Brownian Motion

Motivation: Brownian Motion is a universal object

Let $S_n = X_1 + ... + X_n$ be a random walk. We know that

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

- What if we need more path properties, e.g.

$$P(T_a > t) \sim ? \text{ where } T_a = \inf \{ n: S_n \geq a \}$$

- $\frac{S_{nt}}{\sqrt{n}} \xrightarrow{d} N(0,1)$ for fixed $t$. What if we look at the function $\{ t \to \frac{S_{nt}}{\sqrt{n}} \}$?

- Brownian Motion is the scaling limit of large class of R.W.s
- It has scaling properties
- Some computations are easier
- More difficult questions about R.W.s can be approximated, e.g.

$$P(S_n \geq x) \approx e^{-x^2/2n}$$

Definition: Brownian motion is a random function $\mathbb{R}_+: t \mapsto B_t$ s.t.

1. $\forall 0 \leq t_1 < t_2 < ... < t_k$ the increments

$$(B(t_{i+1}) - B(t_i))_{i} \text{ are independent, } N(0, t_{i+1} - t_i)$$

2. $t \mapsto B(t)$ is continuous

Question: Does B.M. exist? How to construct it?

Hints: Use the definition to show that for any $0 \leq t_1 < ... < t_k$

$$(B(t_1), ..., B(t_k))$$

has multivariate normal distribution
Let $\beta(t)$ be a B.H. i.i.d Uniform on $[0,1]$

$$\beta(t) = \begin{cases} 
\beta(t) & t \notin U \\
0 & t = U 
\end{cases}$$

Show that $\beta(t)$ i.i.d have the same f.d.d. but $\hat{\beta}(t)$ is a.s. not continuous.

Construction of B.H. à la Lévy

We will construct the B.M. on $[0,1]$, but the extension is easy.

$$D_n := \left\{ \frac{k}{2^n} : k = 0, 1, \ldots, 2^n \right\}$$

$$D := \bigcup_{n=1}^{\infty} D_n \text{ the dyadic points of } [0,1]$$

We will construct a sequence of piecewise linear continuous functions with the right f.d.d. The limit will be the B.M.

\[ n = 0 \]

\[ n = 1 \]

$$\beta(t) := \frac{\beta(0) + \beta(t)}{2} + 2^{-1} Z_{1/2}$$

\[ \text{Var}(\beta(t)) = \text{Var}\left(\frac{\beta(0) + \beta(t)}{2}\right) + 2^{\varepsilon} \text{Var}(Z_{1/2}) \]

\[ n = 2 \]

\[ n = 3 \]

$$Z_{1/4} = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

So, for $t = 1/4, 3/4$ we have

\[ \beta(t) = \frac{\beta(t-2^{-n}) + \beta(t + 2^{-n})}{2} + 2^{-n} Z_{2^{-n}} \]

Computing the variance, we will see that it is the right one at $t = 1/4, 3/4$.

Proceed similarly at the next level of dyadic points why.

Need to show: 1) $B_n(t)$ converges uniformly to a function with the B.M. properties.
Observations:
1) $\mathcal{B}_n(r)$ is continuous
2) $\forall s, r \in \mathbb{D}_n$, $\mathcal{B}_n(r) - \mathcal{B}_n(s)$ and $\mathcal{B}_n(r') - \mathcal{B}_n(s')$ are independent.

Explanation:

(Sort of obvious since the $\mathcal{B}_n(r) - \mathcal{B}_n(s)$ involves randomness only from the Gaussians between $(s, r)$ and $\mathcal{B}_n(r') - \mathcal{B}_n(s')$ between $(r, s')$.)

To check telescopic $\mathcal{B}_n(r) - \mathcal{B}_n(s)$ with the closest to $1, 0$ dyadic points in $\mathbb{D}_{n-1}$:

$$\mathcal{B}_n(r) - \mathcal{B}_n(s) = \left( \mathcal{B}_n(r) - \mathcal{B}_n(r - 2^{-n}) \right) + \left( \mathcal{B}_n(r - 2^{-n}) - \mathcal{B}_n(s + 2^{-n}) \right)$$

$$+ \left( \mathcal{B}_n(s + 2^{-n}) - \mathcal{B}_n(s) \right)$$

$$= \left( \frac{\mathcal{B}_{n-1}(r) - \mathcal{B}_{n-1}(r - 2^{-n})}{2} + 2^{-n+1} \mathcal{Z}_{r - 2^{-n}} \right) + \left( \frac{\mathcal{B}_{n-1}(s)}{2} + 2^{-n+1} \mathcal{Z}_s \right)$$

So the increment $\mathcal{B}_n(r) - \mathcal{B}_n(s)$ uses increments on disjoint intervals at level $n-1$ of the construction and the $\mathcal{Z}$-randomness inside $(s, r)$. Using an induction argument we have the conclusion.

3) $\mathcal{B}_n(r) - \mathcal{B}_n(s) \sim N(0, 1-\delta)$ for $r, s \in \mathbb{D}_n$. The argument is similar: Assume that $r, s \in \mathbb{D}_n \setminus \mathbb{D}_{n-1}$

$$\mathcal{B}_n(r) - \mathcal{B}_n(s) = \left( \mathcal{B}_n(r) - \mathcal{B}_n(r - 2^{-n}) \right) + \left( \mathcal{B}_n(r - 2^{-n}) - \mathcal{B}_n(s + 2^{-n}) \right)$$

$$+ \left( \mathcal{B}_n(s + 2^{-n}) - \mathcal{B}_n(s) \right)$$

By induction & Observation 2 all the $(\cdot)$ are independent Gaussians. So $\mathcal{B}_n(r) - \mathcal{B}_n(s)$ is Gaussian, mean $0$. 

Compute the variance:
\[
\text{Var}(\beta_n(r) - \beta_n(s)) = \text{Var}(\mathbf{1}) + \text{Var}(\mathbf{2}) + \text{Var}(\mathbf{3}) = \\
\left\{ \begin{array}{l}
\text{Var}(2^{-\frac{m+1}{2}} Z_r) + \text{Var} \left( \frac{\beta_{n-1}(r+2^{-m}) - \beta_{n-1}(r-2^{-m})}{2} \right) \\
+ \left\{ \begin{array}{l}
\text{Var}(2^{-\frac{m+1}{4}} Z_s) + \text{Var} \left( \frac{\beta_{n-1}(s+2^{-m}) - \beta_{n-1}(s-2^{-m})}{2} \right) \\
\text{by induction}
\end{array} \right. \\
\text{we induction again}
\end{array} \right.
\]
\[
\frac{2^{-(m+1)}}{4} + \frac{1}{2} 2^{-m+1} + (r-s) - 2^{-m+1} \\
+ 2^{-(m+1)} + \frac{1}{4} 2^{-m+1}
\]
\[
= r-s.
\]

So, if we define \( \beta(t) := \beta_n(t) \) when \( t \in D_n \), then we have defined a function with the correct statistics on dyadic points. Since this is a dense set, it remains to show that

CLAIM: \( (\beta_n(t)) \) converges in \( L^1 \phi \).

Consider the functions

- \( F_0(t) := \left\{ \begin{array}{l}
Z_1 \quad t=t=1 \\
0 \quad t=0
\end{array} \right. \)

- \( F_1(t) := \left\{ \begin{array}{l}
2^{-1/2} Z_{1/2} \quad t=t=1/2 \\
0 \quad t=0, 1
\end{array} \right. \)

- \( F_2 := \left\{ \begin{array}{l}
2^{-3/4} Z_{3/4} \quad t=t=3/4 \\
0 \quad t=0, 1, 1/2
\end{array} \right. \)

in general

- \( F_n(t) := \left\{ \begin{array}{l}
2^{-\frac{m+1}{2}} Z_t \quad t \in D_n - D_{n-1} \\
0 \quad t \in D_{n-1}
\end{array} \right. \)
Claim: $\beta(t) = \sum_{i=0}^{n} F_i(t) \quad \forall \, t \in D$

$\left(= \sum_{i=1}^{\infty} F_i(t) \quad \text{if} \quad t \in D_n \right)$

Show this by induction! Let $t \in D_n$

$\beta(t) = F_1(t) + \sum_{i=1}^{n-1} F_i(t)$

if $t \in D_{n-1}$, then $F_{n}(t) = 0$ \quad $\Rightarrow \beta(t) = \sum_{i=1}^{n-1} F_i(t) \quad \checkmark$

if $t \in D_n \cap D_{n-1}$, then

$\beta(t+2^{-m})$

$\beta(t) = 2^{-\lambda(t)/\epsilon} \sum_{i=1}^{n-1} F_i(t + 2^{-m})$

$\frac{1}{2} \left( \beta_{n-1}(t+2^{-m}) + \beta_{n-1}(t-2^{-m}) \right)$

by def., $\beta_{n}(t)$.

**Claim:** $\sum_{i=1}^{\infty} F_i(t)$ converges in $\mathbb{W}_{\text{law}} - \mathbb{P}$ a.s. (so $p(t)$ is a.s. continuous.)

For this, we need to show that $\sum_{i=1}^{\infty} \mathbb{W}_{\text{fill}} \leq \mathbb{P}$ a.s.

We have that $\mathbb{W}_{\text{fill}} \leq 2^{-\lambda/2}$ more $12_+1$ & estimate $\mathbb{P}(12_+1 > c \sqrt{n}) \leq \sum_{n} \sum_{t \in D_n} \mathbb{P}(12_+1 > c \sqrt{n})$

Check

$\leq \sum_{t} \left( 2^{\lambda(t)} \right) e^{-c^2/2} \leq \mathbb{P}$ if $c > \sqrt{2 \log 2}$.

So, by Borel-Cantelli, for all large enough $n$: $12_+1 < c \sqrt{n}$

and $\sum_{t} \mathbb{W}_{\text{fill}} \leq \sum_{n} 2^{-\lambda/2} c \sqrt{n} < \infty$. 
So we have shown that \( \beta(s) \) is a.s. continuous. It also has the right statistics:

**Claim:** \( \beta(s) \) has independent Gaussian increments

because we have shown this if \( t,s \in D \), it must hold

\[
D \ni t \rightarrow \left( \frac{\beta(t+h) - \beta(t)}{\sqrt{n}} \right) \quad \text{in distribution}
\]

\[
D \ni s \rightarrow \left( \frac{\beta(t+h) - \beta(s)}{\sqrt{n}} \right) \quad \text{in distribution}
\]

\[
\xi_i(t) := \frac{\beta(t+h) - \beta(t)}{\sqrt{n}} \quad \text{in distribution}
\]

\[
\xi_i(s) := \frac{\beta(t+h) - \beta(s)}{\sqrt{n}} \quad \text{in distribution}
\]

\[
\xi_i(t) \quad \text{in distribution}
\]

\[
\xi_i(s) \quad \text{in distribution}
\]

**Alternative Approach (Kolmogorov)**

Let \( X \subset C([0,1]) \) be the \( \sigma \)-algebra generated by f.d.d.'s

Assume we have a family of measures \( \{ \mu_F \} \) determining the f.d.d.'s.

Then \( \mu \) s.t. \( \mu([-1,1]) = \mu(\mathbb{R}) \)

\[
\int |x-y|^p \mu_{x,y} \, dx \, dy \leq C \int |t-s|^{1+\alpha}
\]

there exist measure \( \tilde{\mu} \) on \((X,\mathcal{B})\) s.t. \( \tilde{\mu} F = \mu_F \)

\[
\tilde{\mu} \left\{ \sup_{s,t} \frac{|x(t) - x(s)|}{|t-s|^{1/\alpha}} < \infty \right\} = 1
\]

**Corollary** B.H. is (at least) \( \alpha \)-Hölder continuous for any \( \alpha < \frac{1}{2} \)

**Proof** Compute \( \mathbb{E} \left[ \left| \frac{\beta(t+h) - \beta(t)}{\sqrt{n}} \right|^2 \right] \leq \text{const.} \mathbb{E} \left[ \left| \frac{\beta(t+h) - \beta(t)}{\sqrt{n}} \right|^2 \right] \)

Kolmogorov's then implies that

it is Hölder \( \frac{n-1}{2n} \rightarrow \frac{1}{2} \) as \( n \rightarrow \infty \).
Some basic properties of $B.M.$

1) Scaled

* $\forall a > 0 : \beta(at) = a \beta(t)$

because the function $t \to \beta(at)$ are continuous

$E(\beta(at) - \beta(at_0)) = N(0, a(t-t_0)) = a \beta(t-t_0)$

**Remark:** Be aware that the above scaling seem to lead to a contradiction (or confusion):

$$\int_{\beta(t_0)}^{\beta(t)} \beta(s)ds = \int_{\beta(t_0)}^{\beta(t)} \sqrt{\beta(s)^2} \cdot \frac{d\beta(s)}{ds} = \beta(t) - \beta(t_0)$$

**What's wrong?**

2) Inversion

Let $\beta(t)$ be B.M. Then

$$\beta(t) = \begin{cases} 0, & \text{if } t < 0 \\ \int_0^t \beta(s)ds, & \text{if } t \geq 0 \end{cases}$$

because we check first the odd's

$$E[\beta(t_1)] = 0 \quad \forall t > 0$$

$$E[\beta(t_1)^2] = ts \quad E[\beta(t_1)^4] = t^2 s^2$$

- Check also continuity: Let a sequence $t_n \downarrow t$

we have already checked that $\lim_{t_n \to t} \beta(t_n) = \beta(t)$

$$P\left(\lim_{n \to \infty} \beta(t_n) > y\right) = P\left(\bigcup_{n \geq 1} \{\beta(t_n) > y\}\right)$$

$$= \lim_{n \to \infty} P\left(\bigcup_{m \geq n} \{\beta(t_m) > y\}\right) = \lim_{n \to \infty} P\left(\bigcup_{m \geq n} \{\beta(t_m) > y\}\right)$$

$$= \lim_{n \to \infty} \lim_{m \to \infty} P\left(\bigcup_{m \geq n} \{\beta(t_m) > y\}\right) \underbrace{\text{reverse the steps}}$$

$$= P\left(\lim_{n \to \infty} \beta(t_n) > y\right) = 0 \quad \text{(by the continuity of } \beta(t)\text{)}$$

thus $\beta(t)$ is also B.M.
We have shown that B.H. is a.s. Hölder with exponent \( \alpha \) for any \( 0 < \alpha < 1/2 \). However,

\[
P \left( \sup_{t, s \in [0, 1]} \frac{|R(s) - R(t)|}{|t - s|^{1/2}} > \epsilon \right) = 0.
\]

Can you strengthen this to:

"B.H. is a.s. nowhere, Hölder 1/2."

The following theorems study the modulus of continuity.

**Thm 1** \( \exists \epsilon_0 > 0 \text{ s.t. a.s. for all } \epsilon \text{ sufficiently small (depending on the realization of B.H.)}

\[
\mathbb{P} \left( \sup_{t \in [0, 1]} \frac{|R(t + \epsilon) - R(t)|}{\sqrt{\log 1/\epsilon}} > \epsilon \right) 
\]

**Thm 2** \( \forall \epsilon < \sqrt{2} \)

\[
\mathbb{P} \left( \forall \epsilon > 0, \exists \epsilon_0 > 0, \delta > 0 \text{ s.t. a.s. for all } \epsilon \text{ sufficiently small (depending on the realization of B.H.)}

\[
\mathbb{P} \left( \sup_{t \in [0, 1]} \frac{|R(t + \epsilon) - R(t)|}{\sqrt{\log 1/\epsilon}} > \epsilon \right) < \epsilon
\]

We will prove these theorems using the construction of Lévy.

**Proof of Thm 1.** We have represented B.H. as

\[
\beta(t) = \sum_{n=0}^{\infty} \mathcal{F}_n(t)
\]

Recall:

\[
\mathcal{F}_0(t) : \quad \begin{array}{c}
0 \quad 2^{-1/2} \quad 2
\end{array}
\]

\[
\mathcal{F}_1(t) : \quad \begin{array}{c}
2^{-1/2} \quad 2
\end{array}
\]

\[
\mathcal{F}_2(t) : \quad \begin{array}{c}
2^{-1/2} \quad 2
\end{array}
\]
We will estimate

\[ |f(t+h) - f(t)| \leq \sum_{n=0}^{\infty} |f_n(t+h) - f_n(t)| \leq \sum_{n=0}^{\infty} l_{n+1} |f_n'(t)| + 2 \sum_{n=2}^{\infty} \frac{\|f_{n+1}\|_{L^p}}{\|f_n\|_{L^p}} \]

but (look at the figures)

\[ \|f_n\|_{L^p} \leq \frac{2^{\frac{n+1}{2}} \max_{t \in [a,b]} |f(t)|}{2^{n}} \quad \text{a.s.} \]

by Borel-Cantelli

\[ \leq C \sqrt{n} 2^{-\frac{n}{4}} \quad \text{for all large } n \text{, say } n \geq N \text{ with no random.} \]

Use this into 1 (after you first decompose further)

\[ \frac{1}{h} \sum_{n=0}^{N_0} \|f_n\|_{L^p} + \frac{1}{h} \sum_{n=N_0+1}^{L} C \sqrt{n} 2^{\frac{n}{4}} + 2 \sum_{n=N_0+1}^{\infty} \frac{\|f_{n+1}\|_{L^p}}{\|f_n\|_{L^p}} \]

\[ \leq \frac{1}{h} \sum_{n=0}^{N_0} \|f_n\|_{L^p} + \text{const.} \frac{L^{\frac{1}{2}}}{\sqrt{L^{\frac{1}{2}} + \text{const.}} \sqrt{2} 2^{-\frac{1}{4}}} \]

we choose \( L \) so that this sum is optimized (we \( L = \log \frac{1}{h} \))

\[ = \frac{1}{h} \sum_{n=0}^{N_0} \|f_n\|_{L^p} + \text{const.} \left( \frac{1}{h} \log \frac{1}{h} \right)^{\frac{1}{4}} \]

for all \( L \) small enough (depending on \( N_0 = N_0(\epsilon) \)) we have

that the above is \( \approx \text{const.} \left( \frac{1}{h} \log \frac{1}{h} \right)^{\frac{1}{4}} \)

We can tighten up this estimate by identifying the right constant \( c \).

Above the constant \( c \) was essentially the \( \sqrt{2\log 2} \) obtained from the

application of the Borel-Cantelli.

We can actually have the (optimal) estimate:
Refinement of Ilem 1: For every $c > \sqrt{2}$ it holds that

$$\mathbb{P}\left( \limsup_{n \to \infty} \sup_{x \in \mathbb{R}} \frac{1}{\sqrt{\log n}} \left| \frac{1}{n} \beta^{(n+1)} - \frac{1}{n} \beta^{(n)} \right| < c \right) = 1$$

Proof: Before proving it, it is important to understand why the previous estimate is not optimal. The functions $(F_n(x))_{n \in \mathbb{N}}$ represent the scales within Brownian Motion $\mathcal{L}$ where we estimated $\sup_{x \in \mathbb{R}} |Z_{+}^{(n)}|$ (they will with Bara-Cantelli) to control to $\mathbb{D}_n$ uniformly on all the scales.

The tighter estimate will come from sort of bypassing the uniform estimate $\mathcal{L}$ also bootstrap the $1/\log n$ order obtained before.

Define: $\Lambda_n (u_1) := \{ (u-\mathbf{1}_n) 2^{-u/a}, (u+\mathbf{1}_n) 2^{-u/n} \}$, $a, b \in \{ 0, \frac{1}{m}, \ldots, \frac{m}{m} \}$, $\mathcal{L} \in \{ 1, 2, \ldots, 2^n \}$.

$\Lambda_n (u_1) := \bigcup \Lambda_n (u_1)$

Compute:

$$\mathbb{P}\left( \sup_{u_1 \in \mathbb{R}} \sup_{u \in \{0, \frac{1}{m}, \ldots, \frac{m}{m} \}} \left| \frac{1}{n} \beta^{(u-\mathbf{1}_n) 2^{-u/n}} - \beta^{(u+\mathbf{1}_n) 2^{-u/n}} \right| > c \sqrt{2^{-u/n} \log 2^{-u/n}} \right)$$

$c$ will be $> \sqrt{2}$, but arbitrary $> c \sqrt{2^{-u/n} \log 2^{-u/n}}

\leq 2^n n^2 \mathbb{P}\left( \left| \beta^{(n+1)} - \beta^{n+1} \right| > c \sqrt{2^{-u/n} \log 2^{-u/n}} \right)

\leq 2^n n^2 \mathbb{P}\left( |\beta^{n+1}| > c \sqrt{\log 2^{-u/n}} \right)

\leq 2^n n^2 e^{-\frac{1}{2} c^2 \log 2^{-u/n}} = n^2 \exp \left\{ \log 2 \left( \frac{e^2}{2} (n-\alpha + \eta) \right) \right\}

\text{which is summable if } c > \sqrt{2}.

$$\mathbb{P}\left( |\hat{\beta}(u) - \hat{\beta}(u)| < c \sqrt{(t-s) \log \frac{1}{|t-s|}} \text{ for all } t, s \in \Lambda_n (u), \forall \eta > 0 \right) = 1,$$

if $c > \sqrt{2}$. 
This is enough, because for arbitrary $t$, s.e. $(\alpha, \eta)$ we can interpolate with t.s.e $\lambda_n(\eta)$ for some $\eta, \eta$.

(this is the purpose of the bit complicated definiton of $\lambda_n(\eta)$ involving the $a, b$-s.)

**Proof of** Thm 2 (lower bound).

Want to show that $\mathbb{P}\left( \liminf \sup \frac{|\beta(t+n) - \beta(t)|}{\sqrt{n \log n}} > \epsilon \right) = 1 \ \forall \epsilon > 0$

the probability is bounded below by

$$\mathbb{P}\left( \sup_{\eta, \eta} \frac{|\beta(\kappa\eta) - \beta(\kappa)\eta)|}{ce^{-\frac{\eta^2}{2}}} > \epsilon \right)$$

Compute

$$\mathbb{P}\left( \sup_{\eta, \eta} \frac{|\beta(\kappa\eta) - \beta(\kappa)\eta)|}{ce^{-\frac{\eta^2}{2}}} > \epsilon \ \forall \ k, \eta \right)$$

By Borel-Cantelli it suffices to estimate

$$\sum_{\eta} \mathbb{P}\left( \sup_{\kappa} \frac{|\beta(\kappa\eta) - \beta(\kappa)\eta)|}{ce^{-\frac{\eta^2}{2}}} < \epsilon \ \forall \ k, \eta \right)$$

$$\sum_{\eta} \mathbb{P}\left( |\beta(\kappa)\eta < c \eta^2 \right)$$

$$= \sum_{\eta} \left( 1 - \mathbb{P}(|\beta(\kappa)\eta > c \eta^2) \right)$$

$$\leq \sum_{\eta} \exp \left\{ -c^2 \epsilon^2 \left( 1 - 1 - \frac{c \epsilon^2}{2} \right) \right\}$$

If $c < \sqrt{2}$

We have, thus, prove that

(Levy's modulus of continuity)

$$\frac{\sup_{t \in [0, n]} |\beta(t+h) - \beta(t)|}{\sqrt{n \log n}} \xrightarrow{a.s.} \sqrt{2}$$

**Remark:** the Tim can be strengthened to lim
Next we will look at the "nowhere differentiability" properties of B.M.

We will start with a warm up:

**Prop.** B.M. is nowhere increasing.

**Proof.** Suppose that there is an interval \((a, b)\) of increase.

Split this interval into an arbitrary number of subintervals \((a_i, a_{i+1})\) with \(n\) arbitrary.

We have:

\[
P(\beta_i(t) \text{ is increasing in } (a_i, a_{i+1})) \leq
\]

\[
\leq \prod_{i=1}^{n} P(\beta(t_{i+1}) > \beta(t_i))
\]

\[
= \prod_{i=1}^{n} \left( \beta(t_i) > 0 \right)^{n-i} \to 0 \quad \text{as} \quad n \to 0.
\]

Thus Brownian motion is nowhere differentiable. In particular, we have that a.s. for all \(t\)

\[
D^\beta \beta(t) := \limsup_{h \to 0} \frac{\beta(t+h) - \beta(t)}{h} = +\infty
\]

or

\[
D_\beta \beta(t) := \liminf_{h \to 0} \frac{\beta(t+h) - \beta(t)}{h} = -\infty
\]

**Proof.** Assume that (a.s.) there exists \(t_0 \in \mathbb{R}\) s.t.

\[
-t_0 < D^\beta \beta(t_0) \leq D_\beta \beta(t_0) < +t_0
\]

Then

\[
\sup_{h} \left| \frac{\beta(t+h) - \beta(t)}{h} \right| \leq M \quad \text{for some} \quad M
\]

This will imply that for any \(j\)

\[
|\beta((n+i)2^{-n}) - \beta((n+i-1)2^{-n})| \leq |\beta((n+i)2^{-n}) - \beta(t_0)| +
\]

\[
+ |\beta((n+i-1)2^{-n}) - \beta(t_0)| \leq (2j+1)2^{-n}M
\]
Every though this is true for all $j$, we will only need it for $j=1, 2, 3$. We have:

$$
\Pr \left( \exists \epsilon > 0 : \frac{\left| \beta(h_0+\epsilon) - \beta(h_2) \right|}{\epsilon} \geq \frac{1}{n} \right)
\leq \Pr \left( \forall \epsilon > 0 : \left| \beta(h_0+h_j^{2^{-m}}) - \beta(h_2) \right| \leq (2j+1)2^{-m} \right)
\quad \text{for } j=1, 2, 3
$$

We want to show that this is $0$ a.s. so we use Borel-Cantelli:

Compute

$$
\sum_{n=1}^{\infty} \Pr \left( \exists \epsilon > 0 : \left| \beta(h_0+h_j^{2^{-m}}) - \beta(h_2) \right| \leq (2j+1)2^{-m} \right)
\leq \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} \Pr \left( \left| \beta(h_0+h_j^{2^{-m}}) - \beta(h_2) \right| \leq (2j+1)2^{-m} \right)
\leq \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} \left( \frac{1}{k} \right)^3 (14H 2^{-\epsilon})^3
\leq \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} (14H 2^{-\epsilon})^3
\leq \sum_{n=1}^{\infty} 2^n (14H 2^{-\epsilon})^3 < \infty.
$$

(because $\int_{-a}^{a} e^{-x^2} dx \leq 2\epsilon$)
Prop: \( \exists t \in (0,1) \) s.t. D^+(t) = 0

Proof: Define the function

\[ g(x) = \sup \{ t \in (\beta(x), \sup \beta(x)] : \beta(x) > t \} \]

Some properties of \( g(x) \):

1. \( g \) is strictly decreasing
2. \( g \) is left continuous (because if not, there would be accumulated times where B.H. would go arbitrarily close to a previously reached level, but would never reach it. This would violate continuity).
3. \( \beta(g(x)) = x \)
4. Jump points of \( g \) are dense
   (otherwise it would violate the no increasing property — look at the second picture).

Define \( V_\eta := \{ x : g(x-\eta) - g(x) \geq \eta \text{ for some } \eta \in (0,1) \} \)

- \( V_\eta \) is open (why?)
- \( V_\eta \) is dense: Suppose that \( I \) interval such that
  \( \forall \eta \in V_\eta \) in that interval \( g(x-\eta) - g(x) < \eta \).

Use (3) to write the inequality as

\[ g(x-\eta) - g(x) \leq \eta \left( \beta(g(x)) - \beta(g(x-\eta)) \right) \]

Call \( \frac{\eta}{\Delta^+} < \varepsilon \) for some \( \varepsilon \)

\( \Rightarrow \delta < \varepsilon \left( \beta(\varepsilon) - \beta(\varepsilon+\delta) \right) \) for all \( \varepsilon \) in an interval. But such an event has prob 0, or otherwise it would violate the non-increasing property.
By the Basic Category law

\[ \sqrt{\bigwedge_{i=1}^{n} x_i} \text{ is dense} \]

This implies that for all \( x \in V \),

\[ \exists x_{n+1} \text{ s.t.} \]

\[ g(x_{n+1}) > \gamma (x-x_{n}) = \]

\[ = \gamma \left( \psi(g(x))-\psi(g(x_{n+1})) \right) \]

Set \( t = g(x_{n+1}) \)

\[ t - t > \gamma \left( \beta(t)-\beta(t+n) \right) \]

\[ \Rightarrow \quad -\frac{1}{\gamma} \left( t+n \right) < \beta(t+n)-\beta(t) \]

let \( t \to 2 \)

\[ 0 \leq \lim_{t \to n} \frac{\beta(t+n)-\beta(t)}{t+n-t} = D^* \quad \]

For \( t = g(x_{n+1}) \), with \( x \in V \), the opposite inequality

\[ D^*(t) \leq 0 \]

is also true: Why?

---

**Quadratic Variations & Variaton**

Thus:

Let \( 0 = t_0 \leq t_1 \leq \ldots \leq t_{m-1} \leq t_m = t \)

a nested partition, i.e., at each stage points are added.

\[ \Rightarrow \max \left( \begin{array}{c}
\frac{t_j}{t_{j+1}}
\end{array} \right) \to 0 \]

Then

\[ D_t : = \sum_{j=1}^{m} \left( \psi(t_j) - \psi(t_{j+1}) \right)^2 \frac{P_{n-3}}{n-3} \to t \]

**Proof**

First, compute the expected value & they have variance.

\[ \text{E} D_t = \sum_{j=1}^{m} \text{E} \left( \psi(t_j) - \psi(t_{j+1}) \right)^2 = \]

\[ = \sum_{j=1}^{m} \left( \frac{t_j}{t_{j+1}} \right)^2 = t. \]
\[ \text{Var}(D_n) = \sum \text{Var} \left( \beta (t_j^n - \beta(t_j^{n-1}))^2 \right) \]
\[ = \sum \text{Var} \left( \beta (t_j^n - \beta(t_j^{n-1}))^2 \right) - \left( \text{Var} \left( \beta (t_j^n - \beta(t_j^{n-1})) \right) \right)^2 \]
\[ = 2 \sum \text{Var} \left( t_j^n - t_j^{n-1} \right)^2 \leq 2 \max \sum \frac{y_j}{\delta} \sum (t_j^n - t_j^{n-1}) \]
\[ \to 0. \]

**Remark:** The above computations show that \( D_n \) converges to 0 in probability 2 for this we don't need the nested assumption. If also \( \| t_j^n - t_j^{n-1} \| \to 0 \) fast enough, e.g., \( t_j = j 2^{-n} \), so that \( \sum \sum (t_j^n - t_j^{n-1})^2 \to 0 \), then Park-Carroll implies a.s. convergence.

Interestingly, we need this nested assumption if we want the a.s. convergence but we can wave it if \( \| t_j^n \| \) is nested. This uses a HG argument:

**a.s. convergence via HG argument:** Let \( X_n := \sum \text{Var} \left( \beta (t_j^n - \beta(t_j^{n-1}))^2 \right) \)
\[ \mathbb{E} [ X_n | G_n ] \]
\[ = \sum \text{Var} \left( \beta (t_j^n - \beta(t_j^{n-1}))^2 \right) \]
\[ = \sum \text{Var} \left( \beta (t_j^n - \beta(t_j^{n-1})) \right)^2 + \text{Var} \left( \beta (t_j^n - \beta(t_j^{n-1})) \right) \]
\[ = \sum \text{Var} \left( \beta (t_j^n - \beta(t_j^{n-1})) \right)^2 + \text{Var} \left( \beta (t_j^n - \beta(t_j^{n-1})) \right) \]
\[ = X_n \]
We will need to explain why the cross expectation equals 0, but before let's conclude:

\[ E[X_t | G_t] = X_t \quad \Rightarrow \quad X_t = E[X_t | G_t] \]

and that \((X_t)_{t \in \mathbb{N}}\) is a reverse MG. By a theorem of Levy, a reverse MG has a.s. a limit, thus \( \lim X_t \) exists a.s. and actually equals \( E[X_t | G_\infty] \). For a reference regarding the backward MG convergence, you may look at Durrett's "Probability: theory & examples".

Now, let's check why the cross expectation vanishes. Let's denote

\[ X_t := \beta(t_j^n) - \beta(t_{j-1}^n) \quad \text{and} \quad Y_t := \beta(t_j^n) - \beta(t_j^n) \]

These are independent Gaussian. We can write the cross expectation as

\[ E[XY | G_t] = E[E[XY | X_t] | G_t] \quad \text{and it suffices to show that} \quad E[XY | X_t] = 0. \]

This can be actually reduced to checking that \( E[XY | X^2 + Y^2] = 0 \) for \( X, Y \) independent standard normal if this is explained in class.

Ex: Show that the nesting condition is necessary for a.s. convergence, by showing that \( \exists (t_j^n) \) with mesh going to 0 s.t.

\[ \lim_{n \to \infty} \sum_{j=1}^{X_n} (\beta(t_j^n) - \beta(t_{j-1}^n))^2 = \infty \quad \text{a.s.} \]

Ex: Shows that B.H. has a.s. 100% variance
**B.H. as a Markov Process**

**Def** (d-dimensional B.H.)

Let $\beta_1, \ldots, \beta_d$ be independent B.M. starting at $\beta_1, \ldots, \beta_d$. Then $(\beta_1(t), \beta_2(t), \ldots, \beta_d(t)) =: \beta(t)$ is the d-dim. B.M. starting at $\beta_1, \ldots, \beta_d$.

**Notation**: $P_x, x \in \mathbb{R}^d$, will denote the distribution of d-dim B.M. starting at $x$ (Wiener measure).

**Prop** Let $\beta(t): t \geq 0$ be the B.M. on $\mathbb{R}^d$. Then $t \geq 0$

$\beta(t+\tau) - \beta(s) : t \geq 0$ is B.M. starting at 0 i.e.

it is independent of $\beta(t): 0 \leq t \leq s$.

**Prop** B.H. is Markov

**Proof**: refer to the book of Mörters-Peres.

**Def** Filtration is a sequence of increasing $\sigma$-algebras

\[ F_t = \sigma(\beta(2^{-n}, t \leq n)) \]

Notice that $\beta(t)$ is measurable w.r.t. $F_t$ i.e. $\beta(t)$ is adapted.

Gery $\sigma$-algebra:

\[ F^+(s) = \bigwedge_{s \leq t} F_t \]

Then $\beta(t+\tau) - \beta(s) : t \geq 0$ is independent of $F^+(s)$

**Proof**: \[ \beta(t+\tau) - \beta(s) : i = 1, \ldots, k \]

and each of the RVs is independent of $F^+(s)$.

The same will hold in the limit $\sum \nu_s$. 

They (Blumenthal's 0-1 law)

\[ \forall A \in \mathcal{F}^+(0) \text{ it holds that } \mathbb{P}(A) \in \{0,1\}. \]

**Proof**

Any \( A \in \mathcal{G}(\beta(t): t \geq 0) = \bigcup_{t \geq 0} \mathcal{G}(\beta(t): t \geq 0) \)

but all these are independent of \( \mathcal{F}^+(0) \). So \( A \) is independent of itself. Hence

\[ \mathbb{P}(A) = \mathbb{P}(A \cap A) = \mathbb{P}(A)^2 \rightarrow \mathbb{P}(A) \in \{0,1\}. \]

They (Tail events)

Let \( x \in \mathbb{R}^n \) and \( A \in \mathcal{T} := \bigcap_{t \geq 0} \mathcal{G}(\beta(s): s \geq t) \). Then \( \mathbb{P}_x(A) \in \{0,1\} \)

**Proof**

Suppose \( x = 0 \).

\[ K(t) := \begin{cases} \mathbb{P}(\beta(t) \leq 0) : t \geq 0 \\ 0 : t > 0 \end{cases} \]

is a B.M. A tail event for \( \beta(\cdot) \) is a germ event for \( x(\cdot) \) and its probability is \( \{0,1\} \).

**Stopping Times**

**Def** \( \tau \) is a stopping time w.r.t. a filtration \( \mathcal{F}_t \) if

\[ \mathbb{F}_\tau \subseteq \mathcal{F}_\tau \forall \tau. \]

**Ex** 1) any deterministic time is a stopping time.

2) if \( \{ \tau_n \} \) are stopping times \& \( \tau_1 \leq \tau_2 \) then \( \tau_2 \) is a stopping time

\[ \text{[because } \mathbb{F}_\tau \subseteq \bigcap_{n \geq 1} \mathcal{F}_{\tau_n} \text{]} \]

3) a stopping time \( \tau \) w.r.t. \( \mathcal{F}_\tau = \mathcal{G}(\beta(s): s \geq t) \) is a stopping time w.r.t. to \( \mathcal{F}_\tau^+ \) as well

\[ \text{[since } \mathcal{F}_\tau^+ \subseteq \mathcal{F}_\tau \text{]} \]
4) Hitting Times
Let \( \mathcal{H} \) be a closed set, then

\[ T := \inf \{ t > 0 : \beta(t) \in \mathcal{H} \} \]

is \( \mathcal{F}_t \)-stopping time.

[ because \( \{ \exists t > 0 : \beta(t) \in \mathcal{H} \} \] for some set \( \mathcal{H} \)

\[ = \bigcap_{s \in \mathbb{Q} \cap (0, T)} \bigcup_{x \in \mathbb{Q}^d \cap \mathcal{H}} \mathcal{F}_t(e) \]

picturewise

\[ \text{ } \]

5) Let \( \mathcal{G} \subset \mathbb{R}^d \) open. \( T := \inf \{ t > 0 : \beta(t) \in \mathcal{G} \} \)

They \( T \) is a stopping time w.r.t. \( \mathcal{F}_t^+ \) but not necessarily w.r.t. \( \mathcal{F}_t^+ \) ⊆ \( \mathcal{F}_t \) ▽

[ \{ \exists t > 1 : 0 < t < 1 \} \]

picturewise

\[ \text{ } \]

Remark: \( \mathcal{F}_t^+ \) is right continuous.

[ You should actually think that what this statement means.

\( \mathcal{F}_t^+ \) is the information you obtain from looking at BM at \( t^+ \) or right continuous should be deformed on \( \mathcal{F}_t^+ \) ⊆ \( \mathcal{F}_t^+ \)

Then the right continuity should be obvious but if you want a proof, here it is:
\[ \bigcap_{t_0 \geq 0} \mathcal{F}_{t_0}^+ = \bigcap_{t \geq 0} \bigcap_{t+t \in E} \mathcal{F}_{t+t}^+ = \bigcap_{t \geq 0} \mathcal{F}_{t+t}^+ = \mathcal{F}_t^+ \]

**Def:** Let \( \tau \) a stopping time. Define
\[ \mathcal{F}_\tau^+ = \{ A : \forall \tau \leq t \in \mathcal{F}_t^+ \} \]

In words: \( \mathcal{F}_\tau^+ \) is the information until time \( \tau \).
We use \( \mathcal{F}_\tau^+ \) instead of \( \mathcal{F} \) because of the situation with entrance times, cf. Ex. (5) above.

**Thm (Strong Markov Property)**
For every a.s. finite st. time \( \tau \) the process
\[ \{ \beta(t+\tau) - \beta(\tau) : t \geq 0 \} \]
is standard B.M. independent of \( \mathcal{F}_\tau^+ \).

**Corollary**
\[ \mathbb{E} \left[ \mathcal{F}_\tau \left( \beta(t+\tau) : t \geq 0 \right) \bigg| \mathcal{F}_\tau^+ \right] = \mathbb{E}_{\beta(\tau)} \left[ \mathcal{F}_\tau \left( \beta(t) : t \geq 0 \right) \right] \]

**Proof of Thm**
We start with a discretization. We define
\[ \tau_n = \sum_{k=1}^n \frac{\kappa}{2^n} \mathbf{1}_{E \left( \frac{k-1}{2^n}, \frac{k}{2^n} \right]} \]

Let \( \tau_n = \frac{\kappa}{2^n} \mathbf{1}_{E \left( \frac{k-1}{2^n}, \frac{k}{2^n} \right]} \).
\( \tau_n \) is what a mathematician would see as \( \tau \).

Let us look at \( \beta(t+\tau_n) - \beta(\tau_n) \) which equals to
\[ \beta(t+\kappa 2^n) - \beta(\kappa 2^n) \mathbf{1}_{E \left( \frac{k-1}{2^n}, \frac{k}{2^n} \right]} \]

& Let \( E \in \mathcal{F}_\tau^+ \), we will check that
\[ \{ \beta(t+\tau_n) - \beta(\tau_n) : t \in E \} \bigcap \mathcal{A} \] is independent of \( E \).
\[ P(\beta_{i} + \omega_{n} - \beta(\omega_{n}) \in A | \omega_{n}) = \sum_{k=0}^{\infty} P(\beta_{i} + \omega_{n} - \beta(\omega_{n}) \in A | \omega_{n}, \omega_{n} = \omega_{n}/2^k) \]

\[ = \sum_{k=0}^{\infty} \frac{1}{2^k} P(\beta_{i} + \omega_{n} - \beta(\omega_{n}) \in A | \omega_{n} = \omega_{n}/2^k) \]

Harren

\[ = \sum_{k=0}^{\infty} P(\omega_{n} = \omega_{n}/2^k) P(\beta_{i} \in A) \]

\[ = P(\omega_{n}) P(\beta_{i} \in A) \]

\[ \Rightarrow P(\beta_{i} + \omega_{n} - \beta(\omega_{n}) \in A | \omega_{n}) = P(\beta_{i} \in A) \]

and this implies that \( \{\beta_{i} + \omega_{n} - \beta(\omega_{n})\} \) is independent of \( F_{\omega_{n}} \supset F_{\omega_{n}+1} \).

Taking the limit \( n \to \infty \) we have that \( \omega_{n} \to 2 \) a.s. \( \omega_{n} \) by the continuity of \( B.H. \) \( \{\beta_{i} + \omega_{n} - \beta(\omega_{n})\} \xrightarrow{a.s.} \{\beta_{i} + \omega_{1} - \beta(\omega_{1})\} \)

which implies that the latter is also independent of \( F_{\omega_{1}} \).

*Q.* Find a process which is Harren but not Strong Harren.

**Brownian Motion as a H.M.**

(a quick reminder) **Def.** A process \( x(t) \) is a H.M. w.r.t. \( F_{t} \) if

\[ E[x(t + t/F)] = x(t) \text{ for } t > 0 \]

\[ \text{sub H.M.: } E[x(t + t/F)] \geq x(t) \]

\[ \text{sup H.M.: } E[x(t + t/F)] \leq x(t) \]

**Example.** B.M. is a H.M.

for set \( E[\beta(t)^{+}|F_{s}] = E[(\beta(t) - \beta(s)) + \beta(s)^{+}|F_{s}] = \beta(s) + E[\beta(t) - \beta(s)|F_{s}] \)

= \beta(s) + E[\beta(t) - \beta(s)|F_{s}] \text{ put by the Harren properly}.\]
Let \( x \) be a continuous MG \( \mathbb{E} \) \( \mathbb{F} \) stopping times

\[
\mathbb{E} \left [ x(t) \mid \mathbb{F}_t \right ] = x(t) \text{ with } x \text{ integrable}
\]

We will not prove this but there is a nice characterization of B.M. due to Lévy as a MG:

- B.M. is the unique continuous process \( \beta(t) = \beta(0) + \int_0^t \alpha(s) \, ds \)

But there are many more MGs associated to B.M. The following is very useful:

They let \( f : \mathbb{R}^d \to \mathbb{R} \) twice continuously differentiable \& \( x(t) \) a \( d \)-dimensional B.M. Also

\[
\mathbb{E} \left [ f(p_t) \right ] < \infty \quad \forall t > 0
\]

\[
\mathbb{E} \int_0^t \Delta f(p_s) \, ds < \infty \quad \forall t > 0
\]

They

\[
\frac{d}{dt} f(p_t) + \frac{1}{2} \int_0^t \Delta f(p_s) \, ds
\]

is a MG.

Before proving let's look at how we could guess this. To do so let's look at a RW \( (S_n)_{n \geq 1} \) and compute

\[
\mathbb{E} \left [ f(S_n) \mid S_{1}, \ldots, S_{n-1} \right ] = f(S_n) \quad \mathbb{E} \left [ f(S_n) \mid S_n \right ] = f(S_n) - \frac{1}{2} f(S_n + 1) + \frac{1}{2} f(S_n - 1) = f(S_n)
\]

\[
\frac{1}{2} (\Delta f)(S_n) \quad \text{where here } \Delta \text{ is the discrete Laplacian.}
\]
Proof: Let denote by $M_t := \mathbb{E}_x [\tau (B \cap I_1) - \mathbb{E}_x \left[ \int_0^t \Delta t (B (x)) \, dr \right] ]$

For $s < t$

$$E_x \left[ \mathbb{E}_x \left[ \tau (B \cap I_1) \mid I_s \right] - \frac{1}{2} \mathbb{E}_x \left[ \int_s^t \Delta t (B (x)) \, dr \mid I_s \right] \right] \leq 0 \quad \text{(1)}$$

The first term equals, by the Markov property,

$$E_x \left[ \mathbb{E}_x \left[ \tau (B \cap I_1) \mid I_s \right] \right] = \int \mathbb{P} (\tau (B \cap I_1) \leq t) \, d\mathbb{P} \left( \left. \frac{\rho \left( \frac{t}{s} \right) - \rho \left( \frac{s}{s} \right) }{s} \right\} \right) \quad \text{(2)}$$

where $P_+ (x, y) := \frac{1}{(2\pi t)^{d/2}} \exp \left( -\frac{|x - y|^2}{2t} \right)$

is the heat kernel.

The second term equals

$$\frac{1}{2} \int_0^S \Delta t (B (x)) \, dr + \frac{1}{2} \int_0^t \mathbb{E}_x \left[ \Delta t (B (x)) \mid I_s \right] \, dr$$

$$= \frac{1}{2} \int_0^S \Delta t (B (x)) \, dr + \frac{1}{2} \int_0^t \int \mathbb{P} (\tau (B \cap I_1) \leq t) \, d\mathbb{P} \left( \left. \frac{\rho \left( \frac{t}{s} \right) - \rho \left( \frac{s}{s} \right) }{s} \right\} \right) \quad \text{(3)}$$

Integrating by parts the second integral gives

$$\frac{1}{2} \int_0^S \Delta t (B (x)) \, dr + \frac{1}{2} \int_0^t \int f(t) \, d\mathbb{P}_0 \left( \left. \frac{\rho \left( \frac{t}{s} \right) - \rho \left( \frac{s}{s} \right) }{s} \right\} \right) \quad \text{(4)}$$

Keeping in mind that the heat kernel solves the heat equation

$$\frac{\partial \rho}{\partial t} = \frac{1}{2} \Delta \rho$$

$$\rho_0 (x,y) = \delta (x-y)$$

gives that the above equals

$$\frac{1}{2} \int_0^S \Delta t (B (x)) \, dr + \int_0^t \int f(t) \, d\mathbb{P}_0 \left( \left. \frac{\rho \left( \frac{t}{s} \right) - \rho \left( \frac{s}{s} \right) }{s} \right\} \right) \, ds$$

Interchanging the integrals equals

$$\frac{1}{2} \int_0^S \Delta t (B (x)) \, dr + \int_0^t \int f(t) \, d\mathbb{P}_0 \left( \left. \frac{\rho \left( \frac{t}{s} \right) - \rho \left( \frac{s}{s} \right) }{s} \right\} \right) \, ds$$

$$= \frac{1}{2} \int_0^S \Delta t (B (x)) \, dr + \int_0^t \int f(t) \, d\mathbb{P}_0 \left( \left. \frac{\rho \left( \frac{t}{s} \right) - \rho \left( \frac{s}{s} \right) }{s} \right\} \right) \, ds$$

$$= \frac{1}{2} \int_0^S \Delta t (B (x)) \, dr + \int_0^t \int f(t) \, d\mathbb{P}_0 \left( \left. \frac{\rho \left( \frac{t}{s} \right) - \rho \left( \frac{s}{s} \right) }{s} \right\} \right) \, ds$$

$$= \frac{1}{2} \int_0^S \Delta t (B (x)) \, dr + \int_0^t \int f(t) \, d\mathbb{P}_0 \left( \left. \frac{\rho \left( \frac{t}{s} \right) - \rho \left( \frac{s}{s} \right) }{s} \right\} \right) \, ds$$

$$= \frac{1}{2} \int_0^S \Delta t (B (x)) \, dr + \int_0^t \int f(t) \, d\mathbb{P}_0 \left( \left. \frac{\rho \left( \frac{t}{s} \right) - \rho \left( \frac{s}{s} \right) }{s} \right\} \right) \, ds$$

$$= \frac{1}{2} \int_0^S \Delta t (B (x)) \, dr + \int_0^t \int f(t) \, d\mathbb{P}_0 \left( \left. \frac{\rho \left( \frac{t}{s} \right) - \rho \left( \frac{s}{s} \right) }{s} \right\} \right) \, ds$$
Putting 1, 2, 3 together, we see that

$$E \int \mathcal{H}(t, \tau, x) = \mathcal{H}(t, x) - \frac{1}{2} \int_0^t \Delta \mathcal{H}(t, \tau, x) \, d\tau = M(t)$$

which is the MG property.

This then is important as it can provide many MGs. E.g. any $f$ that satisfies $\Delta f = 0$ gives a MG via

$$f(p(t))$$

for example $f(x, y) = e^{x+y} \cos x - y$.

The then will also provide the link to PDEs.

**Exercise:** Let $f \in C^{1,2}(\mathbb{R}^d \times \mathbb{R}^d)$. Show that

$$f(t, p(t)) - \int_0^t \left( \partial_t + \frac{1}{2} \Delta \right) f(t, p(t)) \, dt$$

is a MG.

Use this to show that $p(t)^2 + t$ is MG.

**Boundary Value Problems**

Thue: Let $\Omega$ is an open subset of $\mathbb{R}^d$

with smooth boundary $\partial \Omega$ and $u \in C^2(\Omega) \cap C(\bar{\Omega})$ is a solution of

$$\Delta u + f \circ u = 0$$

where $\Omega$ is the hitting time of $\partial \Omega$.  

**Proof:** For the moment, let’s assume that $\Omega$ is as (we will check this later). We know that

$$u(p(t)) - \frac{1}{2} \int_0^t \Delta u(t) \, dt$$

is a MG since $\Delta u = 0$.

Also, $u(p(t))$ is a MG. But also $u(p(t) + \tau)$ is a MG. So, for any $t \geq 0$,

$$\lim_{\tau \to 0} E u(p(t + \tau)) = 0$$

for any $x \in \Omega$.

Then

$$\lim_{\tau \to 0} E u(p(t + \tau)) - E [\lim_{\tau \to 0} u(p(t + \tau))] = -E u(p(t)).$$
What about the opposite?

Then $u(x_1) - E_x[t(\beta(x))]$ is a $C^2(L) \cap C(\overline{L})$ solution of $\Delta u = 0$.

\textbf{Proof:} \\
$u(x_1) = E_x [t(\beta(x))] = E_x [E_x [t(\beta(x)|E_x]]$ \\
where $\xi$ is the hitting time of $\partial B(x;r)$ with $r$ arbitrary (it will be taken no). Using the strong Markov property we write

the above as

$u(x_1) = E_x [E_x [t(\beta(x))] = E_x [u(\beta(x))$ \\
$= \int_{\partial B(x_1, r)} u(y) \pi(x;dy) = \int_{\partial B(0, r)} u(x+t\xi) \pi_0 (dz) \pi_0 (dy)$

This mean value property will imply that $u$ is smooth and satisfies $\Delta u(x) = 0$. For this we will need the observation that

$\pi_0 [dt ASN] = dt$ because of the symmetry of $B_M$. So let's check:

\textbf{$\Delta u(x) = 0$: Take } r \text{ small enough and expand by Taylor:}

$u(x_1) = \int_{\partial B(0, r)} u(x+t\xi) \pi_0 (dz) = \int_{\partial B(0, r)} u(x) + D_u x \cdot z + \frac{1}{2} \pi_0 (dz) \frac{1}{2} D_u z + o(r^2)$

$= u(x) + \frac{1}{2} D_u u \frac{1}{2} \pi_0 (dz)$

[we used the symmetry of $\pi_0$, which would imply that

$\int_{\partial B(0, r)} \pi_0 (dz) = \int_{\partial B(0, r)} D_u u \frac{1}{2} \pi_0 (dz) = 0$]

Deciding the relation by $r^2$ and letting $r \to 0$, we obtain that $\Delta u(x) = 0$. \qed
Smoothness: Consider the integral
\[
\int_{B(0,1)} \psi(1/r^d) u(x+z) \pi(dz)
\]
where \(\psi\) is a mollifier. Writing the integral in polar coordinates, we get it equals
\[
\text{const. } \int_0^1 \int_0^{2\pi} \psi(r^d) r^{d-1} u(x + rz) \lambda(B(r)) dr d\theta
\]
\[
= \text{const. } \int_0^1 \int_0^{2\pi} \psi(r^d) u(x) r^{d-1} dr d\theta
\]
\[
= \text{const. } u(x) \pi(dz)
\]
where \(\text{const.}\) are (different) constants. Since \(\psi u\) is smooth (this is a standard result), it follows that \(u\) is also smooth in \(\Omega\).

So, we have checked that \(u\), as defined in this way, is smooth and satisfies \(\Delta u = 0\) in \(\Omega\). We need to check that it is continuous up to the boundary. In general, the continuity up to the boundary might fail. This can happen if the boundary is not smooth, e.g., the part \(\partial\) of the boundary has the property that we can fit no cone with vertex the irregular point, which will belong exclusively in \(\Omega^c\). Such points are called irregular and have the property that
\[
P_x(\mathcal{Z}_{\Omega^c} = 0) > 0.
\]
So a B.H. that starts at \(x\) will a.s. exit somewhere far away from \(x\).
On the other hand, if \( D \mathcal{C} \) is smooth, then at each point \( \partial D \mathcal{C} \) we can fit a cone lying exclusively in \( \mathcal{C} \). This point will have the property that

\[
\mathbb{P}_x (z = 0) = 1
\]

with angle \( \alpha \).

\[
\text{[Why: by brownian scaling } \mathbb{P}_x (\beta(t) \in \mathcal{C}, \alpha) = C \alpha \]
\]

\[
\mathcal{C} = \lim_{t \to 0} \mathbb{P}_x (\beta(t) \in \mathcal{C}) \leq \lim_{t \to 0} \mathbb{P}_x (z > 0) \leq 1
\]

\[
\text{by Blumenthal, this implies that } \mathbb{P}_x (z = 0) = 1
\]

**Def.:** Any point which satisfies \( \mathbb{P}_x \) is called regular.

The above computation shows that any cone point is regular (in particular, any point which belongs to a smooth boundary). For such points we have:

**Prop.** If \( y \in \partial D \mathcal{C} \) is regular then

\[
\lim_{r \to y} \mathbb{E}_x [\tau_{\beta(y)}] = \tau_y
\]

**Proof.** By continuity of \( \mathbb{P}_x \), it suffices to show that

\[
\lim_{r \to y} \mathbb{P}_x (z < 0, \beta(t) \in \mathcal{B}(y; \delta)) = 1
\]

The last probability is

\[
\mathbb{P}_x (z < 0, \beta(t) \in \mathcal{B}(y; \delta)) \leq \mathbb{P}_x (z < 0, \beta(t) \in \mathcal{B}(y; \delta))
\]

\[
\leq \mathbb{P}_x (z < 0, \beta(t) \in \mathcal{B}(y; \delta))
\]

In the limit \( x \to y \) the first term is

\[
\mathbb{P}_x (z < 0) \text{ by lower semi-continuity}
\]

(to prove later)

So \( \mathbb{P}_x \) is
\[ \lim_{y \to -} P_y(t \geq s) - P_0 \left( \sup_{x \in \mathbb{R}^+} \beta(s, x) \geq 1 \right) \geq \lim_{y \to -} x - y \]

Let now \( t \to 0 \): the second term converges to 0 by the continuity of B.M. 2 the first to 1 by the assumption that \( y \) is regular.

**Proof of l.s.c.**

Use the Markov property to write
\[ P_x(t \leq 0) = \lim_{C \to 0} \int P_e(x, z) P_C(2 \leq t - e) \]

Since the heat kernel \( p_t(x, z) \) is smooth the above convolution is also a smooth function in \( x \), while the whole integral is increasing (to see this look at the line above before using the Markov). The increasing limit of continuous function is l.s.c.

**Prop (B.M. will exit a bounded domain in thick time)**

Let \( D \) a bounded domain \( 2: x = \text{int} \{ z: B(t) \subset D \} \)

Then \( P_x(t \leq 0) = \inf_{x \in \mathbb{R}} \]

**Proof**

\[ P_x(t \geq s) = E_x \left[ P_x(t \geq s \mid I) \mid I \right] \]

\[ = E_x \left[ P_{B(s)}(2 \leq s - 1) \mid I \right] \]

\[ \leq \sup \frac{P_s(2 \leq s - 1)}{P_x(2 \leq s)} \]

where \( R \): diameter of \( D \) & the inequality is because if B.M. has not exited \( D \) by time 1, then it could not exit the bigger ball \( B(x, R) \subset D \). But by translation invariance the last prob is independent of \( x \): \( P_x = \frac{P_0(B(x, R) \geq s) < 1}{2} \) we have that

\[ \sup_{x \in \mathbb{R}} P_x(t \geq s) < \infty \sup_{x \in \mathbb{R}} P_x(t \geq s - 1) < ... \]
Reflecting Principle:

Let $\beta(t)$ be a Brownian motion. Then
\[
P_0\left(\sup_{s \leq t} \beta(s) > a\right) = 2P_0(\beta(t) > a)
\]

\[P_0(\beta(t) > a) = P_0(\beta(t+2a) \geq 0, \tau \leq t)
\]
\[= E_0\left[P_0(\beta(t+2a) \geq 0, \tau \leq t) \mid \tau \leq t \right]
\]
\[= E_0\left[P_0(\beta(t+2a) \geq 0) \mid \tau \leq t \right]
\]
\[= \frac{1}{2} E_0 [1_{\tau \leq t}] = \frac{1}{2} P_0(\tau \leq t)
\]
\[= \frac{1}{2} P_0\left(\sup_{s \leq t} \beta(s) > a\right)
\]

Recurrence & Transience

Prop. Brownian motion is recurrent in $d \geq 2$ and transient in $d < 3$.

Proof

Let $u(x) = P_x(\tau \geq y)$. We know that $u$ is the solution to the b.v.p.
\[
\begin{cases}
\Delta u = 0 & \text{in } \mathbb{R}^2 \\
u = 0 & \text{on } \partial \mathbb{R}^2 \\
u = 1 & \text{on } \partial \mathbb{R}^2
\end{cases}
\]

Due to spherical symmetry, the solution should only depend on the radial part $u = u(r, \theta) = u(\theta)$. The Laplacian in polar coordinates is
\[
\frac{\partial^2 u}{\partial r^2} + \frac{1}{r} \frac{\partial u}{\partial r} + \frac{1}{r^2} \frac{\partial^2 u}{\partial \theta^2} = 0
\]

\[\Rightarrow r \frac{d}{dr} \left( r \frac{d}{dr} u \right) + \frac{\partial^2 u}{\partial \theta^2} = 0
\]

\[u(r) = \int_0^{2\pi} C_\theta \log r + C_\theta \, d\theta
\]

\[u(r) = \frac{C_1}{r} + C_2, \quad r \neq 0
\]
The boundary conditions imply that

\[ u(x) = \frac{\log |x| - \log R}{\log r - \log R}, \quad \text{if } d = 2 \]

\[ u(x) = \frac{1}{2} \frac{1 - x^d}{R^d - x^d}, \quad \text{if } d > 2 \]

We then have that

\[ P_x(\tau_\rho < \infty) = \lim_{R \to \infty} P_x(\tau_\rho < 2R) \]

\[ = \lim_{R \to \infty} \left\{ \begin{array}{ll}
\frac{\log |x| - \log R}{\log r - \log R} = 1, & \text{if } d = 2 \\
\frac{2^d - R^d}{2^d - R^d} = \left( \frac{1}{2} \right)^2, & \text{if } d > 2
\end{array} \right. \]

Prop: For \( d > 2 \), \( |\beta(t)| \to \infty \) a.s.

Let \( A_n = \{ |\beta(t)| \geq \sqrt{n}, \quad t \geq t_n \} \)

where

\[ t_n = \inf \left\{ t : |\beta(t)| \geq n \right\} \]

Compute for \( \lambda \) finite

\[ P_x(A_n^c) = E_x \left[ P_{\beta(t_n)}(2, \sqrt{n}) \right] \]

\[ = E_x \left[ \left( \frac{|\beta(t_n)|}{\sqrt{n}} \right)^{2-d} \right] \]

\[ = \left( \frac{\lambda}{\sqrt{n}} \right)^{2-d} \xrightarrow{n \to \infty} 0 \]

We then have

\[ P_x(\{ A_n \text{ i.o.} \}) = \lim_{n \to \infty} P_x(\bigcup_{m \geq n} A_m) = \lim_{n \to \infty} P(A_m) = 1 \]

and this implies the result.
let \( B \subset \mathbb{R}^d \). The occupation measure

\[
\mu(B) := \int_0^\infty 1_B(\beta(s))ds
\]

is the total amount of time that \( \beta \cdot M \) spends in the set \( B \). It should be intuitively clear that for any ball \( B \subset \mathbb{R}^d \)

\[
\begin{align*}
\mu(B) &< \infty, \quad \text{if } d > 3. \\
\mu(B) &= \infty, \quad \text{if } d \leq 2.
\end{align*}
\]

Let's compute

\[
E_x[\mu(B)] = E_x\int_0^\infty 1_B(\beta(s))ds = \\
= \int_0^\infty E_x[1_B(\beta(s))]ds = \\
= \int_0^\infty \int_B d\gamma P_x(x\gamma) = \\
= \int_0^\infty \int_B d\gamma \frac{1}{(2\pi s)^{d/2}} e^{-\frac{|x-\gamma|^2}{2s}} = \\
= \begin{cases} \\
\int_B d\gamma \text{ const.}(d) \left| x-\gamma \right|^{2-d} & \text{if } d > 3 \\
\infty & \text{if } d = 1, 2 \text{ because } \\
\int_0^\infty P_x(x\gamma) ds = \infty \text{ in this case.}
\end{cases}
\]

Let us, now, try to perform a more advanced computation. Let \( D \subset \mathbb{R}^d \) a bnd. domain \( \beta \notin \partial D \in C(D) \)

\[
\mathcal{U}(x) = E_x \int_{\partial D} \hat{f}(\beta(s))ds.
\]
Using the NG formulation, we can relate \( u \) to the solution of the b.s.p.

\[ \Delta u = f \text{ in } D \]

\[ u = 0 \text{ on } \partial D \]

But let us also do a different computation that will lead to a concept familiar from PDE's.

\[
u(x) = \mathbb{E}_x \int_0^\infty \int_D f(y) d\nu_s \Delta \text{c}_2 \text{d}s
\]

\[ = \int_0^\infty \mathbb{E}_x \left[ f(y) \text{c}_2 \right] \text{d}s
\]

\[ = \int_0^\infty \text{d}s \int d\gamma \ f(\gamma) P_\gamma \left( x = \gamma ; x_c = x \right)
\]

\[ = \int d\gamma \ f(\gamma) \mathbb{E}_x \left[ f(y) \right]
\]

\[ = \int d\gamma \ f(\gamma) G_D(x,\gamma)
\]

\( G_D(x,\gamma) \) is the Green's function corresponding to the Laplacian in domain \( D \) and formally is the solution to b.s.p.

\[ \Delta G_D(x,\gamma) = \delta_x \text{ in } D \]

\[ G_D(x,\gamma) = 0 \text{ on } \partial D \]

for any fixed \( x \in D \). It is not easy, in general, to obtain a closed form for \( G_D \) but we can prove the following properties:

1. \( G_D > 0 \)
2. \( G_D(x,\gamma) = G_D(\gamma,\gamma) \quad \text{if } x,\gamma \in D \)
3. \( G_D(x,\gamma) \approx \begin{cases} \frac{1}{|x-\gamma|^{d-1}} & \text{if } d \geq 3 \\ -\log |x-\gamma| & \text{if } d = 2 \\ |x-\gamma| & \text{if } d = 1 \end{cases} \), when \( x \approx \gamma \).
Let \( V \in C^2 (\mathbb{R}^d) \). Then

\[
\begin{aligned}
    u(t,x) &= E_x \left[ \frac{f(\beta(t))}{t} \exp \left( \int_0^t V(\beta(s)) \, ds \right) \right] \\
    \text{solves } \quad &\begin{cases}
    \mathcal{D}u + Vu = \frac{1}{2} \Delta u + Vu, \quad t > 0, \ x \in \mathbb{R}^d \\
    u(0,x) = f
    \end{cases}
\end{aligned}
\]

Proof. We first expand the exponential in the F.V.P. representation:

\[
\begin{aligned}
    1 + \sum_{n=1}^{\infty} \frac{1}{n!} E_x \left[ \frac{f(\beta(t))}{t} \left( \int_0^t V(\beta(s)) \, ds \right)^n \right] &
    = 1 + \sum_{n=1}^{\infty} \frac{1}{n!} E_x \left[ f(\beta(t)) \int_0^t \cdots \int_0^t V(\beta(s_1)) \cdots V(\beta(s_n)) \, ds_1 \cdots ds_n \right] \\
    \overset{\text{Symmetry}}{=} 1 + \sum_{n=1}^{\infty} \frac{1}{n!} \int_0^t \cdots \int_0^t E_x \left[ V(\beta(s_1)) \cdots V(\beta(s_n)) f(\beta(t)) \right] \, ds_1 \cdots ds_n \\
    = 1 + \sum_{n=1}^{\infty} \int_0^t \cdots \int_0^t E_x \left[ V(\beta(s_1)) \cdots V(\beta(s_n)) f(\beta(t)) \right] \, ds_1 \cdots ds_n
\end{aligned}
\]

Let us denote the \( n \)-th term in the above series by \( I_n(t,x) \).

By the change of variable \( S_i = t - y_i \), we can write it as

\[
I_n(t,x) = \int_0^t \cdots \int_0^t E_x \left[ V(\beta(t-y_1)) \cdots V(\beta(t-y_n)) f(\beta(t)) \right] \, dy_1 \cdots dy_n
\]

Use the Markov property at time \( t - y_1 \) to write the above as

\[
I_n(t,x) = \int_0^t \cdots \int_0^t E_x \left[ V(\beta(t-y_1)) \left( \int_{y_1}^t E_{\beta(t-y_1)} \left[ V(\beta(t-y_1)) \cdots V(\beta(t-y_n)) f(\beta(t)) \right] \, ds \right) \, ds_1 \cdots ds_n \right] \, dy_1 \cdots dy_n
\]

Differentiate w.r.t. \( x \), w.r.t. \( t \),
\[ \frac{d}{dt} I_{\eta}(t,x) = \int_{c \eta t - c \eta L}^{c \eta t + c \eta L} V(x) E_x \left[ \frac{\nu(\beta(x,y)) - \nu(\beta(x,y_n))}{\frac{d}{dt} \nu(\beta(y))} \right] \]

\[ + \int_{c \eta t - c \eta L}^{c \eta t + c \eta L} \frac{d}{dt} E_x \left[ \nu(\beta(x,y)) E_{\beta(x,y)} \left[ V(\beta(y_1,y_2)) - \nu(\beta(y_1,y_2)) \frac{d}{dt} \nu(\beta(y_2)) \right] \right] \]

but denoting \( F(y) := V(y) E_y \left[ V(\beta(y_1,y_2)) - \nu(\beta(y_1,y_2)) \frac{d}{dt} \nu(\beta(y_2)) \right] \)

the expectation in the last integral can be compactly written as

\[ E_x \left[ F(\beta(x,y)) \right] \]

& which solves the heat equation, i.e.

\[ \frac{d}{dt} E_x \left[ F(\beta(x,y)) \right] = \frac{1}{2} \Delta E_x \left[ F(\beta(x,y)) \right] \]

So, we have that

\[ \frac{d}{dt} I_{\eta}(t,x) = V(x) I_{\eta-1}(t,x) + \frac{1}{2} \Delta I_{\eta}(t,x) \]

Summing this over \( \sum_{n=1}^{\infty} I_{n}(t,x) \) we have

\[ \partial_t \sum_{n=1}^{\infty} I_{n}(t,x) = \frac{1}{2} \Delta \sum_{n=1}^{\infty} I_{n}(t,x) + \nu(x) \sum_{n=2}^{\infty} I_{n}(t,x) \]

Adding the constant term \( 1 \) inside the derivatives leads to

\[ \partial_t \sum_{n=0}^{\infty} I_{n}(t,x) = \frac{1}{2} \Delta \sum_{n=0}^{\infty} I_{n}(t,x) + \nu(x) \sum_{n=0}^{\infty} I_{n}(t,x) \]

or \( \partial_t = \frac{1}{2} \Delta u + \nu u \).
Using the F.-K. formula we will compute:

**Example (Arrival law)**

Let \( n_t := \frac{1}{t} \int_0^t 1_{(0,\infty)}(\theta_s) ds \) be the fraction of time \( \theta_s \) spends in the positive axis.

\[
\mathbb{P}_0(n_t \leq x) = \int_0^x \frac{1}{\pi \sqrt{y(1-y)}} dy = \frac{2}{\pi} \sin^{-1}(\sqrt{x})
\]

First note that

\[
\frac{1}{t} \int_0^t 1_{(0,\infty)}(\theta_s) ds = \frac{1}{x} \int_0^x 1_{(0,\infty)}(\theta_s) dy = \frac{1}{x} \int_0^x 1_{(0,\infty)}(\beta(s)) dy = \frac{1}{x} \int_0^x 1_{(0,\infty)}(\beta(s)) dy = \frac{1}{x} \int_0^x 1_{(0,\infty)}(\beta(y)) dy
\]

Let, now, \( \sqrt{x} := \int_0^x 1_{(0,\infty)}(\beta(y)) dy \). We will compute the Laplace transform of \( n_t \): \( E_x \exp \left\{ \int_0^t f(s) ds \right\} = E_x \exp \left\{ \int_0^x 1_{(0,\infty)}(\beta(y)) dy \right\} \)

but in order to make use of Feynman-Kac we will instead compute

\[
U_x(t, x) := E_x \exp \left\{ \int_0^t 1_{(0,\infty)}(\beta(y)) dy \right\}
\]

which by F.-K. solve

\[
\frac{\partial}{\partial t} U_x = \frac{1}{2} \Delta U_x - 6 V U_x
\]

\( U_x(0, x) = 1 \)

To solve this we use the Laplace method i.e. consider the transform

\[
\tilde{g}_c(A, x) := \int_0^x U_x(t, x) e^{-A t} dt
\]

This will lead to

\[
(A + 6 V) \tilde{g}_c - \frac{1}{2} \partial_x^2 \tilde{g}_c = 0
\]

\( \Rightarrow \quad \int_0^x (A + 6 V) \tilde{g}_c - \frac{1}{2} \partial_x^2 \tilde{g}_c = 0 \), \( x \geq 0 \)

\[
\partial_x^2 \tilde{g}_c - \frac{1}{2} \partial_x^2 \tilde{g}_c = 1 \quad \text{at} \quad x = 0
\]
Each one of the above is an ODE, which is easily solved

\[
\eta_c(A,x) = \begin{cases} 
\frac{1}{2 + c} + A e^{\sqrt{2(2+c)} x} + B e^{-\sqrt{2(2+c)} x}, & x > 0 \\
\frac{1}{c} + C e^{\sqrt{2c} x} + D e^{-\sqrt{2c} x}, & x < 0
\end{cases}
\]

The boundedness of \( \eta_c \) at \( x = \pm \infty \) and the continuity up to first derivatives at \( x = 0 \), determine the constants \( A, B, C, D \). Determining the constants & looking at the solution at \( x = 0 \), we get that

\[
\eta_c(A,0) = \frac{1}{\sqrt{2(2+c)}},
\]

By definition

\[
\eta_c(A,0) = \int_0^\infty e^{-\lambda t} u_c(x,\lambda) dt = \int_0^\infty e^{-\lambda t} E_0 e^{-\lambda} \int_0^{+\infty} V(\rho) \rho d\rho dt
\]

\[
= E_0 \int_0^\infty e^{-\lambda t} - c \int_0^{+\infty} V(\rho) \rho d\rho dt
\]

\[
= E_0 \int_0^\infty c e^{-\lambda t} (A + C) dt
\]

\[
= E_0 \left[ \frac{1}{A + C} \right]
\]

Setting \( \lambda = 1 \) & expanding both sides in \( c \) we obtain

\[
E_0 \frac{r_1}{\pi} = \frac{1}{2} \int_0^1 x^{n-t} (1-x)^{-t} dx
\]

\[
= \int_0^1 x^n \rho(x) dx
\]

\( \rho(x) \) is the arcsize law.
Let \((X_n)_{n \geq 1}\) be zero mean, variance 1 i.i.d. variables. Let
\[ S_n = \sum_{k=1}^{n} X_k, \quad n \geq 1. \]

By doing a linear interpolation we can consider this sequence as a piecewise linear function, which we denote by
\[ S^*(t) = \frac{S\lfloor nt \rfloor}{\sqrt{n}} + \frac{S\lfloor nt \rfloor - S\lfloor nt \rfloor}{\sqrt{n}}. \]

Donsker's theorem is a functional generalization of CLT.

**Theorem (Donsker's Invariance Principle)**

The sequence of continuous functions
\[ \{ S^*(t); t > 0 \} \xrightarrow{n \to \infty} \{ \beta(t); t > 0 \} \]
\[ \text{in } C([0,1]; \mathbb{R}) \]

Before proving this we will need to recall what is meant by weak convergence. If \((X_n)\) is a sequence of r.v.'s defined on some Polish space \(X\), then
\[ X_n \xrightarrow{w} X \quad \text{if for any } f \in C_b(X) \]
\[ E f(X_n) \to E f(X). \]

We recall the following theorem that provides equivalent checks for weak convergence.
They let \((X_n), \ X\) random variables on a metric space \((\mathbb{E}, \mathbb{P})\). They the following are equivalent

1. \(X_n \xrightarrow{d} X\)

2. \(\forall K \subseteq \mathbb{E} \text{ closed}
\limsup_{n \to \infty} \mathbb{P}(X_n \in K) \leq \mathbb{P}(X \in K)\)

3. \(\forall G \subseteq \mathbb{E} \text{ open}
\liminf_{n \to \infty} \mathbb{P}(X_n \in G) \geq \mathbb{P}(X \in G)\)

4. \(\forall A \text{ Borel with } \mathbb{P}(X \in \mathbb{E} \setminus A) = 0\), then
\(\lim_{n \to \infty} \mathbb{P}(X_n \in \mathbb{E} \setminus A) = \mathbb{P}(X \notin A)\)

By considering the distribution measures
\[\mu_n(A) := \mathbb{P}(X_n \in A)\]

we can recast the notion of weak convergence of r.v.'s to weak convergence of measures. Thus we say that
\[
\mu_n \Rightarrow \mu
\]

if one of the condition of the above thing is satisfied or if \(\forall f \in C_0(\mathbb{E}, \mathbb{P})\)
\[
\int f \ d\mu_n \to \int f \ d\mu.
\]

Recall the notion of tightness: A sequence of measures is tight if no mass escapes to the boundary of the metric space or to \(\infty\) (compare this with condition 4 in the above thing). The mathematical definition of tightness is that \(\forall \varepsilon > 0 \exists K \subseteq \mathbb{E} \text{ compact such that}
\sup_{n} \mu_n(K^c) < \varepsilon\).
A standard thing in probability says that if a sequence \( f_n \) is tight then it weakly converges subsequences.

If we want to show convergence of a sequence of measures or random variables, we first show that it is tight. Then we have to show that every subsequential limit is the same (i.e., uniqueness).

In the Donsker setting, in order to talk about tightness we first need to understand what are the compact sets of \( C ([0,1]; \mathbb{R}_+^d) \). We stress that tells us that these are the families of equicontinuous functions. For example Hölder continuous families with the same Hölder constant are equicontinuous. Essentially, this can reduce to checking the convergence of finite dimensional distributions, i.e., for \( t_1 < t_2 < \ldots < t_k \) it holds that

\[
\left( S_{n}^{*}(t_1), \ldots, S_{n}^{*}(t_k) \right) \xrightarrow{n \to \infty} (\beta(t_1), \ldots, \beta(t_k))
\]

which is just a consequence of the CLT.

Even though this is the most systematic approach to weak convergence, we will follow a different approach to prove Donsker's theorem.

Before going into the proof let's look at some applications...
Example 1. Let \( \{X_n\}_{n \geq 1} \) be i.i.d. mean 0, variance 1. Let
\[
S_n = X_1 + \cdots + X_n.
\]
Let \( M_n = \max \{ S_k : k \leq n \} \). They
\[
\Pr \left( \frac{M_n}{\sqrt{n}} > x \right) \to \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-\frac{1}{2}y^2} dy.
\]
This is a consequence of Donsker's invariance. But in order to apply it we first need to ask the question of whether
\[
M_n = \max \{ x(t) : 0 \leq t \leq 1 \}
\]
is a continuous functional of the path \( x(t) \). It is easy to see that the answer to this question is yes. They, Donsker implies that
\[
\frac{M_n}{\sqrt{n}} \to \max \left\{ \frac{1}{\sqrt{n}} S_k : k \leq n \right\}
\]
which they implies that
\[
\Pr \left( \frac{M_n}{\sqrt{n}} > x \right) \to \Pr \left( \max \{ \beta(t) : 0 \leq t \leq 1 \} > x \right) =
\]
\[
= 2 \Pr \left( \beta(1) > x \right) =
\]
\[
= \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-\frac{1}{2}y^2} dy.
\]

Example 2. Let \( T_n = \min \{ n : S_n \leq b \sqrt{n} \} \). We are tempted to say that \( T_n \) converges in distribution to the law of \( \inf \{ t : \beta(t) \leq b \} \). But, again, in order to apply Donsker we have to ask whether \( T \) is a continuous function of the underlying path. This is actually not the case. See for example the following figure:
The red and orange paths remain very close to each other for all times but they can have a very different first hitting time. However, even if this can certainly happen for “regular”, smooth paths, it cannot happen for Brownian like paths. That is
\[ P \left\{ X(t) : T_0(X(t)) \text{ is discontinuous} \right\} = 0. \]

In such situation Donsker’s Theorem can be still applied, and in this case we have
\[ P^{\text{law}} \left( T_{b_{0}} < t \right) \rightarrow P^{\text{law}} \left( T_{b_{0}} < t \right). \]

**Proof of Donsker’s Theorem**

We will need a lemma (which will be related to Skorohod embedding). Let \( \varepsilon > 0 \) be a fixed standard \( \beta = 1 \) stable random variable with
- \( X \) mean 0, variance 1, there exist stopping times
- \( 0 = T_{0} \leq T_{1} \leq T_{2} \leq \ldots \) \( S_{A} \).

6. \( \beta \left( T_{n} \right); n \geq 0 \) \( = \sum_{i=1}^{n} X_{i} \), \( S_{n} = X_{1} + \ldots + X_{n} \), \( n \geq 0 \).

(b) Consider the R.W. \( (S_{n})_{n \geq 0} = \left( \beta \left( T_{n} \right) \right)_{n \geq 0} \) and \( S_{n}^{\ast} (t) \) the linear interpolation scaled by \( \sqrt{n} \).

Then \( P^{\text{law}} \left( \sup_{0 \leq t \leq 1} \left| \frac{\beta(t)}{\sqrt{n}} - S_{n}^{\ast} (t) \right| > \varepsilon \right) \rightarrow 0. \)
let us assume the lemma for the moment and prove Denker.

**Proof of Denker**

Denote $W_n(t) := \frac{\beta(t)}{\sqrt{n}}$ for $t \geq 0$. By scaling invariance $W_n(t)$ has the same distribution like a standard BM.

Let $\mathcal{K} \subset C([a_1; b_1];\mathbb{R})$ closed and $\mathcal{K}_\varepsilon$ its $\varepsilon$-cover as

$$\mathcal{K}_\varepsilon := \{ f \in C([a_1; b_1];\mathbb{R}) : \| f - g \|_\infty < \varepsilon \ \forall g \in \mathcal{K} \}.$$  

They

$$P^x(\mathcal{S}_n \in \mathcal{K}) = P^x_{W_n(\mathcal{K})}(\mathcal{S}_n \in \mathcal{K})$$

$$\leq P(W_n \in \mathcal{K}) + P(\|W_n - \mathcal{S}_n\| > \varepsilon)$$

because

$$P(W_n = d\beta)$$

$$\xrightarrow{\varepsilon \downarrow 0} 0$$

by the previous lemma.

So we have that

$$\limsup_{n \to \infty} P^x(\mathcal{S}_n \in \mathcal{K}) \leq P(\beta \in \mathcal{K}) \downarrow P(\beta \in \mathcal{K})$$

[ because $\lim_{\varepsilon \downarrow 0} P(\beta \in \mathcal{K}_\varepsilon) = P(\beta \in \mathcal{K}) = P(\beta \in \mathcal{K})$]

The proof is now completed by the Thm with the equivalent criteria for weak convergence.

The proof of the lemma (see Márkus-Pálos pg.132) we use

**Thm (Skorohod embedding)**

If $\beta(t)$ is standard 1d BM and $X$ real valued with

$E[X] = 0$, $\text{Var}(X) < \infty$, there $\exists$ stopping time $\tau$ st.

- $\beta(\tau) = dX$
- $\tau \overset{W}{=} X$
- $E[\tau] = E[X^2]$
Proof of Skorohod (The Attia-Yor version)

Let's first look at the easiest case i.e. $X$ is mean-zero Bernoulli $X \sim \mathcal{B}(\alpha)$, with prob $p$

To satisfy $EX=0$ we must have $\alpha p = b(1-p) = c(p) = \frac{b}{b-a}$.

The embedding consists of stopping $\mathcal{H}$, starting from $0$ when it hits either $a$ or $b$, i.e. $\tau$ is the exit time from $(a,b)$. Since $P_{\alpha}(\tau_{a} < \infty) = \frac{b}{b-a}$ we have the embedding in this case.

- Assume wlog that $X$ takes values $x_1, x_2, \ldots, x_n$ with probabilities $p_1, p_2, \ldots, p_n$.

Define stopping times

$$T_i := \inf \{ t > 0 : \mathcal{H}(t) \notin (x_{i-1}, x_i) \}$$

$$T_i := \inf \{ t > 2T_{i-1} : \mathcal{H}(t) \notin (x_{i-1}, x_i) \}$$ for $i = 1, 2, \ldots, n-1$

Let the sought stopping will be $T := T_{n-1}$. Notice that $\beta(T)$ can take any of the values $x_1, x_2, \ldots, x_n$:

It can be equal to $x_1$, if $\beta(T_{i-1}) = x_1$ in which case $\beta(T_{i-1} + T_i) = x_1$

or it can equal $x_n$, if $\beta(T_{n-1}) = x_n$, in which case $\beta(T_{n-1} + T_{n-2}) = \beta_{n-1} = x_n$

$\beta_{n-1} = x_n$, $\beta_{n-2} = x_2$ in which case $\beta(T_{n-1} + T_{n-2}) = \beta_{n-2} = x_2$

$\beta(T_{n-2}) = \beta(T_{n-2} + \tau_{x_2}) = \beta_{n-2}$

$\beta(T_{n-2} + \tau_{x_2}) = \beta(T_{n-2} + \tau_{x_2} + \tau_{x_1}) = \beta_{n-2}$, and so on.
We have already computed the probability
\[ P(\beta(1) = x_1) = P(\beta(2) = x_1) = P(\gamma_1 = x_1) = p_1 \]

Let's compute
\[ P(\beta(1) = x_2) = P(\beta(2) = x_2) = P(\gamma_1 = x_1, \gamma_2 = x_2) P(\lambda_2 = x_2) \]
\[ = \frac{-x_1}{\gamma_2 - \gamma_1} \cdot \frac{\gamma_2 - x_1}{\gamma_2 - x_2} \]

Doing the computations using \( \gamma_i = \mathbb{E}[X | X=x_{i-1}] \)
\[ \sum_{j \geq i-1} \frac{x_j}{p_j} \frac{p_j}{\sum_{j \geq i-1} p_j} \]

you see that the above equals \( p_2 \).

Similar is the computation in the general case
\[ P(\beta(1) = x_1) = P(\beta(2) = x_1) = \ldots = P(\beta(n) = x_1) \]

This proves the first claim of Skorohod (in the case that \( X \) has a finite number of values).

- What about the claim that \( \mathbb{E}Z = \mathbb{E} \beta(2)^2 \)? This follows directly from the fact that \( Z \) is a stopping time and \( \beta(t+1) \) is a UMD.

- What about the general case of a real variable \( X \)?
In this case we define \( \mathbb{V}(x):= \mathbb{E} \left[ X | X \geq x \right] \)
\[ Z := \inf \left\{ t \geq 0 : H(t) \geq \mathbb{V}(\beta(t)) \right\} \]

where \( H(t) := \sup_{s \leq t} \beta(s) \)

We will not derive this, which can be done by approximation, but check that this definition is consistent with the above.
Local Times

Local time is the time that 1d S.M. has spent at a point \( x \in \mathbb{R} \) until time \( t \).

Notice that this requires some proper definition! But formally we can think of the local time (denoted by \( L_t(x) \))

\[
L_t(x) = \int_0^t \delta_x (\beta(s)) \, ds
\]

Even though this would be satisfactory in physics, some care is required so it turns out that local time can only be given a meaningful definition in dimension 1. One way to define is to use an approximation of the delta function \( \delta_x \)

\[
L_t(x) = \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} \int_0^t 1_{(x-\varepsilon, x+\varepsilon)} (\beta(s)) \, ds
\]

A meaning to this limit should be provided. We will follow an alternative way, that of downcrossings. We define this by first defining a sequence of stopping times:

Let \( a < b \)

\[
\tau_1 := \inf \{ s > 0 : \beta(s) = a \}
\]

\[
\tau_2 := \inf \{ s > \tau_1 : \beta(s) = a \}
\]

\[
\tau_3 := \inf \{ s > \tau_2 : \beta(s) = a \}
\]

\[
\tau_4 := \inf \{ s > \tau_3 : \beta(s) = a \}
\]

In pictures

The part of S.M. between times \( (\tau_1, \tau_2) \)
(cite the red part) are called the downcrossings.
The number of downcrossings by time $t$ is defined by

$$D(a, b; t) := \max \{ j : z_j \leq t \}$$

We will define the local time at zero as the limit

$$\lim_{a \to 0, b \to 0} \frac{2(b-a)D(a, b; t)}{a^2}$$

We will also look at the local time as a process $L$ shows that it is Hölder continuous, relate to PDEs with Neumann boundary condition and show that it is related to the process

$$M_t := \sup_{s \leq t} \beta(s).$$

**Lemma:** Let $\alpha \in (0, \infty)$, $(\beta(t)) : t \geq 0$ standard $\mathbb{B}M$.

$$T := \inf \{ t > 0 : \beta(t) = a \}$$

- $D := \# \text{downcrossings of } [a, b] \text{ up to } T$
- $D_1 := \# \text{downcrossings of } [a, m] \text{ up to } T$
- $D_u := \# \text{downcrossings of } [m, b] \text{ up to } T$

They

$$D_1 = X_0 + \sum_{j \geq 1} D_j$$

$$D_u = Y_0 + \sum_{j \geq 1} Y_j$$

where $X_0, X_1, \ldots \in \mathcal{Y}_0, Y_1, \ldots$ are independent sequences of independent variables, independent of $D$.

AND

$X_0 = \# \text{downcrossings of } [a, b] \text{ before first downcrossing of } [a, b]$

$Y_0 = \# \text{downcrossings of } [a, m] \text{ after last downcrossing of } [a, b]$
\[ X_j = \text{Geom} \left( N_+ \right), \quad p = \frac{u - a}{b - a} \quad j \geq 1 \]
\[ Y_j = \text{Geom} \left( N_+ \right), \quad p = \frac{b - u}{b - a} \quad j \geq 1 \]

**Proof**

The downcrossings of \((a, b)\) decompose the path into a number of independent sub-paths. The (green) downcrossings of \((u, b)\) is the number of attempts before a full downcrossing (orange path) of \((a, b)\) happens & the (red) downcrossings of \((a, u)\) are the attempts to go below level \(a\) & then above level \(b\). The probability of a green path is \( P_m(\mathcal{Z}_u < \mathcal{Z}_a) = \frac{u - a}{b - a} \). The probability of a red path is

\[ P_m(\mathcal{Z}_a < \mathcal{Z}_b) = \frac{b - u}{b - a} \, . \]

\( X_j \) is then the number of green paths in between the \( j^{th} \) & \( (j+1)^{th} \) \((a, b)\)-downcrossing.

\( Y_j \) is the number of red paths in between the \( j^{th} \) & \( (j+1)^{th} \) downcrossing.
Lemma: For any two sequences \( a, b \) and \( x, y \), the process \( \sum_{n=1}^{\infty} 2(b_{n+1} - a_n) D(a_n, b_{n+1}; T) \) is a sub-Markov w.r.t. to its own filtration (\( T \) is taken to be the hitting time of \( b \)).

Proof: Assume \( a_n = a, b_n = b \). Compute

\[
\mathbb{E} \left[ 2(b_{n+1} - a_n) D(a_n, b_{n+1}; T) \mid D_n \right] =
\]

\[
= \mathbb{E} \left[ 2(b_{n+1} - a_n) D(a_n, b_{n+1}, T) \mid D_n \right]
\]

where we define \( D_n := D(a_n, b_n; T) \).

#orange = \( D_n \). In the notation of the previous lemma

\[
D(a_n, b_{n+1}; T) = D_n = x_0 + \sum_{j=1}^{D_n} x_j
\]

so the above conditional expectation can be written as

\[
\mathbb{E} \left[ 2(b_{n+1} - a_n) \left( x_0 + \sum_{j=1}^{D_n} x_j \right) \mid D_n \right]
\]

\[
\geq 2(b_{n+1} - a_n) \mathbb{E} \left[ \sum_{j=1}^{D_n} x_j \mid D_n \right]
\]

\[
= 2(b_{n+1} - a_n) \sum_{j=1}^{D_n} \mathbb{E} \left[ x_j \mid D_n \right]
\]

\[
= 2(b_{n+1} - a_n) \sum_{j=1}^{D_n} \mathbb{E} \left[ x_j \right]
\]

\[
= 2(b_{n+1} - a_n) D_n \cdot \frac{b_n - a_n}{b_{n+1} - a_n} = 2(b_{n+1} - a_n) D_n,
\]