E}(\text{leb}(Z)) = \int_0^\infty \mathbb{P}(B_t = 0)dt = 0$

Let $\tau_s = \inf\{t \geq s : B_t = 0\}$ and then $(B_{\tau_s+t} - B_{\tau_s})$ is a new Brownian motion and the Law of the Iterated Logarithm gives that τ_s is not isolated. Let $\hat{Z} = \{\tau_s : s \in \mathbb{Q}\}$. We ask whether $\hat{Z} = Z$. This is no, because we can take the last zero before a point and this will not be of the form τ_s for some s.

Pick $\tau \in Z \setminus \hat{Z}$ and then τ is not isolated from the left. Then pick a sequence $s_k \to \tau$ and then $s_N \leq \tau_{s_N} < \tau$ for all N and so $\tau_{s_k} \to \tau$.

Z is closed as it is the preimage under a continuous map of a closed set. Thus by the deterministic lemma below we have our result. *Q.E.D.*

Lemma 2.25 *Suppose that $A \subset \mathbb{R}$ is closed and has no isolated points. Then A is uncountable.*

Proof Pick $t_0 < t_1$ in A and choose $B_0 = B(t_0, \varepsilon_0)$ and $B_1 = B(t_1, \varepsilon_1)$ disjoint. Then choose points t_{00}, t_{01} inside B_0 and A and t_{10}, t_{11} inside B_1 and A and again choose disjoint balls around these points. Continue this process.

We now have a chain $B_1 \supseteq B_{10} \supseteq ...$ of balls and so there exists a unique point t_a in the infinite intersection. A is closed so $t_a \in Z$. If $a \neq b$ then $t_a \neq t_b$ so the set is uncountable. *Q.E.D.*

2.3 Arcsine Laws for Brownian motion

M is the unique time in $[0,1]$ where it is equal to its supremum, then

$$\mathbb{P}(M \in dt) = \frac{1}{\pi\sqrt{t(1-t)}}dt$$

but why is M symmetric?

The last zero $T = \sup\{t \leq 1 : B_t = 0\}$ then

$$\mathbb{P}(T \in dt) = \frac{1}{\pi\sqrt{t(1-t)}}dt$$

Let $L = \int_0^1 \chi(B_s > 0)ds$, i.e. the time it is above zero. Then

$$\mathbb{P}(L \in dt) = \frac{1}{\pi\sqrt{t(1-t)}}dt$$

These are called arcsine laws due to the following:

$$\int_0^t \frac{1}{\pi\sqrt{t(1-t)}}dt = \frac{1}{\pi}\int_0^{\sin^{-1}(\sqrt{t})} \frac{2\sin\theta\cos\theta}{\sin\theta\cos\theta}d\theta = \frac{2}{\pi}\sin^{-1}(\sqrt{t})$$

For the second law, note

$$\mathbb{P}(T \leq t) = \mathbb{P}(X \text{ does not hit } -B_t \text{ before time } 1-t)$$

and using the following

$$\begin{aligned}
\mathbb{P}(a + X \text{ does not hit the origin by time } 1-t) &= \mathbb{P}(B \text{ does not hit a by time } 1-t) \\
&= 1 - \mathbb{P}(B \text{ does hit a by time } 1-t) \\
&= 1 - 2\mathbb{P}(B_{1-t} > a) \\
&= 1 - 2\mathbb{P}(N(0,1-t) > a) \\
&= 1 - 2\mathbb{P}(N(0,1) > \frac{a}{\sqrt{1-t}}) \\
&= 1 - 2\int_{\frac{a}{\sqrt{1-t}}}^{\infty} \frac{1}{\sqrt{2\pi}}e^{-z^2/2}dz \\
&=: 1 - 2\Phi\left(\frac{a}{\sqrt{1-t}}\right)
\end{aligned}$$

and so

$$\begin{aligned}
\mathbb{P}(T \leq t) &= \mathbb{E}(1 - 2\Phi\left(\frac{a}{\sqrt{1-t}}\right)\Big|_{a=B_t}) \\
&= \mathbb{E}(1 - 2\Phi\left(\frac{\sqrt{t}|N(0,1)|}{\sqrt{1-t}}\right)) \\
&= -\int_{-\infty}^{\infty} 1 - 2\Phi\left(\frac{\sqrt{t}|x|}{\sqrt{1-t}}\right)\frac{1}{\sqrt{2\pi}}e^{-x^2/2}dx
\end{aligned}$$

and if you consider densities you get

$$\mathbb{P}(t \in dt) = -\frac{4}{\sqrt{2\pi}} \int_0^\infty e^{-\frac{tx^2}{1-t}} e^{-x^2/2} \left(\frac{1}{2\sqrt{t}} \frac{x}{\sqrt{1-t}} + \frac{\sqrt{t}x}{(1-t)^{\frac{3}{2}}} \right) dx$$

and this is of the form $\int_0^\infty axe^{-bx^2}dx$ which eventually gives the answer.

Theorem 2.26 (Lévy) *Suppose $S_t = \sup_{s\leq t} B_s$. Then $X_t = S_t - B_t$ is a reflected Brownian motion.*

Then this would give that the alst zero for X is equal to the time of maximal value for B, and so the first two arcsine laws one would expect to be the same.

Corollary 2.27 $T \overset{D}{=} M$

Remark There is almost a random walk analogue. If we do the same thing, then we get a reflected random walk, but it is " sticky " at the origin. We expect the set of sticky times to be small though, because

$$\int_0^t \chi(B_s = 0) ds = 0$$

Proof It order to prove Lévy's theorem, it is left to the reader to check that X is a Markov process with correct transition kernel as given before.

Then one can find $\mathbb{P}(X_{t_1} \in A_1 ... X_{t_N} \in A_N)$. Then

$$\mathbb{P}(X_t \in dy | \mathcal{F}_s^X) = \mathbb{P}(X_t \in dy | \mathcal{F}_s^B) =: I + II|_{a=S_s - B_s}$$

Where I and II are as below. Call $Y_r = X_{s+r} - X_s$ adn this is a new Brownian motion independent of $\mathcal{F}_s^B$. We then have two possibilities. Either I, it attains its maximum after s, or it attains it before. We then have

$$I = \int_a^\infty \mathbb{P}(S_{t-s}^Y \in dz, B_{t-s} \in dz - y) \qquad II = \mathbb{P}(S_{t-s}^Y \leq a, B \in a - dy)$$

Q.E.D.

3 Brownian Martingales

Definition 3.1 *Fix a filtration $(\mathcal{F}_t, t \geq 0)$. A process $(M_t, t \geq 0)$ is called a $(\mathcal{F}_t)$-* **martingale if**

$$\mathbb{E}(M_t | \mathcal{F}_s) = M_s \qquad \forall s \leq t$$

Observe that this requires $\mathbb{E}|M_t| < \infty$ for all $t \geq 0$

Theorem 3.2 (Optional Stopping Theorem) *If $(M_t, t \geq 0)$ is an $\mathcal{F}_t$ martingale and T is a bounded stopping time then*

$$\mathbb{E}(M_T) = \mathbb{E}(M_0)$$

There is a financial interpretation of this. M_t could be considered as a stock price. Then the martingale property is the natural assumption, in discounted money. In other words the best way to predict the future is to consider the present. Then the OST tells you that the expected selling value M_T equals the initial value M_0, you cannot make money from a martingale.

Some books use the OST as the definition of a martingale.

T bounded means that there exists a $K \in \mathbb{R}$ such that $\mathbb{P}(T \leq K) = 1$

Consider $T_a = \inf\{t : B_t = a\}$. This is not bounded. Then $B_{T_a} = a$ so $\mathbb{E}(B_{T_a}) = a$ but $\mathbb{E}(B_0) = 0$ and these are not equal.

Example 3.1 *(B_t) is a martingale because $\mathbb{E}(B_t|\mathcal{F}_s^B) = \mathbb{E}(B_s + B_t - B_s|\mathcal{F}_s^B) = B_s$ Take $T = T_a \wedge T_b$ with $a < 0$ and $b > 0$. This is the minimum of two stopping times and so is a stopping time. It is also bounded. We now use the Optional Stopping theorem.. We have $\mathbb{E}(B_{T\wedge N}) = 0$ and so by DCT we have $\mathbb{E}(B_T) = 0$ since $|B_{T\wedge N}| \leq \max(|a|, |b|)$. Note to always check that this can be done. We then have*

$$0 = \mathbb{E}(B_T) = b\mathbb{P}(T_b < T_a) + a\mathbb{P}(T_a < T_b)$$

and we also have that

$$1 = \mathbb{P}(T_b < T_a) + \mathbb{P}(T_a < T_b)$$

Solving these gives

$$\mathbb{P}(T_b < T_a) = \frac{-a}{b-a} \qquad \mathbb{P}(T_a < T_b) = \frac{b}{b-a}$$

Example 3.2 *$(B_t^2 - t, t \geq 0)$ is a martingale.*

$$\mathbb{E}(B_t^2|\mathcal{F}_s^B) = \mathbb{E}(B_s^2 + 2B_s(B_t - B_s) + (B_t - B_s)^2|\mathcal{F}_s^B) = B_s^2 + t - s$$

and so $\mathbb{E}(B_t^2 - t|\mathcal{F}_s^B) = B_s^2 - s$ as we want for a martingale. The OST says that

$$\mathbb{E}(B_{T\wedge N}^2 - T \wedge N) = 0$$

and using DCT and MCT we get that

$$\mathbb{E}(B_T^2 - T) = 0$$

and this gives that

$$\mathbb{E}(T) = \mathbb{E}(B_T^2) = b^2\mathbb{P}(T_b < T_a) + a^2\mathbb{P}(T_a < T_b) = \frac{-b^2a}{b-a} + \frac{a^2b}{b-a} = -ab$$

Example 3.3 *$(e^{\theta B_t - t\theta^2/2}, t \geq 0)$ is a martingale for all $\theta \in \mathbb{R}$ But how is this a martingale when it seems to be small always. The reason is the process goes to zero but the expectation doesn't. We now check that this is a martingale.*

$$\mathbb{E}(e^{\theta B_t}|\mathcal{F}_s^B) = \mathbb{E}(e^{\theta B_s}e^{\theta(B_t - B_s)}) = e^{\theta B_s}e^{\theta^2/2(t-s)}$$

and then rearranging gives it as a martingale. We then use the DCT to $T_a \wedge N$ and $e^{\theta B_t - t\theta^2/2}$ to get, with OST,

$$\mathbb{E}(e^{\theta B_{T_a\wedge N)} - T_a \wedge N\theta^2/2)} = 1$$

and $T_a \wedge N \to T_a$ and $B_{T_a \wedge N} \to B_{T_a} = a$ and then as $N \to \infty$ we get $\mathbb{E}(e^{-\theta^2/2T_a}) = e^{-\theta a}$. Then if we let $\lambda = \theta^2/2$ we get

$$\mathbb{E}(e^{-\lambda T_a}) = e^{-\sqrt{2\lambda}a}$$

Before we showd that $\mathbb{P}(T_a \leq t) = 2\mathbb{P}(B_t \geq a)$ and we get here

$$\mathbb{E}(e^{-\lambda T_a}) = \int_0^\infty e^{-\lambda t} \frac{a}{t^{3/2}} \frac{1}{\sqrt{2\pi}} e^{-a^2/2t} dt = ...$$

and try this for yourself.

The above examples give a hint that many interesting facts about Brownian motion can be found from the Optional stopping theorem, as applied to some stopping time.

Example 3.4 *$T_a \wedge T_b$ can be done similarly and inversion can be used to give $\mathbb{P}(T_a \wedge T_b \in dt)$.*

Recall that $\mathbb{E}(e^{-\lambda T}) = \phi(\lambda) = 1 + a_1\lambda + a_2\lambda^2 + ... = 1 - \lambda\mathbb{E}(T) + \lambda^2/2\mathbb{E}(T^2) + ...$ by comparing coefficients.

If we define $X_t = B_t - ct$ then $X_t \to -\infty$. Find $\mathbb{P}(T_a < \infty)$. We consider the martingale $e^{\theta B_t - \theta^2/2t} = e^{\theta X_t} e^{(\theta c - \theta^2/2)t}$ and define $T_a = \int\{t : X_t = a\}$. Then OST gives you

$$\mathbb{E}(e^{\theta X_{T_a \wedge N}} e^{(\theta c - \theta^2/2)T_a \wedge N}) = 1$$

Now $e^{\theta X_{T_a \wedge N}} \to \begin{cases} e^{\theta a} & \{T_a < \infty\} \\ 0 & \{T_a = \infty\}, \theta > 0 \end{cases}$ and this is dominated by $e^{\theta a}$.

Now $e^{(\theta c - \theta^2/2)T_a \wedge N} \to \begin{cases} e^{(\theta c - \theta^2/2)T_a} & \{T_a < \infty\} \\ 0 & \{T_a = \infty\}, \theta c - \theta^2/2 < 0 \end{cases}$. If we use $\theta c - \theta^2/2 = \lambda$ then we get

$$\mathbb{E}(e^{-\lambda T_a} \chi_{\{T_a < \infty\}}) = e^{-\theta a}$$

if $\theta > 0$ and $\theta c - \theta^2/2 < 0$. Now choose $\theta \to 2c$ and so $\mathbb{P}(T_a < \infty) = e^{-2ca}$. This makes sense. If a was big, then the chance is small, and if c is large, i.e. the drift is strongly negative, then the chance is also small.

Theorem 3.3 (OST version 1) *Suppose $(M_t : t \geq 0)$ is an $\mathcal{F}_t$ martingale with continuous paths. Suppose also that T is a bounded stopping time and M is bounded. Then*

$$\mathbb{E}(M_T) = \mathbb{E}(M_0)$$

Proof Suppose first that T is discrete, and $T \leq K$. Let $T \in \{t_1, .., t_N\}$ where $t_1 < t_2 < ... < t_N \leq K$. Then

$$\mathbb{E}(M_T) = \sum_{k=1}^N \mathbb{E}(M_{t_k} \chi_{\{T = t_k\}}$$

and also $\mathbb{E}(M_t | \mathcal{F}_s) = M_s$ and so $\mathbb{E}(M_t \chi_A) = \mathbb{E}(M_s \chi_A)$ for $A \in \mathcal{F}_s$. Also $\{T = t_k\} = \{T \leq t_k\} \setminus \{T \leq t_{k-1}\} \in \mathcal{F}_{t_k}$ and so

$$\sum_{k=1}^N \mathbb{E}(M_{t_k} \chi_{\{T = t_k\}} = \mathbb{E}(M_k) = \mathbb{E}(M_0)$$

Suppose now that we have a general stopping time T. Define $T_N = \frac{j}{N}$ if $T \in [\frac{j-1}{N}, \frac{j}{N})$. T_N are discrete stopping times and $T_N \to T$. Also $M_{T_N} \to M_T$ by the continuity of paths and so

$$\mathbb{E}(M_0) = \mathbb{E}(M_{T_N}) \to \mathbb{E}(M_T)$$

using the DCT since $|M_t(\omega)| \le K$ for some K, and for all t. *Q.E.D.*

$B_t = (B_t^{(1)}, ..., B_t^{(d)})$ is a Brownian motion on $\mathbb{R}^d$. Suppose $T = \inf\{t : B_t^{(2)} = 1\}$. Find $\mathbb{P}(B_t^{(1)} \in dx)$. Solve using T and $B^{(1)}$ are independent and you know $\mathbb{P}(T \in dt)$.

Proposition 3.4 (Lots of Brownian martingales) *Suppose that B is a d dimensional Brownian motion, then*

$$f(B_t, t) - \int_0^t \left(\frac{\partial f}{\partial s}(B_s, s) + \frac{1}{2}\Delta f(B_s, s) \right) ds$$

is a martingale if $f \in C^{2,1}$ and $f, \frac{\partial f}{\partial s}, \frac{\partial f}{\partial x_j}, \frac{\partial^2 f}{\partial x_i \partial x_j}$ are of at most exponential growth.

For example, $B_t^2 - t$ with $f(x,t) = x^2$ then $\frac{1}{2}\Delta f = 1$. Also B_t^4 isn't but if we subtract $\int_0^t 6B_s^2 ds$ it is. Also for $e^{\theta B_t - \theta^2/2t}$ we choose $f(x,t) = e^{\theta x - \theta^2/2t}$ and we subtract nothing.

If $B_t^x = x + B_t$ is d dimensional Brownian motion started at x then the above is the same with B_t replaced with B_t^x. We will prove it soon, and the proof uses the fact that the Gaussian density solves $\frac{\partial \phi_t}{\partial t} = \frac{1}{2}\Delta \phi_t$

Of special importance are $f(x,t)$ satisfying $\frac{\partial f}{\partial s} + \frac{1}{2}\Delta f = 0$ or $\Delta f = 0$.

Consider the **Dirichlet problem**. Let $D \subset \mathbb{R}^d$ be open. Then find u such that

$$\begin{cases} \Delta u(x) = 0 & \text{in } D \\ u(x) = f(x) & \text{on } \partial D \end{cases} \tag{3.1}$$

with f a given function.

Theorem 3.5 *Suppose D is a bounded open set of $\mathbb{R}^d$. Suppose $u \in C^2(\mathbb{R}^d)$ and solves*

$$\begin{cases} \Delta u(x) = 0 & \textit{in } D \\ u(x) = f(x) & \textit{on } \partial D \end{cases}$$

then

$$u(x) = \mathbb{E}(f(B_T^x))$$

where $T = \inf\{t : B_t^x \in \partial D\}$.

Proof We can modify u outside of D so that it has at most exponential growth. Then we get, by the above proposition that $u(B_t^x) - \int_0^t \frac{1}{2}\Delta u(B_s^x) ds$ is a martingale. Then the optional stopping theorem, for $T \wedge N$, gives

$$u(x) = \mathbb{E}\left(u(B_{T\wedge N}^x) - \int_0^{T\wedge N} \frac{1}{2}\Delta u(B_s^x) ds \right) = \mathbb{E}(u(B_{T\wedge N}^x))$$

and by the DCT we get $u(x) = \mathbb{E}(u(B_T^x)) = \mathbb{E}(f(B_T^x))$ since $u = f$ on ∂D. *Q.E.D.*

Example 3.5 *Let* $u(x) = \begin{cases} \frac{1}{|x|^{d-2}} & d \geq 3 \\ \log|x| & d = 2 \end{cases}$ *and this satisfies* $\Delta u = 0$ *except at* $x = 0$*. I leave this as a check for the reader, as it has been seen before many times. Let* $D = \{x \in \mathbb{R}^3 : a < |x| < b\}$ *and let* $T_a = \inf\{t : |B_t^x| = a\}$ *and* $T_b = \inf\{t : |B_t^x| = b\}$*. Then let* $T = T_a \wedge T_b$*, and so*

$$\frac{1}{|x|} = \mathbb{E}(\frac{1}{|B_T^x|}) = \frac{1}{a}\mathbb{P}(T_a < T_b) + \frac{1}{b}\mathbb{P}(T_b < T_a)$$

with $1 = \mathbb{P}(T_a < T_b) + \mathbb{P}(T_b < T_a)$ *by the law of the iterated logarithm, and so*

$$\mathbb{P}(T_a < T_b) = \frac{\frac{1}{|x|} - \frac{1}{b}}{\frac{1}{a} - \frac{1}{b}}$$

If we let $b \to \infty$ *and we get* $\mathbb{P}(T_a < T_b) \to \mathbb{P}(T_a < \infty) = \frac{a}{|x|}$.

Corollary 3.6 $|B_t| \to \infty$ *as* $t \to \infty$ *in* $d \geq 3$*.*

Proof If B_t doesnt tend to infinity then there exists K and $t_N \to \infty$ where $|B_{t_N}| \leq K$. Let $T_N = \inf\{t : |B_t| \geq N\}$ then the law of the iterated logarithm says $T_N < \infty$. Then $X_t^{(N)} = B_{T_N+t} - B_{T_N}$ is a Brownian motion and $\mathbb{P}(X^N \text{ hits } B(0,K)) = \left(\frac{K}{N}\right)^{d-2}$ and thus

$$\mathbb{P}(\cap_{N=K}^{\infty}\{X^N \text{ hits } B(0,K)\}) = 0$$

and so

$$\mathbb{P}(\cup_{N=1}^{\infty} \cap_{N=K}^{\infty} \{X^N \text{ hits } B(0,K)\}) = 0$$

Q.E.D.

For $d = 2$ let $b \to \infty$ and then $\mathbb{P}(T_a < \infty) = 1$ and then if $a \to 0$ then $\mathbb{P}(T_{\{0\}} < T_b) = 0$ and if we now let $b \to \infty$ then $\mathbb{P}(T_{\{0\}} < \infty) = 0$

Corollary 3.7 *In* $d = 2$*,* $\mathbb{P}(B$ *ever hits* $\{x\}) = 0$ *for* $x \neq 0$

Corollary 3.8 $\text{Leb}(B_t : t \geq 0) = 0$ *a.s. and so is not a space filling curve.*

Proof $\text{Leb}(B_t : t \geq 0) = \int_{\mathbb{R}^2} \chi_{x \in (B_t, t \geq 0)} dx$ and we consider expectations.

$$\mathbb{E}(\text{Leb}(B_t : t \geq 0)) = \mathbb{E}\left(\int_{\mathbb{R}^2} \chi_{x \in (B_t, t \geq 0)} dx\right) = \int_{\mathbb{R}^2} \mathbb{P}(x \in B_t) dx = 0$$

Q.E.D.

Corollary 3.9 *The range of* $(B_t, t \geq 0)$ *is dense in* $\mathbb{R}^2$

The Poisson problem is the following, for $D \subset \mathbb{R}^d$.

$$\begin{cases} \frac{1}{2}\Delta u = -g & \text{in } D \\ u = 0 & \text{on } \partial D \end{cases} \tag{3.2}$$

and here g represents pumping heat in or out, and u is the steady state temperature.

Theorem 3.10 (Poisson version 1) *Suppose $u \in C^2(\mathbb{R})$ solves equation (3.2) with D bounded. Then*

$$u(x) = \mathbb{E}(\int_0^T g(B_s^x)ds)$$

with $T = \inf\{t : B_t^x \in \partial D\}$

Example 3.6 $u(x) = \frac{R^2 - |x|^2}{d}$ *solves the Poisson problem on a ball of radius R with $g = 1$. Then $u(x) = \mathbb{E}(T)$.*

Example 3.7 *Suppose $D = (0, \infty) \subset \mathbb{R}$ and $u(x) = -x^2/2$. However, $\mathbb{E}(T) \neq u(x)$*

Proof We can modify u away from D to have exponential growth. Then

$$u(B_t^x) - \int_0^t \frac{1}{2}\Delta u(B_s^x)ds$$

is a martingale. The OST then gives

$$u(x) = \mathbb{E}(u(B_{T\wedge N}^x) - \int_0^{T\wedge N} \frac{1}{2}\Delta u(B_s^x)ds) = \mathbb{E}(u(B_{T\wedge N}^x) + \int_0^{T\wedge N} g(B_s^x)ds)$$

and the former tens to 0 by DCT and the latter tends to $\int_0^T g(B_s^x)ds$ by domination with $||g||_\infty T$, and so we need $\mathbb{E}(T) < \infty$ but $T \leq \inf\{t : |B_t^{(1)}| = K\}$ and the right hand side is finite. *Q.E.D.*

PDE people generally look for solutions to these problems that are in $C^2(D) \cap C(\partial D)$. Sadly, one cannot in general extend these to $C^2(\mathbb{R}^d)$ or even $C^2(\bar{D})$.

Example 3.8 $\frac{1}{2}\Delta u = -1$ *on $(0,1)^2$ and $u = 0$ on the boundary. Then u is not $C^2([0,1]^2)$ because of the effects at the corners, as at $(0,0)$ we have $\Delta u = 0$.*

We can improve though, to allow solutions in $C^2(D) \cap C(\partial D)$ but still $u(x)$ solves the Dirichlet or Poisson problem. We do this as follows.

Shrink the domain in by ε by defining $D_\varepsilon = \{x : d(x, D^C) > \varepsilon\}$ and then $D_\varepsilon \to D$ and we can then mollify. Let $T_\varepsilon = \inf\{t : B_t \in \partial D_\varepsilon\}$ and then $B_{T_\varepsilon} \to B_T$. Then one can believe that you can modify u outside of D_ε to be $C^2(\mathbb{R}^d)$ with exponential growth. Then apply version 1 to D_ε to get $u(x) = \mathbb{E}(u(B_{T_\varepsilon})) \to \mathbb{E}(f(B_T))$ as $\varepsilon \to 0$.

By exponential growth, we mean $|g(t,x)| \leq C_0 e^{C_1|x|}$ for all x, t.

Proof (of theorem 3.4) We first show that $\mathbb{E}(|M_t|) < \infty$

$$\mathbb{E}(|f(B_t^x, t)|) \leq C_0\mathbb{E}(e^{C_1|B_t^x|}) \leq C_0\mathbb{E}(e^{C_1 B_t^x} + e^{-C_1 B_t^x}) \leq C_0 e^{C_1 x} e^{C_1^2 t/2} < \infty$$

and similarly $\mathbb{E}(|\int_0^t Lf(B_s^x, s)ds|) < \infty$ WLOG we can take $x = 0$ since we can shift and we still have exponential growth.

$$\begin{aligned}
\mathbb{E}(M_t|\mathcal{F}_s^B) &= \mathbb{E}(f(B_t, t) - \int_0^t Lf(B_r, r)dr|\mathcal{F}_s^B) \\
&= \mathbb{E}\left(f(B_s, s) + (f(B_t, t) - f(B_s, s)) - \int_0^s Lf(B_r, r)dr - \int_s^t Lf(B_r, r)dr|\mathcal{F}_s^B\right) \\
&= M_s + \mathbb{E}(f(B_s + X_{t-s}, t) - f(B_s, s) - \int_0^{t-s} Lf(X_r + B_s, s + r)dr|\mathcal{F}_s^B) \\
&= M_s + \mathbb{E}(f(Z + X_{t-s}, t) - f(B_s, s) - \int_0^{t-s} Lf(X_r + Z, s + r)dr|\mathcal{F}_s^B)|_{Z=B_s}
\end{aligned}$$

where $X_r = B_{s+r} - B_s$. The result now follows from $\mathbb{E}(g(X_t,t) - g(0,0) - \int_0^t Lg(X_r,r)dr) = 0$ where we have used $g(y,t) = f(z+y,t+s)$. Alternatively this is

$$\mathbb{E}(g(X_t,t) - g(0,0) = \int_0^t Lg(X_r,r)dr)$$

which is the same as

$$\frac{d}{dt}\mathbb{E}(g(X_t,t)) = \mathbb{E}(Lg(X_t,t))$$

Then

$$\begin{aligned}
\frac{d}{dt}\mathbb{E}(g(X_t,t)) &= \frac{d}{dt}\int_{\mathbb{R}^d} g(x,t)\frac{1}{(2\pi t)^{d/2}}e^{-|x|^2/2t}dx \\
&= \int_{\mathbb{R}^d}\left(\frac{\partial g}{\partial t}(x,t)\phi(x,t) + \frac{\partial \phi}{\partial t}g\right)dx \\
&= \int_{\mathbb{R}^d}\left(\frac{\partial g}{\partial t}(x,t)\phi(x,t) + \frac{1}{2}(\Delta\phi)g\right)dx \\
&= \int_{\mathbb{R}^d}\left(\frac{\partial g}{\partial t}(x,t)\phi(x,t) + \frac{1}{2}(\Delta g)\phi\right)dx \\
&= \mathbb{E}(Lg(X_t,t))
\end{aligned}$$

as required. *Q.E.D.*

Theorem 3.11 (Heat Equation) *If $u \in C^{1,2}([0,\infty)\times\mathbb{R}^d)$ is of exponential growth and solves*

$$\begin{cases} \frac{\partial u}{\partial t} = \frac{1}{2}\Delta u & t>0, x\in\mathbb{R}^d \\ u(0,x) = f(x) & x\in\mathbb{R}^d \end{cases}$$

then

$$u(x,t) = \mathbb{E}(f(B_t^x)) = \int_{\mathbb{R}^d} f(y)\frac{1}{(2\pi t)^{d/2}}e^{-|x-y|^2/2t}dx$$

Theorem 3.12 (Heat Equation on a region) *Suppose $D \subset \mathbb{R}^d$ is bounded, and $u \in C^{1,2}([0,\infty)\times D)\cap C([0,\infty)\times\bar{D})$ solves*

$$\begin{cases} \frac{\partial u}{\partial t} = \frac{1}{2}\Delta u & t>0, x\in D \\ u(0,x) = f(x) & x\in D \\ u(t,x) = g(x) & x\in\partial D, t>0 \end{cases}$$

and let $T = \inf\{t : B_t^x \in \partial D\}$ and then

$$u(x,t) = \mathbb{E}\left(g(B_T^x)\chi_{\{T\le t\}} + f(B_t^x)\chi_{\{T>t\}}\right)$$

We prove both theorems simultaneously. **Proof** Fix $t>0$ and consider the map

$$s \mapsto u(B_s^x, t-s) - \int_0^s \left(-\frac{\partial u}{\partial r} + \frac{1}{2}\Delta u\right)(B_r^x, r)dr$$

in other words we run backwards in time. Then in $\mathbb{R}^d$ $u(B_s^x, t-s)$ is a martingale so using the OST we get

$$u(x,t) = \mathbb{E}(u(B_t^x,0)) = \mathbb{E}(f(B_t^x))$$

as required.

On D we stop at $T \wedge t$ which is a bounded stopping time. Then $u(B_s^x, s-t)$ is a martingale so

$$u(x,t) = \mathbb{E}(u(B^x_{T\wedge t}, t - T\wedge t)) = \mathbb{E}(\chi_{\{T\le t\}} g(B^x_T) + \chi_{\{T>t\}} f(B^x_t))$$

as required. *Q.E.D.*

Example 3.9 *Brownian motion staying in a tube. Suppose we write $u(x,t) = \mathbb{P}(|B_t| < 1$ for all $s \le t)$. Then $u(x,t)$ solves the equation*

$$\begin{cases} \frac{\partial u}{\partial t} = \frac{1}{2}\frac{\partial^2 u}{\partial x^2} & x \in (-1,1), t \in (0,t) \\ u(0,x) = 1 \\ u(t,x) = 0 & x = \pm 1 \end{cases}$$

but note that $u \notin C([0,t]\times[-1,1])$ and the solution is

$$u(x,t) = \sum_{k=1}^{\infty} a_k \cos\left(\frac{(2k-1)\pi x}{2}\right) e^{-\frac{(2k-1)^2\pi^2 t}{2}}$$

Inspired by this, we consider the first term $\cos\frac{\pi x}{2} e^{-\pi^2 t/8}$ and then this solves the above equation with initial data $u(0,x) = \cos\frac{\pi x}{2}$ and then

$$u(t,0) = e^{-\pi^2 t/8} = \mathbb{E}\left(\cos\frac{B_t^x \pi}{2}\chi_{\{T>t\}}\right) \le \mathbb{P}(\{T>t\})$$

and so $\mathbb{P}(T>t) \ge e^{-\pi^2 t/8}$, and this is correct asymptotically.

Example 3.10 *Find $\mathbb{P}(B_s^{x_1}, ..., B_s^{x_d}$ do not collide by time $t)$. Let $B_t^x = (B_s^{x_1}, ..., B_s^{x_d})$ be a d-dimensional Brownian motion. Let $V_d = \{x \in \mathbb{R}^d : x_1 < x_2 < ... < x_d\}$ which is called a cell, and then $\partial V_d = \{x \in \mathbb{R}^d : x_i = x_{i+1}$ for some $i\}$. This solves*

$$\begin{cases} \frac{\partial u}{\partial t} = \frac{1}{2}\Delta u & x \in V_d, t > 0 \\ u(0,x) = f(x) & x \in V_d \\ u(t,x) = 0 & x = \partial V_d \end{cases}$$

and this has a famous solution

$$u(x,t) = \int_{V_d} f(y_1,\dots,y_d) \det \begin{vmatrix} \phi_t(x_1-y_1) & \dots & \phi_t(x_1-y_d) \\ \vdots & & \vdots \\ \phi_t(x_d-y_1) & \dots & \phi_t(x_d-y_d) \end{vmatrix} dy_1 \dots dy_d$$

where $\phi_t(z) = \frac{1}{\sqrt{2\pi t}} e^{-z^2/2t}$

Definition 3.13 $\phi : \mathbb{R} \to \mathbb{R}$ *is* **convex** *means $\phi(x) = \sup\{L(x) | L$ is linear $, L \le \phi\}$*

Note that if $\phi \in C^2(\mathbb{R})$ then ϕ is convex if and only if $\phi'' \ge 0$.

Lemma 3.14 (Jensen's Inequality) *If $\mathbb{E}(|X|) < \infty$ and $\phi : \mathbb{R} \to \mathbb{R}$ is convex then*

$$\mathbb{E}(\phi(X)) \ge \phi(\mathbb{E}(X))$$

$$\mathbb{E}(\phi(X)|\mathcal{F}) \ge \phi(\mathbb{E}(X|\mathcal{F}))\ a.s.$$

Proof

$$\begin{aligned}\mathbb{E}(\phi(X)|\mathcal{F}) &\geq \mathbb{E}(L(X)|\mathcal{F})\\ &= \mathbb{E}(aX+b|\mathcal{F})\\ &= a\mathbb{E}(X|\mathcal{F})+b\\ &= L(\mathbb{E}(X|\mathcal{F}))\text{ a.s.}\end{aligned}$$

and then taking a supremum over $L \leq \phi$ gives the result. Note we need to ensure somehow that we only need countably many Ls. *Q.E.D.*

Corollary 3.15 *Suppose (M_t) a martingale and ϕ is convex, then*

$$\mathbb{E}(\phi(M_t)|\mathcal{F}_s) \geq \phi(\mathbb{E}(M_t|\mathcal{F}_s)) \geq \phi(M_s)$$

Lemma 3.16 (OST for $\phi(M_t)$) *Suppose $(M_t, t \geq 0)$ is a martingale with continuous paths and $\phi \geq 0$ is a convex function. Suppose $T \leq K$ is a bounded stopping time. Then*

$$\mathbb{E}(\phi(M_T)) \leq \mathbb{E}(\phi(M_K))$$

Proof First assume that T is discrete, so $T \in \{t_1, ..., t_M\}$ with $t_1 < t_2 < ... < t_M$. Then

$$\mathbb{E}(\phi(M_T)) = \sum_{k=1}^{M} \mathbb{E}(\phi(M_{t_k})\chi_{\{T=t_k\}}) \leq \sum_{k=1}^{M} \mathbb{E}(\phi(M_K)\chi_{\{T=t_k\}}) = \mathbb{E}(\phi(M_K))$$

Now for general T, find discrete stopping times $T_N \leq K$ so that $T_N \to T$ and then we have $\mathbb{E}(\phi(M_{T_N})) \leq \mathbb{E}(\phi(M_K))$ and using Fatou's lemma we get

$$\mathbb{E}(\phi(M_T)) \leq \mathbb{E}(\phi(M_K))$$

Q.E.D.

We now prove a more general OST, one without boundedness of the martingale

Theorem 3.17 *Suppose $(M_t, t \geq 0)$ is a continuous martingale, and $T \leq K$ is a bounded stopping time. Then*

$$\mathbb{E}(M_T) = \mathbb{E}(M_0)$$

Proof For discrete T this works as above and is fine.

We now consider $M_t \geq 0$ and assume we have discrete stopping times $T_N \to T$ and we have $\mathbb{E}(M_{T_N}) = \mathbb{E}(M_0)$. We then write $x = x \wedge L + (x-L)^+$ and we have

$$\mathbb{E}(M_0) = \mathbb{E}(M_{T_N}) = \mathbb{E}(M_{T_N} \wedge L) + \mathbb{E}((M_{T_N} - L)^+)$$

and now $\mathbb{E}(M_{T_N} \wedge L) \to \mathbb{E}(M_T \wedge L)$ by DCT and $\mathbb{E}(M_T \wedge L) = \mathbb{E}(M_T) - \mathbb{E}((M_T - L)^+)$

Fix $\varepsilon > 0$ Choose L large so that

$$\mathbb{E}((M_{T_N} - L)^+) \leq \mathbb{E}((M_K - L)^+) < \varepsilon$$

and then take N large so that

$$|\mathbb{E}(M_{T_N} \wedge L) - \mathbb{E}(M_T \wedge L)| \leq \varepsilon$$

Finally truncate a general martingale at $+L$ and $-L$. This is a bit messier though. *Q.E.D.*

We now have some final remarks on the Dirichlet problem.

Can we define $u(x) = \mathbb{E}(f(B_T^x))$ and check that it solves the Dirichlet problem. Suppose that $D = \{x \in \mathbb{R}^2 : 0 < |x| < 1\}$ the punctured disc. If we suppose that $u = 1$ on the set $\{x : |x| = 1\}$ and 0 at 0 then the solution given by the Brownian motion formula is $u \equiv 1$ but this doesn't solve the boundary conditions.

However, if u is of that form, then $u \in C^\infty(D)$ and $\Delta u = 0$ on D always.

We call a point $y \in \partial D$ **regular** if $u(x) \to f(y)$ as $x \to y$ in D. Thus if all point of ∂D are regular then u is a solution. We thus need a sufficient condition fo $y \in \partial D$ to be regular.

To show the first point, we observe that u being harmonic is the same as u satisfying the ball averaging property, or the sphere averaging property, namely if $B(x, \varepsilon) \subset D$ then

$$u(x) = \frac{1}{|B(x,\varepsilon)|} \int_{B(x,\varepsilon)} u(y)dy$$

or

$$u(x) = \frac{1}{SA_\varepsilon} \int_{\partial B(x,\varepsilon)} u(y)dS$$

Sphere averaging for our formula is almost obvious. If we let $S_\varepsilon = \inf\{t : B_t^x \in \partial B(x, \varepsilon)\}$, $X_t = B_{S_\varepsilon + t} - B_{S_\varepsilon}$ and $T' = \inf\{t : X_t + B_{S_\varepsilon}^x \in \partial D\}$ then

$$\begin{aligned} u(x) &= \mathbb{E}(f(B_T^x)) \\ &= \mathbb{E}(f(B_{S_\varepsilon}^x + B_T^x - B_{S_\varepsilon}^x)) \\ &= \mathbb{E}(f(B_{S_\varepsilon}^x + X_{T'})) \\ &= \mathbb{E}(\mathbb{E}(f(B_{S_\varepsilon}^x + X_{T'})|\mathcal{F}_{S_\varepsilon}^B)) \\ &= \mathbb{E}(f(z + X_{T'})|_{z=B_{S_\varepsilon}^x}) \\ &= \mathbb{E}(u(z)|_{z=B_{S_\varepsilon}^x}) \\ &= \mathbb{E}(u(B_{S_\varepsilon}^x)) \end{aligned}$$

which is sphere averaging,

To do the part on regular points, the following is an equivalent definition of regular. $\mathbb{P}(T_x > \varepsilon) \to 0$ as $x \to y$ where $T_x = \inf\{t : B_t^x \in \partial D\}$ for all $\varepsilon > 0$.

This is equivalent to $\mathbb{P}(\sigma^y = 0) = 1$ where $\sigma^y = \inf\{t > 0 : B_t^y \in D^C\}$. The 0-1 law says that this is either 0 or 1. Remember Lebesgue's thorn. If y is on a thin enough spike then $\mathbb{P}(\sigma^y > 0) = 1$ and then you cannot solve the Dirichlet problem.

A sufficient condition for y to be a regular point is the cone condition, namely if there exists a cone in D^C with vertex at $y \in \partial D$ then y is regular. This is because

$$\mathbb{P}(\sigma^y \le \varepsilon) \ge \mathbb{P}(B_\varepsilon \in \text{ cone }) = p(\alpha) > 0$$

where α is the solid angle of the cone. Then letting $\varepsilon \to 0$ gives

$$\mathbb{P}(\sigma^y = 0) \ge p(\alpha) > 0$$

and so the 0-1 law gives $\mathbb{P}(\sigma^y = 0) = 1$

Book by Richard Bass on this sort of stuff

4 Donsker's theorem

The idea is to show random walks converge to Brownian motion. Throughout this chapter we have the following. $Z_1, Z_2, \ldots$ are IID random variables with $\mathbb{E}(Z_i) = 0$ and $\mathbb{E}(Z_i^2) = 1$ and we define $S_N = \sum_{k=1}^{N} Z_k$ and this is the position at time N. We can interpolate the S_N to get S_t and this is given by

$$S_t = (N + 1 - t)S_N + (t - N)S_{N+1} \qquad\qquad t \in [N, N + 1]$$

We also define $X_t^{(N)} = \frac{S_{Nt}}{\sqrt{N}}$ but we cannot hope that $X_t^{(N)} \to B_t$. However, we do have that $X_t^{(N)} \stackrel{D}{\to} B_t$. This is the aim of this section.

We know that

$$\frac{S_N}{\sqrt{N}} \to N(0,1)$$

by the central limit theorem, and so

$$X_t^{(N)} = \frac{S_{Nt}}{\sqrt{N}} = \frac{S_{\lfloor Nt \rfloor}}{\sqrt{N}} + \text{ error } = \frac{S_{\lfloor Nt \rfloor}}{\sqrt{\lfloor Nt \rfloor}} \frac{\sqrt{\lfloor Nt \rfloor}}{\sqrt{N}} + \text{ error } \to N(0,1)\sqrt{t}$$

we hope. Thus it has the same distribution as Brownian motion.

Similarly $(X_{t_1}^{(N)}, \ldots, X_{t_k}^{(N)}) \to (B_{t_1}, \ldots, B_{t_k})$ but does

$$\max_{t\in[0,1]} X_t^{(N)} \stackrel{D}{\to} \max_{t\in[0,1]} B_t \tag{4.1}$$

$$\int_0^1 Xt^{(N)} dt \stackrel{D}{\to} \int_0^1 B_t dt \sim N(0, 1/3) \tag{4.2}$$

$$\int_0^1 \chi_{\{X_s^{(N)}>0\}} ds \stackrel{D}{\to} \int_0^1 \chi_{\{B_s>0\}} ds \tag{4.3}$$

(4.2) can be rewritten as $\frac{\sum_1^N S_K}{N^{3/2}} \stackrel{D}{\to}$ and (4.3) can be rewritten almost as $\frac{\text{number of times when } S_k>0}{N}$

The plan is to think of $(X_t^{(N)}, t \in [0, 1]) =: X^{(N)}$ and $(B_t, t \in [0, 1]) = B$ as random variables in $C[0, 1]$. We thus need to show that

$$X^{(N)} \stackrel{D}{\to} B$$

on $C[0, 1]$. This is a big improvement of the Central limit theorem. We also show that

$$F(X^N) \stackrel{D}{\to} F(B)$$

and we observe that

$$X^{(N)} \stackrel{D}{\to} B \implies F(X^N) \stackrel{D}{\to} F(B)$$

if F is continuous. In the above questions, the maximum and integral are continuous, but the last is not, as functions $C[0, 1] \to \mathbb{R}$.

Definition 4.1 *(E, d) a metric space. $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (E, d)$ measurable ($X^{-1}(B) \in \mathcal{F}$ for B Borel in E)is called an E-valued random variable.*

We use this with $E = C[0,1]$ with

$$d(f,g) = \sup_{t\in[0,1]} |f(t) - g(t)|$$

and an open ball is $B(f,\varepsilon) = \{g : |g(t) - f(t)| < \varepsilon\}$ and the Borel sets are generated by open balls.

Lemma 4.2 $\mathcal{B}(C[0,1]) = \sigma(\mathcal{F}_0)$ *where*

$$\mathcal{F}_0 = \{f : f(t_1) \in O_1, ..., f(t_N) \in O_N\}$$

for $t_1, ..., t_N$ *and* $N \geq 0$ *and* O_i *open sets.*

Proof It is easy to check that $\mathcal{F}_0$ is a π-system.

We check that $\sigma(\mathcal{F}_0) \subset \mathcal{B}(C[0,1])$. But if $f \in \{f(t_1) \in O_1, ..., f(t_N) \in O_N\}$ then since O_1 is open there is an ε_1 such that if $|f - g| < \varepsilon_1$ then $g(t_1) \in O_1$, and so on, then take $\varepsilon = \min\{\varepsilon_1, ..., \varepsilon_N\}$ and we get $B(f,\varepsilon) \subset \{f(t_1) \in O_1, ..., f(t_N) \in O_N\}$ and so it is open.

We check that $\mathcal{B}(C[0,1]) \subset \sigma(\mathcal{F}_0)$. It is enough to check that $B(f,\varepsilon) \in \sigma(\mathcal{F}_0)$. Now

$$\begin{aligned}\overline{B(f,\varepsilon)} &= \{: |g(t) - f(t)| \leq \varepsilon \ \ \forall t \in [0,1]\} \\ &= \bigcap_{t\in[0,1]\cap\mathbb{Q}} \{: |g(t) - f(t)| \leq \varepsilon\} \\ &= \bigcap_{t\in[0,1]\cap\mathbb{Q}} \bigcap_{N\geq 0} \{: |g(t) - f(t)| < \varepsilon + \frac{1}{N}\} \in \mathcal{F}_0\end{aligned}$$

and also $B(f,\varepsilon) = \bigcup_{N\geq 1} \overline{B(f,\varepsilon - 1/N)}$ *Q.E.D.*

An example of this is $B = (B_t, t \geq 0)$ is a $C[0,1]$ valued variable. One observes that $\{\omega\ B_{t_1} \in O_1, ..., B_{t_N} \in O_N \in \mathcal{F}_0$ Also $X_t^{(N)}$ are random variables in $C[0,1]$. We can show the latter using composition of measurable functions, $\Omega \to \mathbb{R}^N \to C[0,1]$ given by $\omega \mapsto (Z_1(\omega), ..., Z_N(\omega)) \mapsto X_t^{(N)}(\omega)$ and the former is measurable and the latter is continuous.

Definition 4.3 $X^{(N)}, X$ *are* (E,d) *valued variables then* $X_t^{(N)} \stackrel{D}{\to} X$ *means*

$$\mathbb{E}(F(X^{(N)})) \to \mathbb{E}(F(X))$$

for $F : E \to \mathbb{R}$ *bounded and continuous.*

Theorem 4.4 (Donsker)

$$X^{(N)} \stackrel{D}{\to} B$$

on $C[0,1]$

Theorem 4.5 (Continuous Mapping) *If* $X^{(N)} \stackrel{D}{\to} X$ *on* (E,d) *and* $G : (E,d) \to (\tilde{E},\tilde{d})$ *which is continuous then*

$$G(X^{(N)}) \stackrel{D}{\to} G(X)$$

on $(\tilde{E},\tilde{d})$

Proof Take $F : \tilde{E} \to \mathbb{R}$ bounded and continuous. Then

$$\mathbb{E}(F(G(X^{(N)}))) \to \mathbb{E}(F(G(X)))$$

since $F \circ G$ is bounded and continuous. *Q.E.D.*

Corollary 4.6

$$\max_{t\in[0,1]} X_t^{(N)} \to \max_{t\in[0,1]} B_t$$

Proof $F(t) = \max f(t)$ is a continuous map. *Q.E.D.*

Consider $\int_0^1 \chi_{(X_t^{(N)}>0)} dt \to \int_0^1 \chi_{(B_t>0)} dt$ is not continuous. Also recall that if $X^{(N)} \overset{D}{\to} X$ then this does not imply that $\mathbb{P}(X_N \in (a,b)) \to \mathbb{P}(X \in (a,b))$, for example consider $X_N = 1/N$ and $(a,b) = (0,1)$.

Consider the simple random walk S_N. Then $\frac{S_N^2}{N} \overset{D}{\to} N(0,1)$ but $\mathbb{P}(\frac{S_N^2}{N} \in \mathbb{Q}) = 1 \neq 0 = \mathbb{P}(N(0,1) \in \mathbb{Q})$. However, $\mathbb{P}(\frac{S_N^2}{N} \in (a,b)) \to \mathbb{P}(N(0,1) \in (a,b))$

Theorem 4.7 (Extended Continuous Mapping) *If $X^{(N)} \overset{D}{\to} X$ on (E,d) and $F : (E,d) \to \mathbb{R}$ is measurable and* $\text{Disc}(F) = \{f : F \textit{ is discontinuous at} f\}$ *is such that* $\mathbb{P}(X \in \text{Disc}(F)) = 0$ *then*

$$F(X^{(N)}) \overset{D}{\to} F(X)$$

on $(\tilde{E}, \tilde{d})$

This is a big non trivial improvement.

For the indicator problem, we guess the discontinuity set to be F is discontinuous at f if $\int_0^1 \chi_{(f(t)=0)} dt > 0$, but for a Brownian path, $\int_0^1 \chi_{(B_t=0)} dt = 0$ and so we have the result we want.

We now try to prove the approximation. We remind ourselves that $X_t^{(N)} = \frac{S_{Nt}}{\sqrt{N}}$ with S_N the simple symmetric random walk and we take $(Z_1, Z_2, ...)$ IID with mean 0 and variance 1. The aim is to show that $X^{(N)} \overset{D}{\to} B$ on $C([0,1])$ with the supremum norm.

We embed simple random walks as follows. We define $T_1 = \inf\{t : |B_t| = 1\}$ and then inductively define $T_{N+1} = \inf\{t \geq T_N : |B_t - B_{T_N}| = 1\}$ and then linearly interpolate. We have

$$(B_{T_1}, B_{T_2}, ...) \overset{D}{=} (S_1, S_2, ...)$$

We are close to a proof. We know that $\mathbb{E}(T_1) = 1$ and the strong Markov property at T_N implies that $\mathbb{E}(T_{N+1} - T_N) = 1$ and $T_{N+1} - T_N$ is independent of $T_1, T_2, ...$ and so $T_N \approx N \pm O(\sqrt{N})$ and if we take $B_t^{(N)} = \frac{B_{Nt}}{\sqrt{N}}$ then we expect $X^{(N)}$ is close to this.

Lemma 4.8 (Skorokhod) *Take Z with $\mathbb{E}(Z) = 0$. Then there exists a stopping time $T < \infty$ so that $B_T \overset{D}{=} Z$ and $\mathbb{E}(T) = \mathbb{E}(Z^2)$.*

Financial mathematicians love this lemma. There are at least 14 different ways to prove this.

If $B_t^2 - t$ is a martingale and so $\mathbb{E}(B_{T\wedge N}^2 - T \wedge N) = 0$ or $\mathbb{E}(T \wedge N) = \mathbb{E}(B_{T\wedge N}^2)$ and this, by Fatou, gives $\mathbb{E}(T) \geq |E(B_T^2)$.

If we take Z to be independent of B we choose $T = \inf\{t : B_t = Z\}$ and so $B_T = Z$ but $\mathbb{E}(T) = \mathbb{E}(\mathbb{E}(T|\sigma(Z))) = \mathbb{E}(\mathbb{E}(T_a)|_{a=Z}) = \infty$. This is a bit of a silly example.

Proof We first suppose $Z \in \{a, b\}$. Then define T by

$$\begin{cases} \mathbb{P}(B_T = a) = \frac{b}{b-a} \\ \mathbb{P}(B_T = b) = \frac{-a}{b-a} \end{cases}$$

and then $\mathbb{E}(T) = -ab$.

We now take a general Z. Choose random $\alpha < 0, \beta > 0$ independent of B, and use $T = T_{\alpha,\beta}$. We need to find the distribution of α and β. To this end we need to choose

$$\nu(da, db) = \mathbb{P}(\alpha \in da, \beta \in db)$$

and we have the target distribution of

$$\begin{cases} \mu_+(dz) = \mathbb{P}(Z \in dz) & Z \geq 0 \\ \mu_-(dz) = \mathbb{P}(Z \in dz) & Z < 0 \end{cases}$$

If $Z \geq 0$ we need

$$\begin{aligned} \mu_+(dz) &= (B_{T_{\alpha,\beta}} \in dz) \\ &= \mathbb{E}(\mathbb{P}(B_{T_{\alpha,\beta}} \in dz | \sigma(\alpha, \beta))) \\ &= \mathbb{E}\left(\frac{-\alpha}{\beta - \alpha}\chi(\beta \in dz)\right) \\ &= \iint \frac{-a}{b-a}\chi(b \in dz)\nu(da, db) \\ &= \int \frac{-a}{z-a}\nu(da, dz) \end{aligned}$$

and so we choose $\nu(da, dz) = (z - a)\mu_+(dz)\pi(da)$ where $\int -a\pi(da) = 1$ and so we have matched μ_+.

For $Z < 0$ we have

$$\begin{aligned} \mu_-(dz) &= \mathbb{P}(B_{T_{\alpha,\beta}} \in dz) \\ &= \mathbb{E}\left(\frac{\beta}{\beta - \alpha}\chi(\alpha \in dz)\right) \\ &= \iint \frac{b}{b-a}\chi(a \in dz)\nu(da, db) \\ &= \int b\mu_+(db)\pi(dz) \end{aligned}$$

and so we choose

$$\pi(dz) = \frac{\mu_-(dz)}{\int b\mu_+(db)}$$

and so α, β are distributed as

$$\nu(da, db) = \frac{(b-a)\mu_+(db)\mu_-(da)}{\int x\mu_+(dx)}$$

We thus have four things to check:

1. $\mathbb{P}(B_{T_{\alpha,\beta}} \in dz) = \mathbb{P}(Z \in dz)$

2. $\int -a\pi(da) = 1$
3. $\mathbb{E}(T_{\alpha,\beta}) = \mathbb{E}(Z^2)$
4. $\iint \nu(da, db) = 1$

These are all true, but we only check 2 and 3. Observe that 1 is by construction. For 2, we want to show that

$$\int -a\mu_-(da) = \int a\mu_+(da)$$

but we have

$$0 = \mathbb{E}(Z) = \int a\mu_+(da) + \int a\mu_-(da)$$

which is what we want.

For 3, observe that

$$\begin{aligned}
\mathbb{E}(T_{\alpha,\beta}) &= \mathbb{E}(-\alpha\beta) \\
&= \iint -ab\nu(da, db) \\
&= \iint -ab(b-a)\frac{\mu_+(db)\mu_-(da)}{\int x\mu_+(dx)} \\
&= \frac{-\int xd\mu_- \int x^2 d\mu_+ + \int x^2 d\mu_- \int xd\mu_+}{\int xd\mu_+} \\
&= \int x^2 d\mu_+ \underset{\int}{\,} x^2 d\mu_- \\
&= \mathbb{E}(Z^2)
\end{aligned}$$

Q.E.D.

We use a Skorokhod trick:

$$\begin{cases} B_{T_{\alpha,\beta}} \overset{D}{=} Z \\ \mathbb{E}(T_{\alpha,\beta}) = 1 \end{cases}$$

and take IID copies (α_1, β_1), (α_2, β_2) independent of B. Then $T_1 = \inf\{t : B_t \in \{\alpha_1, \beta_1\}\}$,...,$T_{N+1} = \{t \geq T_N : B_t - B_{T_N} \in \{\alpha_{N+1}, \beta_{N+1}\}\}$ and define $(S_1, S_2, ...) = (B_{T_1}, B_{T_2}, ...)$ and then it has the random walk distribution that we want, i.e.

$$X_t^{(N)} = \frac{S_{Nt}}{\sqrt{N}} \qquad B^{(N)} = \frac{B_{Nt}}{\sqrt{N}}$$

The key estimate is

$$\mathbb{P}\left(||X^{(N)} - B^{(N)}||_\infty > \varepsilon\right) \to 0$$

as $N \to \infty$. Assume this, and then fix $F : C[0,1] \to \mathbb{R}$ that is bounded and uniformly continuous. We get

$$\begin{aligned}
|\mathbb{E}(F(X^{(N)})) - \mathbb{E}(F(B^{(N)}))| &\leq |\mathbb{E}(F(X^{(N)}) - F(B^{(N)}))| \\
&\leq \mathbb{E}|F(X^{(N)} - F(B^{(N)}| \\
&\leq 2||F||_\infty \mathbb{P}(\Omega_{N,\varepsilon}) + \mathbb{E}(|F(X^{(N)} - F(B^{(N)}|\chi_{\Omega^C_{N,\varepsilon}})
\end{aligned}$$

Fix $\eta > 0$ and use uniform continuity to choose ε so that the second term is less than or equal to $\eta/2$.

Then choose N large to make the first term less than or equal to $\eta/2$.

We soon check that it is enough to only use uniformly continuous functions.

We now show the key estimate.

$$X^{(N)}_{K/N} := \frac{S_K}{\sqrt{N}} := \frac{B_{T_K}}{\sqrt{N}} = B^{(N)}_{T_K/N} \approx \frac{B^{(N)}_K}{N}$$

where the approximation is the first gap we need to plug. This is the difference of the two at the endpoints. The second gap is the difference at other parts of the paths.

We plug the first gap. Take $T_1, T_2 - T_1, T_3 - T_2$ and so on IID with mean 1. Then

$$\frac{T_N}{N} \overset{a.s.}{\to} 1$$

by the strong law of large numbers. Then

$$\max_{i=1,\dots,N} |\frac{T_k}{N} - \frac{k}{N}| \to 0$$

almost surely, from the above (analysis 1).

Let $\Omega_{N,\delta} = \{\max_{i=1,\dots,N} |\frac{T_k}{N} - \frac{k}{N}| \geq \delta\}$ and then

$$\mathbb{P}(\Omega_{N,\delta}) \to 0$$

as $N \to \infty$.

We now plug the second gap. Suppose $||X^{(N)}_t - B^{(N)}_t||_\infty > \varepsilon$. Then there exists a $t \in [0,1]$ such that

$$|X^{(N)}_t - B^{(N)}_t| \geq \epsilon$$

and suppose that $t \in [K/n, (K+1)/N]$. Then either

$$|B^{(N)}_t - B^{(N)}_{K/N}| \geq \varepsilon$$

or

$$|B^{(N)}_t - B^{(N)}_{(K+1)/N}| \geq \varepsilon$$

and so

$$\begin{aligned} \mathbb{P}(||X^{(N)}_t - B^{(N)}_t||_\infty > \varepsilon) &\leq \mathbb{P}(\Omega_{N,\delta}) + \mathbb{P}(|B^{(N)}_s - B^{(N)}_t| \geq \varepsilon \text{ for } |s-t| \leq \delta + 1/N) \\ &\leq \mathbb{P}(\Omega_{N,\delta}) + \mathbb{P}(|B^{(N)}_s - B^{(N)}_t| \geq \varepsilon \text{ for } |s-t| \leq 2\delta) \end{aligned}$$

and then choose δ small and N large.

This concludes the proof of Donskers theorem, modulo some other minor tidy ups.

Corollary 4.9

$$\int_0^1 X^{(N)}_t dt \overset{D}{\to} \int_0^1 B_t dt$$

We apply $F(f) = \int_0^1 f$. The former isn't really very useful, so we rewrite it as follows:

$$\int_0^1 X^{(N)}_t dt = \sum_0^{N-1} \frac{1}{2N}\left(\frac{S_K}{\sqrt{N}} + \frac{S_{K+1}}{\sqrt{N}}\right) = \frac{1}{N^{3/2}} \sum_0^{N-1} S_K - \frac{S_N}{N^{3/2}}$$

and so we suspect that

$$\frac{1}{N^{3/2}} \sum_0^{N-1} S_K \to N(0, 1/3)$$

We need a tidy up lemma:

Lemma 4.10 *Suppose that* $X_N \xrightarrow{D} X$ *and* $Y_N \xrightarrow{D} 0$. *Then* $X_N + Y_N \xrightarrow{D} X$.

This is not true if Y_N converges to a non zero limit. For example consider $Y_N = \begin{cases} Z & N \text{ even} \\ -Z & N \text{ odd} \end{cases}$.

Proof We consider characteristic functions

$$|\mathbb{E}(e^{i\theta(X_N+Y_N)} - \mathbb{E}(e^{i\theta X})| \leq |\mathbb{E}(e^{i\theta X_N} - e^{i\theta X})| + |\mathbb{E}(e^{i\theta(X_N+Y_N)} - e^{i\theta X_N})|$$

The former tends to zero as $X_N \to X$ and the latter is as follows

$$\begin{aligned} |\mathbb{E}(e^{i\theta(X_N+Y_N)} - e^{i\theta X_N})| &\leq \mathbb{E}(|e^{i\theta Y_N} - 1|) \\ &\leq \sqrt{\mathbb{E}(|e^{i\theta Y_N} - 1|^2)} \\ &\leq \sqrt{\mathbb{E}(1 - e^{i\theta Y_N} - e^{-i\theta Y_N} + 1)} \\ &\to 0 \end{aligned}$$

Q.E.D.

Now back to the problem in hand. We have $F(f) = \int_0^1 f$ and we have $X^{(N)} \xrightarrow{D} B$ and then

$$F(X^{(N)} = \frac{1}{N^{3/2}} \sum_0^{N-1} S_K - \frac{S_N}{N^{3/2}}$$

Then

$$\mathbb{E}\left|\frac{S_N}{N^{3/2}}\right| \leq \sqrt{\mathbb{E}\left|\frac{S_N}{N^{3/2}}\right|^2} = \frac{N}{N^3} \to 0$$

This uses the following

Lemma 4.11 *If* $\mathbb{E}|X_N| \to 0$ *then* $X_N \xrightarrow{D} 0$

and the proof uses

$$|\mathbb{E}(e^{i\theta X_N} - 1)| \leq \theta \mathbb{E}|X_N| \to 0$$

We have another example, with a non continuous F.

Take $T_a = \inf\{N : S_N \geq a\}$ and we guess that

$$\sqrt{T_{\sqrt{N}a}} N \xrightarrow{D} \tau_a = \inf\{t : B_t = a\}$$

and we check this. We take

$$F(f) = \inf\{t : f(t) \geq a\} \wedge 1$$

where the minimum with 1 is for convenience. Then

$$F(X^{(N)}) = \sqrt{T_{\sqrt{N}a}} N + \text{ error}$$

and the error is at most $1/N/$. We need only show that

$$F(X^{(N)}) \to F(B)$$

but the problem is F is not continuous. We thus guess the discontinuity set of F:

$$\text{Disc}(F) \subset \{f : \tau_a(f) < 1, \exists \varepsilon > 0 \text{ such that } f(t) \le a \text{ for } t \in [\tau_a, \tau_a + \varepsilon]\}$$

and our aim is to show that

$$\mathbb{P}(B \in \text{Disc}(F)) = 0$$

If we define $X_t = B_{T_a+t} - a$ then this is a new Brownian motion and $\mathbb{P}(X_t \le 0 \text{ for } t \in [0, \varepsilon]) = 0$ by the time inverted law of the iterated logarithm.

We will check the inclusion of the discontinuity set. We prove the complement statement. Choose $\delta > 0$ and assume that $\sup_{t \le F(f)-\delta} f(t) = a - \varepsilon$ for some $\varepsilon > 0$. Then if

$$||g - f||_\infty \le \varepsilon/2$$

and this implies that $F(g) \ge F(f) - \delta$. In other words, if $g \to f$ then $F(g) \ge F(f)$ and so we have

$$\liminf_{g \to f} F(g) \ge F(f)$$

which is semicontinuity.

We take an f where there exists $S_N \to F(f)$ and $f(X_N) > a$. We need to show that F is continuous at such an f.

If $f(S_N) = a + \varepsilon$ then take a g such that $||g - f||_\infty < \varepsilon/2$ and this will have $g(S_N) > a$. Thus $F(g) \le S_N$ and so $\limsup_{g \to f} F(g) \le F(f)$ and so we are continuous here.

Theorem 4.12 *The following are equivalent for X_N, X over (E, d).*

1. $\mathbb{E}(F(X_N)) \to \mathbb{E}(F(X))$ *for all bounded continuous functions* $F : E \to \mathbb{R}$.
2. $\mathbb{E}(F(X_N)) \to \mathbb{E}(F(X))$ *for all bounded uniformly continuous functions* $F : E \to \mathbb{R}$.
3. $\limsup_{N\to\infty} \mathbb{P}(X_N \in A) \le \mathbb{P}(X \in A)$ *for all closed* A.
4. $\liminf_{N\to\infty} \mathbb{P}(X_N \in A) \ge \mathbb{P}(X \in A)$ *for all open* A.
5. $\mathbb{P}(X_N \in A) \to \mathbb{P}(X \in A)$ *for all* A *such that* $\mathbb{P}(X \in \partial A) = 0$, *where* $\partial A = \overline{A} \setminus A^o$.
6. $\mathbb{E}(F(X_N)) \to \mathbb{E}(F(X))$ *for all measurable* F *such that* $\mathbb{P}(X \in \text{Disc}(F)) = 0$.

Proof $1 \implies 2$ is immediate.

$2 \implies 3$ Define $F_\varepsilon(x) = \left(1 - \frac{d(x,A)}{\varepsilon}\right)^+$ and this is uniformly continuous and converges to χ_A. Then

$$\mathbb{P}(X \in A) \leftarrow \mathbb{E}(F_\varepsilon(X)) = \lim_{N\to\infty} \mathbb{E}(F_\varepsilon(X_N)) \ge \limsup_{N\to\infty} \mathbb{P}(X_N \in A)$$

$3 \implies 4$

$$\mathbb{P}(X \in A) = 1 - \mathbb{P}(X \in A^C)$$

and so if A is open then A^C is closed.

$4 \implies 5$

$$\mathbb{P}(X \in A^o) \le \liminf \mathbb{P}(X_N \in A^o) \le \liminf \mathbb{P}(X_N \in A) \le \limsup \mathbb{P}(X_N \in A) \le \limsup \mathbb{P}(X_N \in \overline{A}) \le \mathbb{P}(X \in$$

and so

$$\mathbb{P}(X \in \overline{A}) - \mathbb{P}(X \in A^o) = \mathbb{X} \in \partial\mathbb{A} = 0$$

$5 \implies 6$ We observe that 5 is a special case with $f = \chi_A$ with Disc$f = \partial A$.
Now choose $\alpha_1 < \alpha_2 < ... < \alpha_N$ with $|\alpha_{i+1} - \alpha_i| < \varepsilon$.
We can approximate a general f by

$$f_\varepsilon(x) = \sum_1^N \alpha_i \chi_{f^{-1}(\alpha_i, \alpha_{i+1}]}$$

and with $|f_\varepsilon(x) - f(x)| \leq \varepsilon$. We will check $\mathbb{E}(f_\varepsilon(X_N)) \to \mathbb{E}(f_\varepsilon(X))$, and this is enough. We then apply part 5 with $A = f^{-1}(\alpha_i, \alpha_{i+1}] \cup \{x : f(x) = \alpha_i\} \cup \{x : f(x) = \alpha_{i+1}\}$.

We claim that

$$\text{Disc}(\chi_A) \subset \text{Disc}f \cup \{x : f(x) = \alpha_i\} \cup \{x : f(x) = \alpha_{i+1}\}$$

and we prove the complemented statement.

Pick $x \notin \text{Disc}f$ with $f(x) \neq \alpha_i$ and $f(x) \neq \alpha_{i+1}$. Take $x_N \to x$ and then as f is continuous we have $f(x_N) \to f(x)$, because $\chi(f(x_N) \in (\alpha_i, \alpha_{i+1}])$ is discontinuous only at α_i and α_{i+1}.

To apply 5 we need to know that $\mathbb{P}(X \in \{x : f(x) = \alpha_i\}) = 0$. To this end let $p_\alpha = \mathbb{P}(f(X) = \alpha) > 0$. Now $\{\alpha : p_\alpha \geq 1/N\}$ has at most N elements, and so $p_\alpha > 0$ for only countably many α. Let $Q = \{\alpha : p_\alpha > 0\}$. Then we need to choose $\alpha_1, ..., \alpha_N$ not lying in Q. As Q is countable this is easy.

$6 \implies 1$ is immediate.

Q.E.D.

5 Up Periscope

Suppose that a box has N balls, half of which are black and the other half are white. We draw at random until the box is empty. Let S_K denote the number of blacks left by draw K minus the number of whites by draw K. Let $X_t^{(N)} = \frac{S_{NT}}{\sqrt{N}}$ and we expect:

Theorem 5.1 $X^{(N)} \overset{D}{\to}$ *Brownian Bridge.*

We consider a population model. Let S_N be the population size at time N. Assume that there is a one half probability of one individual having 2 or zero offspring each time step.

Lemma 5.2

$$\mathbb{P}(S_N) > 0 | S_0 = 1) \sim \frac{2}{N}$$

Corollary 5.3

$$\mathbb{P}(S_N > 0 | S_0 = N) \sim 1 - (1 - \frac{2}{N})^N \sim 1 - e^{-2}$$

If we instead choose $S_0 = N$ and linearly interpolate to get $X_t^{(N)} = \frac{S_{Nt}}{N}$ we get

Theorem 5.4

$$X^{(N)} \overset{D}{\to} X$$

where X solves Feller's equation:

$$\frac{dX}{dt} = \sqrt{X} \frac{dB}{dt}$$

Such theorems are called diffusion approximations. One can split the proof into two parts:

1. Soft nonsense- Show there exists a convergent subsequence $X^{(N_l)}$
2. Characterise the limit points of $X^{(N}$

The first point is to do with compactness.

Theorem 5.5 (Kolmogorov) *Suppose $X^{(N)}$ is a continuous process and*

$$\mathbb{E}|X_t^{(N)} - X_s^{(N)}| \leq C|t-s|^{1+\varepsilon}$$

for all $s, t \leq 1$ and $N \geq 1$. Then $X^{(N)}$ has a convergent subsequence.

We can deduce this from compact sets $K \subset C[0,1]$. For example

$$\{f : |f(t) - f(s)| \leq C_1|t-s|^{\alpha}, |f(0)| \leq C_2\}$$

is compact.

The second point is specific to each convergence. We look only at the population one. How do we characterise Feller's diffusion. We would want

$$X_{t+\Delta t} - X_t \approx \sqrt{X_t} B(0, \Delta t)$$

The key estimate is the following:

$$X^{(N)}_{\frac{K+1}{N}} - X^{(N)}_{\frac{K}{N}} = \frac{S_{K+1}}{N} - \frac{S_K}{N}$$

and the bit on the right hand side has mean zero and variance S_K, which agrees with $\sqrt{S_K} B(0, \Delta t)$.