Advanced Probability

===

(2nd November, Friday)

Chapter 5. Weak Convergence

5.1. Definitions

Let E be a metric space. Whenever we are talking about a metric space, the σ-algebra is given by the Borel σ-algebra. Write $C_b(E)$ for the set of bounded continuous functions on E.

- Let $(\mu_n : n \in \mathbb{N})$ be a sequence of probability measures and let μ be another probability measure on E. We say that $\mu_n \to \mu$ **weakly** (as $n \to \infty$) if $\mu_n(f) \to \mu(f)$ for all $f \in C_b(\mathbb{R})$.

Theorem 5.1.1) The following are equivalent.

(a) $\mu_n \to \mu$ weakly on E

(b) $\liminf_{n\to\infty} \mu_n(U) \geq \mu(U)$ for all U open

(c) $\limsup_{\mu(F)} \leq \mu(F)$ for all F closed.

(d) $\mu_n(B) \to \mu(B)$ for all $B \in \mathcal{B}$ such that $\mu(\partial B) = 0$.(Boundary is the set of limit points of B that are not contained in B.)

    proof) Exercise.

For an example, consider a sequence $(x_n)_n \subset \mathbb{R}$ such that $x_n \to 0$ as $n \to \infty$. We want to have $\delta_{x_n} \to \delta_0$. Indeed, this is true in the weak sense. However, the sequence has $\delta_{x_n}(\{0\}) = 0$ for all n, hence we should have inequality in condition (c).

We have a similar version of the theorem for the real line.

Proposition 5.1.2) Consider the case $E = \mathbb{R}$. TFAE

(a) $\mu_n \to \mu$ weakly for some probability measure μ.

(b) $F_n(x) \to F(x)$ for all $x \in \mathbb{R}$ such that $F(x^-) = F(x)$. (Here, $F(x) = \mu((\infty, x])$ is the **distribution function** of μ.) (Sometimes called convergence of distributions)

(c) There exists a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and random variables X_n, X on Ω such that $X_n \sim \mu_n$, $X \sim \mu$ and $X_n \to X$ almost surely.

    proof) See probability and measure notes.

5.2. Prohorov's Theorem

When does a sequence of probability measures has a converging subsequence?

Let E be a metric space and $(\mu_n : n \in \mathbb{N})$ be a sequence of probability measures on E.

- We say that $(\mu_n)_n$ is **tight** if for all $\epsilon > 0$, there is a compact set $K \subset E$ such that

$$\mu_n(E \backslash K) \leq \epsilon \quad \forall n \in \mathbb{N}$$

For example, the sequence $(\delta_n)_n$ is *not* tight.

Theorem 5.2.1) Let $(\mu_n : n \in \mathbb{N})$ be a sequence of probability measures on a metric space E and suppose that $(\mu_n : n \in \mathbb{N})$ is tight. Then there exists a subsequence $(n_k)_k \subset \mathbb{N}$ and probability measure μ on E such that $\mu_{n_k} \to \mu$ weakly as $k \to \infty$.

This gives a version of weakly sequential compactness of probability measures. We are only going to prove this for $\mathbb{R}$. This theorem is hard to prove in general.(e.g. there is a method using Monge-Kantorovich metric defined for Polish spaces. For this method, see "Topics in Optimal Transport", C.Villani, Ame.Soc.Math. For the general version, see the attached note)

proof for $E = \mathbb{R}$) By a diagonal argument and by passing to a subsequence, it suffices to consider the case where $F_n(x) \to g(x)$ as $n \to \infty$ for all $x \in \mathbb{Q}$ for some $g(x) \in [0,1]$, where F_n is the distribution function of F_n. Now $g : \mathbb{Q} \to [0,1]$ is non-decreasing so g has a non-decreasing extension $G : \mathbb{R} \to [0,1]$, i.e.

$$G(x) = \lim_{q \searrow x, q \in \mathbb{Q}} g(q)$$

which has only countably many discontinuities.(because there should be a rational number in each discontinuity). Now we must have

$$F_n(x) \to G(x) \quad \forall x \text{ s.t. } G \text{is continuous at } x$$

Set $F(x) = G(x^+)$, then F and G have same points of continuity, so $F_n(x) \to F(x)$ for all $x \in \mathbb{R}$.

We are only left to check that $G(x) \to 1$ as $x \to \infty$ using tightness condition.

Since $(\mu_n : n \in \mathbb{N})$ is tight, given $\epsilon > 0$, there exists $R < \infty$ such that $\mu_n(\mathbb{R} \backslash (-R, R)) \leq \epsilon$ for all ϵ so $F_n(-R) \leq \epsilon$, $F_n(R) \geq 1 - \epsilon$. So

$$F(x) \to 0 \quad \text{as } x \to -\infty$$
$$F(x) \to 1 \quad \text{as } x \to \infty$$

So F is distribution function. So there exists a probability measure μ such that $\mu((-\infty, x]) = F(x)$. Then $\mu_n \to \mu$ by **Prop 5.1.2.**

(End of proof) $\square$

5.3. Weak Convergence and Characteristic Functions

Take $E = \mathbb{R}^d$.

- For a probability measures mu on $\mathbb{R}^d$, define its **characteristic function** $\phi : \mathbb{R}^d \to \mathbb{C}$ by

$$\phi(u) = \int_{\mathbb{R}^d} e^{i\langle u,x \rangle} \mu(dx)$$

Lemma 5.3.1) Fix $d = 1$. For all $\lambda \in (0, \infty)$,

$$\mu(\mathbb{R}\backslash(-\lambda, \lambda)) \leq C\lambda \int_0^\lambda (1 - \mathrm{Re}(\phi(u)))du$$

where $C = (1 - \sin(1))^{-1} < \infty$.

proof) Consider for $t \geq 1$. Let $A(t) = t^{-1}\int_0^t (1 - \cos v)dv$. Then

$$A(t) \geq A(0) = 1 - \sin(t)$$

(to see this, observe that $A(t)$ is the average of $(1 - \cos(v))$ on interval $(0, t)$ and divide the cases $|t| \leq \pi/2$ and $|t| \geq \pi/2$)

So $Ct^{-1}\int_0^t (1 - \cos(v))dv \geq 1$. Substitute $v = uy$, $u = v/y$,

$$Ct^{-1}\int_0^{t/y} (1 - \cos(uy))ydu \geq 1$$

Put $t/y = 1/\lambda$, $\lambda = y/t$, $t = y/\lambda \geq 1$ to see

$$C\lambda \int_0^{1/\lambda} (1 - \cos(uy))du \geq 1$$

whenever $t = y/\lambda \geq 1$(this was the assumption we started with). Now for general $y \in \mathbb{R}$, has

$$C\lambda \int_0^{1/\lambda} (1 - \cos(uy))du \geq 1_{|y| \geq \lambda}$$

Now integrate with respect to μ and use Fubini.

$$\mu(\mathbb{R}\backslash(-\lambda, \lambda)) \leq C\lambda \int_{\mathbb{R}} \int_0^{1/\lambda} (1 - \cos(uy))du\mu(dy)$$

$$= C\lambda \int_0^{1/\lambda} \int_{\mathbb{R}} (1 - \cos(uy))du\mu(dy)$$

(End of proof) $\square$

===

(5th November, Monday)

Theorem 5.3.2) Let μ_n, μ be probability measures on $\mathbb{R}^d$ with characteristic functions ϕ_n, ϕ. Then the following are equivalent

(a) $\mu_n \to \mu$ weakly on $\mathbb{R}^d$.

(b) $\phi_n(u) \to \phi(u)$ for all $u \in \mathbb{R}^d$.

We will prove only for the case $d = 1$.

> **proof)** It is clear that (a) implies (b). Suppose (b) holds. We prove via a 'compactness argument'. We aim to show that the sequence $(\mu_n)_n$ tight, and therefore has a converging subsequence, and show that the converging point is in fact μ.
>
> Note that $\phi(0) = 1$ and ϕ is continuous. Given $\epsilon > 0$, there exists $\lambda < \infty$ such that
>
> $$C\lambda \int_0^{1/\lambda} (1 - \mathrm{Re}(\phi(u)))du \le \epsilon/2$$
>
> with $C = (1 - \sin(1))^{-1} < \infty$. By dominated convergence,
>
> $$\int_0^{1/\lambda} (1 - \mathrm{Re}(\phi_n(u)))du \xrightarrow{n\to\infty} \int_0^{1/\lambda} (1 - \mathrm{Re}(\phi(u)))du$$
>
> so for sufficiently large n, by **Lemma 5.3.1,**,
>
> $$\mu_n(\mathbb{R}\backslash(-\lambda, \lambda)) \le C\lambda \int_0^{1/\lambda} (1 - \mathrm{Re}(\phi_n(u)))du \le \epsilon$$
>
> Since ϵ was arbitrary, we see that $(\mu_n : n \in \mathbb{N})$ is tight. By Prohorov's theorem, we have a converging subsequence $\mu_{n_k} \to \nu$ for some probability measure ν.
>
> Suppose for a contradiction that $\nu \neq \mu$. Therefore, there exists $\epsilon > 0$, and $f \in C_b(\mathbb{R}^n)$ such that
>
> $$|\mu_{n_k}(f) - \mu(f)| \ge \epsilon \quad \forall k$$
>
> By above argument, we have $\mu_{n_k} \to \nu$. But then, since e^{inx} is a bounded continuous function,
>
> $$\int_\mathbb{R} e^{inx}\nu(dx) = \lim_{k\to\infty} \phi_{n_k}(n) = \phi(n)$$
>
> which indicates $\mu = \nu$ by uniqueness of characteristic functions (see PM notes), a contradiction.
>
> *(End of proof)* $\square$

In fact, the proof of the theorem implies a slightly stronger statement, which is less useful.

Theorem 5.3.3) *(Lévy's continuity theorem for characteristic functions)* Let $(\mu_n : n \in \mathbb{N})$ be a sequence of probability measures on $\mathbb{R}^n$ with characteristic functions ϕ_n. Suppose $\phi_n(u) \to \phi(u)$ for all u for some function ϕ (not necessarily a characteristic function) such that ϕ is continuous at 0. Then ϕ is the characteristic function of some probability measure μ on $\mathbb{R}^d$ and $\mu_n \to \mu$ weakly on $\mathbb{R}^d$.

6. Large Deviations

6.1. Cramérs theorem

Theorem 6.1.1) Let $(X_n : n \in \mathbb{N})$ be a sequence of integrable *i.i.d.* random variables in $\mathbb{R}$. Set $m = \mathbb{E}(X_1)$, $S_n = X_1 + \cdots + X_n$. We know $S_n/n \to \delta_m$ in probability, so if $(m - \epsilon, m + \epsilon) \cap B = \phi$ then $\mathbb{P}(S_n/n \in B) \to 0$ as $n \to \infty$. Then in fact the convergence rate is given by $\sim \exp(-n\alpha(B))$ for some α. To be precise, for all $a \geq m = \mathbb{E}(X_1)$, as $n \to \infty$,

$$\frac{1}{n}\log \mathbb{P}(S_n \geq na) \to -\psi^*(a)$$

where ψ^* is the *Legendre transform* of the *cumulant generating function* $\psi(\lambda) = \log(\mathbb{E}(e^{\lambda X_1}))$, where Legendre transform is given by

$$\psi^*(x) = \sup_{\lambda \in \mathbb{R}}\{\lambda x - \psi(\lambda)\}$$

In particular, for n sufficiently large and in case $\psi^*(a) < \infty$, we get

$$-\psi^*(a) - \epsilon \leq \frac{1}{n}\log(\mathbb{P}(S_n \geq a)) \leq -\psi^*(a) + \epsilon$$

and therefore

$$e^{-n(\psi^*(a)+\epsilon)} \leq \mathbb{P}(S_n \geq na) \leq e^{-n(\psi^*(a)-\epsilon)}.$$

Note : ψ is always a convex function, so ψ^* is also a convex function.

Examples :

(i) $X_1 \sim N(0,1)$, then $\mathbb{E}(e^{\lambda X_1}) = e^{\lambda^2/2}$, $\psi(\lambda) = \lambda^2/2$ and $\psi^*(x) = x^2/2$. Hence

$$\frac{1}{n}\log(\mathbb{P}(S_n \geq a)) \to -\frac{a^2}{2} \quad \forall a \geq 0$$

Can check this directly, using the fact that $S_n \sim N(0,n)$ in this case.

(ii) $X_1 \sim \text{Exp}(1)$, then

$$\mathbb{E}(e^{\lambda X_1}) = \int_0^\infty e^{\lambda x}e^{-x}dx = \begin{cases} \infty & \text{if } \lambda \geq 1 \\ \frac{1}{1-\lambda} & \text{if } \lambda < 1 \end{cases}$$

so $\psi(\lambda) = -\log(1-\lambda)$ if $\lambda < 1$ and ∞ otherwise, and $\psi^*(x) = x - 1 - \log(x)$ for $x > 0$. Cramér's theorem implies that

$$\frac{1}{n}\log \mathbb{P}(S_n \geq na) \to -(a - 1 - \log(a)) \quad \forall a \geq 1$$

On the other hand, $\text{Var}(X_1) = 1 < \infty$, so $\frac{S_n-n}{\sqrt{n}} \to N(0,1)$ by CLT. So

$$\mathbb{P}(S_n \geq n + a\sqrt{n}) \to \int_a^\infty \frac{1}{\sqrt{2\pi}}e^{-x^2/2}dx$$

so Cramér's theorem gives a result of a different flavour from CLT for distributions with bounded variation : while CLT provides a description for distribution near the average, Cramér gives an explanation of tail distribution of S_n.

preparation for proof of Cramér's theorem) Let $\mu(B) = \mathbb{P}(X_1 \in B)$. Exclude the easy case where $\mu = \delta_m$. Define for $\lambda \geq 0$ with $\psi(\lambda) < \infty$, the **tilted distribution** μ_λ by

$$\mu_\lambda(dx) \propto e^{\lambda x}\mu(dx)$$

For $K \geq m = \mathbb{E}(X_1)$, define the conditional distribution by

$$\mu_K(dx|x \leq K) \propto 1_{\{x \leq K\}}\mu(dx)$$

The CGF(cumulant generating function) of μ_K is then given by

$$\psi_K(\lambda) = \log(\mathbb{E}(e^{\lambda X_1}|X_1 \leq K))$$

==

(7th November, Wednesday)

We now start proving the following theorem.

Theorem 6.1.1) Let $(X_n : n \in \mathbb{N})$ be a sequence of integrable *i.i.d.* random variables in $\mathbb{R}$. Set $m = \mathbb{E}(X_1)$, $S_n = X_1 + \cdots + X_n$. Then for all $a \geq m = \mathbb{E}(X_1)$, as $n \to \infty$,

$$\frac{1}{n}\log\mathbb{P}(S_n \geq na) \to -\psi^*(a)$$

where $\psi(\lambda) = \log(\mathbb{E}(e^{\lambda X_1}))$, and $\psi^*(x) = \sup_{\lambda \in \mathbb{R}}\{\lambda x - \psi(\lambda)\}$.

    proof) *(Upper bound)* For all $\lambda \geq 0$ and $n \geq 1$

$$\mathbb{P}(S_n \geq na) = \mathbb{P}(e^{\lambda S_n} \geq e^{\lambda na}) \leq e^{-\lambda na}\mathbb{E}(e^{\lambda S_n}) = e^{-(\lambda a - \psi(\lambda))n}$$

so $\frac{1}{n}\log\mathbb{P}(S_n \geq na) \leq -(\lambda a - \psi(\lambda))$ and

$$\frac{1}{n}\log\mathbb{P}(S_n \geq na) \leq -\psi^*(a)$$

(Lower bound) It remains to show the lower bound. That is, we aim to prove

$$\liminf_{n\to\infty}\frac{1}{n}\log\mathbb{P}(S_n \geq na) \geq -\psi^*(a)$$

Consider first the case where $\mathbb{P}(X_1 \leq a) = 1$. Then

$$\mathbb{E}(e^{\lambda(X_1-a)}) \xrightarrow{\lambda\to\infty} \mathbb{P}(X_1 = a)$$

Call $p = \mathbb{P}(X_1 = a)$, so $\lambda a - \psi(\lambda) \to -\log(p)$. So in particular,

$$\psi^*(a) \geq -\log(p)$$

Now $\mathbb{P}(S_n \geq na) = p^n$ so

$$\frac{1}{n}\log\mathbb{P}(S_n \geq na) = \log(p) \geq -\psi^*(a)$$

hence we can eliminate the case $\mathbb{P}(X_1 \leq a) = 1$.

    Next consider the case $\psi(\lambda) < \infty$ for all $\lambda \geq 0$ and $\mathbb{P}(X_1 > a) > 0$. Fix $\epsilon > 0$ and set $b = a + \epsilon$, $c = a + 2\epsilon$, choosing ϵ small enough so $\mathbb{P}(X_1 > b) > 0$. Then there exists λ such that $\psi'(\lambda) = b$ - where define the differentiability and the existence is justified in the following proposition :

Proposition 6.1.2) Suppose X is integrable and not a.s. constant. Then

$$\psi_K(\lambda) = \log \mathbb{E}(e^{\lambda X_1}|X_1 \leq K) < \infty \quad \forall K < \infty$$
$$and \quad \psi_K(\lambda) \nearrow \psi(\lambda) \quad \text{as } K \to \infty$$

Moreover in the case $\psi(\lambda) < \infty$ for all $\lambda \geq 0$, ψ has a continuous derivative on $[0, \infty)$ and is C^2 on $(0, \infty)$ with

$$\psi'(\lambda) = \int_{\mathbb{R}} x\mu_\lambda(dx)$$
$$\psi''(\lambda) = \mathrm{Var}(\mu_\lambda) > 0$$

and ψ' is a homeomorphism from $[0, \infty)$ to $[m, \sup(\mathrm{supp}(\mu))$.
proof) (Exercise)

Now we use the idea of tilting the probability measure. Define a new probability measure $\mathbb{P}_\lambda$ by $d\mathbb{P}_\lambda = e^{\lambda S_n - n\psi(\lambda)}d\mathbb{P}$. Then observe that under $\mathbb{P}_\lambda$ the random variables $X_1, \cdots, X_n$ are independent with distributions μ_λ and that $\mathbb{E}_\lambda(X_1) = b$. Consider the event

$$A_n = \{\left|\frac{S_n}{n} - b\right| \leq \epsilon\} = \{(b - \epsilon)n = an \leq S_n \leq (b + \epsilon)n = cn\}$$

By the weak law of large numbers, $\mathbb{P}_\lambda(A_n) \to 1$. So

$$\mathbb{P}(S_n \geq na) \geq \mathbb{P}(A_n) = \mathbb{E}_\lambda\left(1_{A_n} e^{-\lambda S_n + \psi(\lambda)n}\right)$$
$$\geq e^{-\lambda cn + \psi(\lambda)n}\mathbb{P}_\lambda(A_n)$$

So

$$\frac{1}{n}\log \mathbb{P}(S_n \geq na) \geq -\lambda c + \psi(\lambda) + \frac{\log(\mathbb{P}_\lambda(A_n))}{n}$$

and

$$\liminf_{n\to\infty} \frac{1}{n}\log \mathbb{P}(S_n \geq na) \geq -(\lambda c - \psi(\lambda)) \geq -\psi^*(c)$$

Now ψ^* is continuous at a (recall, ψ^* is a Legendre transform of a convex function so is convex, and therefore continuous. Or, see **Lemma 6.1.3**) and $\epsilon > 0$ is arbitrary so the desired lower bound follows on letting $\epsilon \to 0$.

Finally, consider the general case $\mathbb{P}(X_1 > a) > 0$ but allowing $\psi(\lambda) = \infty$ for some $\lambda \geq 0$. For $K > a$, we have $\mathbb{P}(X_1 > a|X_1 \leq K) > 0$ and $\psi_K(\lambda) < \infty$ for all $\lambda \geq 0$. So preceding argument shows

$$\liminf_{n\to\infty} \frac{1}{n}\log \mathbb{P}_K(S_n > na) \geq -\psi_K^*(a)$$

where $\mathbb{P}_K$ is the probability measure given by

$$d\mathbb{P}_K^{(n)} \propto 1_{\{X_1 \leq K, \cdots, X_n \leq K\}}d\mathbb{P}$$

(To see this, note, under $\mathbb{P}_K$, random variables $X_1, \cdots X_n$ are independent with distribution $\mu(\cdot|x \leq K)$). But

$$\mathbb{P}(S_n \geq na) \geq \mathbb{P}(S_n \geq na|X_1 \leq K, \cdots, X_n \leq K) = \mathbb{P}_K(S_n \geq na)$$

and $\psi_K^*(a) \searrow \psi^*(a)$ as $K \to \infty$ (by **Lemma 6.1.3**) so we see

$$\liminf_{n \to \infty} \frac{1}{n} \log \mathbb{P}(S_n \geq na) \geq -\psi_K^*(a) \nearrow -\psi^*(a)$$

(End of proof) □

One different way to see that ψ^* is continuous at a is presented in the following lemma.

Lemma 6.1.3) For all $a \geq m$, with $\mathbb{P}(X_1 > 0) > 0$ we have $\psi_K^*(a) \searrow \psi^*(a)$ as $K \to \infty$. Moreover in the case $\psi(\lambda) < \infty$ for all $\lambda \geq 0$, ψ^* is continuous at a and we have $\psi^*(a) = \lambda^* a - \psi(\lambda^*)$ where λ^* is uniquely determined by $\psi'(\lambda^*) = a$.

proof) Consider first the later case where $\psi(\lambda) < \infty$ for $\lambda \geq 0$. Then by **Proposition 6.1.2** wee see that

$$\psi^*(a) = \lambda^* a - \psi(\lambda^*)$$

where $a = \psi'(\lambda^*)$ and ψ^* is continuous at a with $\lambda^* = (\psi')^{-1}(a)$.

For the first part, note that ψ_K^* is non-increasing in K. For K sufficiently large, we have

$$\mathbb{P}(X_1 > a | X_1 \leq K) > 0$$

and $a \geq m \geq m_K$ (where $m_K = \mathbb{E}(X_1 | \leq X_1 \leq K)$) and $\psi_K(\lambda) < \infty$ for all $\lambda \geq 0$, so we may apply the preceding argument to μ_K to see that

$$\psi_K^*(a) = \lambda_K^* a - \psi_K(\lambda_K^*)$$

where $\lambda_K^* \geq 0$ is determined by $\psi_K'(\lambda_K^*) = a$. Now $\psi_K'(\lambda)$ is non-decreasing in K and λ, so $\lambda_K^* \searrow \lambda^*$ for some $\lambda^* \geq 0$. Also $\psi_K'(\lambda) \geq m_K$ for all $\lambda \geq 0$ so

$$\psi_K(\lambda_K^*) \geq \psi_K(\lambda^*) + m_K(\lambda_K^* - \lambda^*)$$

Then

$$\psi_K^*(a) = \lambda_K^* a - \psi_K(\lambda_K^*) \leq \lambda_K^* a - \psi_K(\lambda^*) - m_K(\lambda_K^* - \lambda^*) \to \lambda^* a - \psi(\lambda^*) \leq \psi^*(a)$$

So $\psi_K^*(a) \searrow \psi^*(a)$ as $K \to \infty$ as claimed.

(End of proof) □

===

(9th November, Friday)

7. Borwnian Motion

7.1. Definition

Let $(X_t)_{t \geq 0}$ is a random process in $\mathbb{R}^d$. We say $(X_t)_{t \geq 0}$ is a **Brownian motion** if :

(i) For all $s, t \geq 0$, the random variable $X_{s+t} - X_s$ is Gaussian, of mean 0 and variance tI and is independent of $\mathcal{F}_s^X = \sigma(X_r : r \leq s)$

(ii) for all $\omega \in \Omega$ the map $t \mapsto X_t(\omega) : [0, \infty) \to \mathbb{R}^d$ is *continuous*.

Condition (i) means that, for all $s \geq 0$, $t > 0$, all Borel sets $B \subset \mathbb{R}^d$ and all $A \in \mathcal{F}_s^X$,

$$\mathbb{P}(\{X_{s+t} - X_s \in B\} \cap A) = \mathbb{P}(A) \int_B (2\pi t)^{-\frac{d}{2}} e^{-|y|^2/2t} dy$$

Or, in terms of conditional expectation, (i) is equivalent to : for all $s, t \geq 0$ and all $f \in C_b(\mathbb{R}^d)$,

$$\mathbb{E}(f(X)_{s+t}|\mathcal{F}_s^X) = P_t f(X_x) \quad \text{a.s.}$$

where P_t is the **heat semigroup**, i.e.

$$P_0 f(x) = \int_{\mathbb{R}^d} p(t, x, y) f(y) dy, \quad p(t, x, y) = (2\pi t)^{-\frac{d}{2}} e^{-|y-x|^2/2t}$$

If $X_0 = x$ then we call $(X_t)_{t \geq 0}$ a **Brownian motion starting from** x. In this case, condition (i) is equivalent following property : for all $t_1, \cdots, t_n \geq 0$ with $t_1 < \cdots < t_n$ and all $B \in \mathcal{B}(\mathbb{R}^{dn})$

$$\mathbb{P}((X_{t_1}, \cdots, X_{t_n}) \in \mathcal{B}) = \int_B \prod_{i=1}^n p(s_i, x_{i-1}, x_i) dx_i$$

where $t_0 = 0$, $x_0 = x$, $s_i = t_i - t_{i-1}$.

Given independent Brownian motions $(X_t^1)_{t \geq 0}, \cdots, (X_t^d)_{t \geq 0}$ in $\mathbb{R}$ starting from 0 and given $x = (x^1, \cdots, x^d) \in \mathbb{R}^d$, the process $(x + (X_t^1, \cdots, X_t^d))_{t \geq 0}$ is a Brownian motion in $\mathbb{R}^d$ starting from x and we obtain all Brownian motion starting from x in $\mathbb{R}^d$ in this way.

7.2. Wiener's theorem

Brownian motion was established as a mathematical object only after 1920's.

Let $W_d = C([0, \infty), \mathbb{R}^d)$, and $x_t : W_d \to \mathbb{R}^d$, $x_t(w) = w(t)$ be the coordinate functions. We may endow W_d with σ-algebra $\mathcal{W}_d = \sigma(x_t : t \geq 0)$.

Given a continuous process $(X_t)_{t \geq 0}$ in $\mathbb{R}^d$ on Ω, we can define

$$X : \Omega \to W_d, \quad X(\omega)(t) = X_t(\omega)$$

then X is $\mathcal{W}_d$-measurable so X has a law on $(W_d, \mathcal{W}_d)$.

Theorem 7.2.1.) *(Wiener)* For all $d \geq 1$ and $x \in \mathbb{R}^d$, there exist a unique probability measure μ_x on $(W_d, \mathcal{W}_d)$ such that $(x_t)_{t \geq 0}$ is a Brownian motion in $\mathbb{R}^d$ staring from x. In particular, Brownian motion exists.

  proof) Conditions (i) and (ii) determine the finite dimensional distributions of a Brownian motion and hence determine the law of any BM on $(W_d, \mathcal{W}_d)$ (with given starting point - hence such probability measure is unique.

  Suppose we have a measure μ_0 on $(W_1, \mathcal{W}_1)$ such that $(x_t)_{t \geq 0} \sim \text{BM}_0$ in $\mathbb{R}$. For $x \in \mathbb{R}$, $(x + x_t)_{t \geq 0} \sim \text{BM}_x$ so could take μ_x as law of this process. Then for $x = (x_1, \cdots, x_d) \in \mathbb{R}^d$, the measure $\mu_{x_1} \otimes \cdots \otimes \mu_{x_d}$ has required properties. So we only have to work in 1 dimension, starting at 0.

Define $\mathbb{D}_N = \{k2^{-N} : k \in \mathbb{Z}^+\}$ and $\mathbb{D} = \bigcup_{N \geq 0} \mathbb{D}_N$. There exists some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a family $(Y_t : t \in \mathbb{D})$ of independent $N(0,1)$ random variables. First define for $t \in \mathbb{D}_0 = \mathbb{Z}^+$,

$$\xi_t = Y_1 + \cdots + Y_t$$

Then $(\xi_n)_{n \in \mathbb{D}_0}$ is Gaussian and $(\xi_{t+1} - \xi_t : t \in \mathbb{D}_0)$ are independent and has distribution $\sim N(0,1)$. We define recursively $(\xi_t)_{t \in \mathbb{D}_N}$ as follows for $t \in \mathbb{D}_{N+1} \backslash \mathbb{D}_N$:

: set $r = t - 2^{-N-1}$, $s = t + 2^{-N-1} \in \mathbb{D}_N$, set $Z_t = 2^{-\frac{N+2}{2}} Y_t$ and define $\xi_t = \frac{\xi_r + \xi_s}{2} + Z_t$.

We will show by induction that for all $N \geq 0$, $(\xi_{t+2^{-N}} - \xi_t : t \in \mathbb{D}_N)$ are independent, $\sim N(0, 2^{-N})$ random variables

: Suppose true for N. Take $t \in \mathbb{D}_{N+1} - \mathbb{D}_N$ and r, s as above. Then

$$\xi_t - \xi_r = \frac{\xi_s - \xi_r}{2} + Z_t, \quad \xi_s - \xi_t = \frac{\xi_s - \xi_r}{2} - Z_t$$

$$\mathrm{Var}\left(\frac{\xi_s - \xi_r}{2}\right) = \frac{1}{4} 2^{-N}, \quad \mathrm{Var}(Z_t) = 2^{-N-2}$$

so

$$\mathrm{Var}(\xi_t - \xi_r) = \frac{1}{4} 2^{-N} + 2^{-N-2} = 2^{-N-1} = \mathrm{Var}(\xi_s - \xi_r)$$

$$\mathrm{cov}(\xi_t - \xi_r, \xi_s - \xi_t) = 0$$

Also for any interval $(u, v]$ disjoint from $(r, s]$ with $u, v \in \mathbb{D}_{N+1}$,

$$\mathrm{cov}(\xi_s - \xi_r, \xi_v - \xi_u) = \mathrm{cov}(\xi_s - \xi_t, \xi_v - \xi_u) = 0$$

So the induction proceeds.

==

(12th November, Monday)

proof of Wiener's theorem continues) We constructed $(\xi_t : t \in \mathbb{D})$ such that for all $s, t \geq 0$, $\xi_{s+t} - \xi_s \sim N(0, t)$ and is independent of $\sigma(\xi_r : r \leq s, r \in \mathbb{D})$. Choose $p > 2$ and set $C_p = \mathbb{E}(|\xi_1|^p)$. Then $C_p < \infty$ and $\mathbb{E}(|\xi_t - \xi_s|^p) = C_p |t - s|^{p/2}$ so by Kolmogorov's lemma(**Theorem 4.2.1**), there exists a continuous process $(X_t)_{t \geq 0}$ such that $X_t = \xi_t$ a.s. for all $t \in \mathbb{D}$.

For $s \geq 0$, $t > 0$ and for any $A \in \mathcal{F}_s^X$ there exist sequence (s_n) in $[0, \infty)$, (t_n) in $(0, \infty)$ and $A_0 \in \sigma(\xi_r : r \leq s, r \in \mathbb{D})$. s.t. $s_n \to s$, $t_n \to t$ as $n \to \infty$ and $1_A = 1_{A_0}$ a.s. (to check this, consider $\{A \in \mathcal{F}_s : $ this holds $\}$). Then for any $f \in C_b(\mathbb{R})$,

$$\mathbb{E}(f(X_{s_n+t_n} - X_{s_n})1_A) = \mathbb{E}(f(\xi_{s_n+t_n} - \xi_{s_n})1_{A_0})$$

$$= \mathbb{P}(A_0) \int_{\mathbb{R}} p(t_0, 0, y) f(y) dy$$

so letting $n \to \infty$ using bounded convergence theorem gives

$$\mathbb{E}(f(X_{t+s} - X_s)1_A) = \mathbb{P}(A) \int_{\mathbb{R}} p(t, 0, y) f(y) dy$$

so X is a Brownian motion as required.

(End of proof) $\square$

10

7.3. Symmetries of Brownian Motion

Proposition 7.3.1) Let $(X_t)_{t\geq 0}$ be a $\mathrm{BM}_0(\mathbb{R}^d)$ and let $\sigma \in (0,\infty)$ and $U \in O(d)$. Then the following processes are also $\mathrm{BM}_0(\mathbb{R}^d)$.

    (i) **(Scaling property)** $(\sigma X_{\sigma^{-2}t})_{t\geq 0}$,

    (ii) **(Rotation invariance)** $(UX_t)_{t\geq 0}$.

In fact $\mathrm{BM}_0(\mathbb{R}^d)$ is characterized among continuous Gaussian processes by its means and co-variances,

$$\mathbb{E}(X_t) = 0, \quad \mathrm{cov}(X_s^i, X_t^j) = \mathbb{E}(X_s^i X_t^j) = \delta_{ij}(s \wedge t)$$

    proof) Exercise.

7.4. Brownian Motion in a Given Filtration

Suppose given a filtration $(\mathcal{F}_t)_{t\geq 0}$ on $(\Omega, \mathcal{F}, \mathbb{P})$.

- We say that $(X_t)_{t\geq 0}$ is a $(\mathcal{F}_t)_{t\geq 0}$-**BM** if

    (a) X_t is $\mathcal{F}_t$-measurable,

    (b) for all $s, t \geq 0$, the random variable $X_{s+t} - X_s \sim N(0, tI)$ and is independent of $\mathcal{F}_s$,

    (c) for all $\omega \in \Omega$, the map $t \mapsto \left(X_t(\omega) : [0,\infty) \to \mathbb{R}^d\right)$ is continuous.

    This implies that X is a BM in the old sense and any process X which is a BM in the old sense is an $(\mathcal{F}_t^X)_{t\geq 0}$-BM.

Proposition 7.4.1) Let $X = (X_t)_{t\geq 0}$ be a $\mathrm{BM}(\mathbb{R}^d)$ and let F be a bounded measurable function on W_d. Define

$$f(x) = \int F(\omega)\mu_x(d\omega), \quad x \in \mathbb{R}^d$$

Then f is measurable on $\mathbb{R}^d$ and $\mathbb{E}(F(X)|\mathcal{F}_0) = f(X_0)$ a.s. (recall, μ_x is the law of BM_x)

    proof) (Was an exercise.) First consider F of form $1_{A^{(a,\underline{\eta})}}$ where $A^{(a,\underline{\eta})} = \{g \in W_d : g(a) \geq (\eta_1, \cdots, \eta_d)\} \subset W_d$. Then

$$\int A^{(a,\eta)}(\omega)\mu_x(d\omega) = \mathbb{P}(B^{(x)}(a) \geq \eta) = \int_{\eta_1}^\infty \cdots \int_{\eta_d}^\infty \frac{1}{(2\pi)^{d/2}} e^{-y^2/2} d^d y$$

    where $B^{(x)}$ is a Brownian motion starting at x, so this is measurable. Also, we easily see that any finite intersection of sets $A^{(a,\eta)}$ if of form

$$B_{(\eta_1,\cdots,\eta_m)}^{(a_1,\cdots,a_m)} = \{g \in W_d : g(a_1) \geq \eta_1, \cdots, g(a_1 + \cdots + a_m) \geq \eta_m\}$$

with $a_1, \cdots a_m \geq 0$ so

$$\int B_{(\eta_1,\cdots,\eta_m)}^{(a_1,\cdots,a_m)}\mu_x(d\omega) = \int A^{(a_1,\eta_1)}(\omega)\mu_x(d\omega)\prod_{k=2}^m \int A^{(a_k,\eta_k-\eta_{k-1})}(\omega)\mu_0(d\omega)$$

    is measurable. We also see that $\mathbb{E}(F(X)|\mathcal{F}_0) = f(X_0)$ a.s. for each case $F = 1_{A^{(a,\eta)}}$ or $F = 1_{B^{(a_1,\cdots,a_m)(\eta_1,\cdots,\eta_m)}}$.

Now notice that the sets $(B^{(a_1,\cdots,a_m)}_{(\eta_1,\cdots,\eta_m)} : m \geq 1, a_j \geq 0, \eta_j \in \mathbb{R})$ forms a π-system generating the σ-algebra $\mathcal{W}_d$. Since any limit of measurable functions is measurable, and $\mathbb{E}(F_n(X)|\mathcal{F}_0) = f_n(X_0) \; \forall n$, $F_n \nearrow F$ implies $\mathbb{E}(F(X)|\mathcal{F}_0) = f(X_0)$, we see from monotone class theorem that the statement also holds for general bounded measurable functions F.

(End of proof) $\square$

7.5. Martingales of BM

Theorem 7.5.1) Let $(X_t)_{t \geq 0}$ be an $(\mathcal{F}_t)_{t \geq 0}$-BM in $\mathbb{R}^d$ and let $f \in C_b^2(\mathbb{R}^d)$. Define

$$M_t = f(X_t) - f(X_0) - \int_0^t \frac{1}{2}\Delta f(X_s)ds$$

Then $(M_t)_{t \geq 0}$ is an $(\mathcal{F}_t)_{t \geq 0}$-martingale.

There are two ways for doing this. One is to use stochastic calculus, and the other is to do Markovian-way(which would be made clear soon). You will see the first method in Stochastic Calculus course next term, and we are going to prove this in the second way.

proof) It is clear that $(M_t)_{t \geq 0}$ is adapted and integrable. Consider for now the case $X_0 = x \in \mathbb{R}^d$. Set

$$m(x) = \mathbb{E}(M_t) = \int_{W_d} \left[f(\omega(t)) - f(\omega(0)) - \int_0^t \frac{1}{2}\Delta f(\omega(s))ds \right] \mu_x(d\omega)$$

Fix $s \in (0, t]$. We have

$$M_t - M_s = f(X_t) - f(X_s) - \int_s^t \frac{1}{2}\Delta f(X_r)dr$$

so using Fubini,

$$\mathbb{E}(M_t - M_s) = \int_{\mathbb{R}^d} p(t,x,y)f(y)dy - \int p(s,x,y)f(y)dy - \int_s^t \int_{\mathbb{R}^d} p(r,x,y) \times \frac{1}{2}\Delta f(y)dy$$

But since p is a heat kernel, has $\frac{d}{dt}p = \frac{1}{2}\Delta p$ so

$$\int_{\mathbb{R}^d} p(r,x,y) \times \frac{1}{2}\Delta f(y)dy = \int_{\mathbb{R}} \frac{1}{2}\Delta_y p(r,x,y)f(y)dy$$
$$= \int_{\mathbb{R}} \frac{d}{dt}p(r,x,y)f(y)dy$$

and

$$\int_s^t \int_{\mathbb{R}^d} p(r,x,y) \times \frac{1}{2}\Delta f(y)dydr = \int_s^t \int_{\mathbb{R}} \frac{d}{dt}p(r,x,y)f(y)dydr$$
$$= \int_{\mathbb{R}^d} (p(t,x,y) - p(s,x,y))dy$$

so $\mathbb{E}(M_t - M_s) = 0$. But $M_s \to 0$ a.s. as $s \to 0$ so by bounded convergence, $\mathbb{E}(M_s) \to 0$. Hence $m(x) = \mathbb{E}(M_t) = 0$ for all x.

12

Return to the case of general X_0. By **Prop 7.4.1**,

$$\mathbb{E}(M_t|\mathcal{F}_0) = m(X_0) = 0 \quad \text{a.s.}$$

Now for any $s, t \geq 0$,

$$M_{s+t} - M_s = f(X_{s+t}) - f(X_s) - \frac{1}{2}\int_0^s \Delta f(X_{s+r})dr$$

and $(X_{s+t})_{t\geq 0}$ is an $(\mathcal{F}_{s+t})_{t\geq 0}$-BM. So we see that

$$\mathbb{E}(M_{s+t} - M_s|\mathcal{F}_s) = 0 \quad \text{a.s.}$$

as required.

(End of proof) $\square$

We are in fact apply the theorem in the case where the assumption $f \in C_b^2(\mathbb{R}^d)$ does not hold. We can actually relax this condition using almost the same proof, e.g. any f of polynomial growth rate would work.

Exercise : see how much we can relax $C_b^2(\mathbb{R}^d)$.

-Every step other than the integration by parts $\int_{\mathbb{R}^d} p(r,x,y)\times\frac{1}{2}\Delta f(y)dy = \int_{\mathbb{R}} \frac{1}{2}\Delta_y p(r,x,y)f(y)dy$ works without assumption that $f \in C_b^2(\mathbb{R}^d)$. Hence, it is enough to have $f \in W^{2,1}(\mathbb{R}^d)$(the Sobolev space) with sub-exponential growth rate.

==

(14th November, Wednesday)

Examples : Let $(X_t^i)_{t\geq 0}$, $i = 1, \cdots, n$ be a BM.

- $f(x) = x^i$ then $\Delta f = 0$, so is a martingale.

- $f(x) = x^i x^j$ then $\Delta f = 0$, so is a martingale.

- $(X_t^i X_t^j - \delta_j t)_{t\geq 0}$ is a martingale.

- $\exp(\lambda X_t - \frac{1}{2}\lambda^2 t)_{t\geq 0}$ is a martingale.

- For all harmonic functions f bounded, $(f(X_t))_{t\geq 0}$ is a martingale.

7.6. Strong Markov Property

Strong Markov Property of Brownian Motion would be particularly useful.

Theorem 7.6.1) Let $(X_t)_{t\geq 0}$ be an $(\mathcal{F}_t)_{t\geq 0}$-BM and let T be a stopping time. Then conditioned on $\{T < \infty\}$, $(X_{T+t})_{t\geq 0}$ is an $(\mathcal{F}_{T+t})_{t\geq 0}$-BM.

proof) Clearly on $\{T < \infty\}$, $(X_{T<t})_{t\geq 0}$ is continuous on $[0, \infty)$. Also on $\{T < \infty\}$, X_{T+t} is $\mathcal{F}_{T+t}$-measurable. Fix $s \geq 0$, $t > 0$, $f \in C_b(\mathbb{R}^d)$, $m, n \in \mathbb{N}$, $A \in \mathcal{F}_{T+s}$ with $A \subset \{T \leq m\}$. For $k = 0, 1, \cdots 2^n m$, set $t_k = k2^{-n}$ and $A_k = A \cap \{t_k - 2^{-n} < T \leq t_k\}$. Note $A_k \in \mathcal{F}_{t_k+s}$ and $A_k = A \cap \{T_n = t_k\}$ where $T_n = 2^{-n}\lceil 2^n T\rceil$, so

$$\mathbb{E}(f(X_{T_n+s+t})1_{A_k}) = \mathbb{E}(f(X_{t_k+s+t})1_{A_k})$$
$$= \mathbb{E}(P_t f(X_{t_k+s})1_{A_k}) = \mathbb{E}(P_t f(X_{T_n+s})1_{A_k})$$

13

Now sum over k and let $n \to \infty$ using bounded convergence to obtain

$$\mathbb{E}(f(X_{T+s+t})1_A) = \mathbb{E}(P_t f(X_{T+s})1_A)$$

But m and A were arbitrary so this implies, on $\{T < \infty\}$,

$$\mathbb{E}(f(X_{T+t+s})|\mathcal{F}_{T+s}) = P_t f(X_{T+s}) \quad \text{a.s.}$$

as required.

(End of proof) □

7.7. Properties of 1-d BM

Proposition 7.7.1) Let $(X_t)_{t\geq 0}$ be a $BM_0(\mathbb{R})$. Set $T_a = \inf\{t \geq 0 : X_t = a\}$. Then

$$\mathbb{P}(T_a < \infty) = 1 \quad \text{for all } a \in \mathbb{R} \quad and$$

$$\mathbb{P}(T_{-a} \leq T_b) = \frac{b}{a+b} \quad \forall a, b \geq 0 \quad and$$

$$\mathbb{E}(T_a \wedge T_b) = ab$$

Moreover T_a has a density f_a on $[0, \infty)$ given by

$$f_a(t) = \frac{1}{\sqrt{2\pi t^3}} e^{-a^2/2t} \quad t \geq 0$$

Moreover the following holds almost surely.

(a) $X_t/t \to 0$ as $t \to \infty$.

(b) $\inf_{t\geq 0} X_t = -\infty$, $\sup_{t\geq 0} X_t = \infty$.

(c) for all $s \geq 0$, there exist $t, n \geq s$ such that $X_t < 0 < X_n$.

(d) for all $s > 0$ there exist $t, n \in [0, s)$ such that $X_t < 0 < X_n$.

proof) Exercise.

Theorem 7.7.2) Let $X \sim BM_0(\mathbb{R})$. Then the following properties hold almost surely :

(a) for all $\alpha < 1/2$, $(X_t)_{t\geq 0}$ is locally Hölder continuous of exponent α.

(b) for all $\alpha > 1/2$ there is no non-trivial interval on which X is Hölder continuous of exponent α.

proof)

(a) For $s, t \geq 0$ with $s < t$, $X_t - X_s \sim (t-s)^{1/2} X_1$ so for all $p < \infty$,

$$\mathbb{E}(|X_t - X_s|^p) = (t-s)^{p/2} C_p$$

where $C_p = \mathbb{E}(|X_1|^p) < \infty$. Given $\alpha < 1/2$ we can find $p < \infty$ so that $\alpha + \frac{1}{p} < \frac{1}{2}$. Then by Kolmogorov's lemma there exist $K \in L^p(\mathbb{P})$ such that

$$|X_t - X_s| \leq K|t-s|^\alpha \quad \text{for all } s, t \in [0, 1]$$

So by scaling of BM, $(X_t)_{t\geq 0}$ is locally Hölder continuous of exponent α a.s.

Finally consider $\alpha_n \nearrow 1/2$, $\alpha_n < 1/2$ to see $(X_t)_{t\geq 0}$ is Hölder continuous of exponent α a.s. for all $\alpha < 1/2$.

(b) Let $m, n \in \mathbb{N}$ with $m \geq n$. For $s, t \in \mathbb{D}_n$, define

$$[X]_{s,t}^m = \sum_{\tau} (X_{\tau + 2^{-m}} - X_\tau)^2$$

where we sum over all $\tau \in \mathbb{D}_n$ with $\tau \in (s, t]$. The random variables $(X_{\tau + 2^{-m}} - X_\tau)^2$ are independent of mean 2^{-m} and variance 2^{-2m+1} (Here we need $\mathbb{E}(X_1^4) = 3$, $\mathrm{Var}(X_1^2) = 2$). So $\mathbb{E}([X]_{s,t}^m) = 2^m(t-s)2^{-m} = t - s$ and $\mathrm{Var}([X]_{s,t}) = 2^m(t-s)2^{-2m+1}$. So $[X]_{s,t}^m \to t - s$ a.s.

Suppose X is Hölder α on $[s, t]$ then for some constant K,

$$|X_{\tau + 2^{-m}} - X_\tau|^2 \leq K^2 (2^{-m})^{2\alpha}$$

so $[X]_{s,t}^m \leq 2^m(t-s)K^2 2^{-2m\alpha} \to 0$ as $m \to \infty$. So $(X_t)_{t \geq 0}$ is not Hölder α on $[s, t]$, which contradicts with our earlier result.

(End of proof) $\square$

===

(16th November, Friday)

7.8. Recurrence and Transience of Brownian Motion

The following theorem tells about recurrence and transience in different dimensions.

Theorem 7.8.1) Let $(X_t)_{t \geq 0}$ is a $\mathbf{BM}_0(\mathbb{R}^d)$.

(a) For $d = 1$, $\mathbb{P}(\{t \geq 0 : X_t = 0\}$ is unbounded$) = 1$. **(Point recurrent)**

(b) For $d = 2$, and all $\epsilon > 0$, $\mathbb{P}(X_t = 0$ for some $t > 0) = 0$ but $\mathbb{P}(\{t \geq 0 : |X_t| \leq \epsilon\}$ is unbounded$) = 1$. **(Neighbourhood recurrent**, but not recurrent**)**

(c) For $d \geq 3$, $\mathbb{P}(|X_t| \to \infty$ as $t \to \infty) = 1$. **(Transient)**

proof)

(a) Is direct from **Prop 7.7.1 (b)**

(b) Note that $x \mapsto \log|x|$ is harmonic in $\mathbb{R}^2 \backslash \{0\}$. Let $(X_t)_{t \geq 0} \sim \mathrm{BM}_x(\mathbb{R}^2)$. For $\epsilon > 0$, set
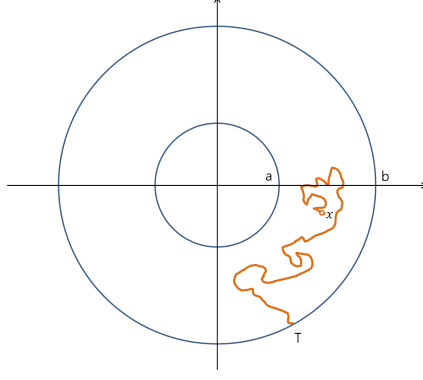
$$p_0(x) = \mathbb{P}(X_t = 0 \text{ for some } t > 0), \quad p_\epsilon(x) = \mathbb{P}(|X_t| \leq \epsilon \text{ for some } t > 0)$$

Consider the case $|x| = 1$. For $a < 1 < b$, there is a function $f \in C_b^2(\mathbb{R}^2)$ such that $f(x) = \log|x|$ for $|x| \in [a, b]$. Also consider the martingale (a martingale by **Theorem 7.5.1**)

$$M_t = f(X_t) - f(X_0) - \int_0^t \frac{1}{2} \Delta f(X_s) ds$$

and the stopping time

$$T = \inf\{t \geq 0 : |X_t| \in \{a, b\}\}$$

15

Note that $\mathbb{P}(T < \infty) = 1$. Also for $t \leq T$, $M_t = \log|X_t|$ so M is bounded up to T. So optional stopping theorem applies to show

$$0 = \mathbb{E}(M_0) = \mathbb{E}(M_T) = \mathbb{E}(\log|X_t|) = p\log a + (1-b)\log b$$

where $p = p(a,b) = \mathbb{P}(|X_T| = a)$. Consider the limit $a \to 0$, then $\log a \to -\infty$, so $p \to 0$. Hence for $|x| = 1$, $p_0(x) = \lim_{b\to\infty}\lim_{a\to 0} p(a,b) = 0$. By scaling, we see $p_0(x) = 0$ for all $x \neq 0$.

In the case $x = 0$, for $n \in \mathbb{N}$, by Markov property,

$$\mathbb{P}(|X_t| = 0 \text{ for some } t \geq \frac{1}{n}) = \int_{\mathbb{R}^2} p(\frac{1}{n}, 0, y)p_0(y)dy = 0$$

Let $n \to \infty$ to see $p_0(0) = 0$.

For $|x| = 1$, let $b \to \infty$ with $a = \epsilon$ fixed. Since $\log b \to \infty$, $p(a,b) \to 1$ so

$$p_\epsilon(x) = \lim_{b\to\infty} p(\epsilon, b) = 1.$$

By scaling $p_\epsilon(x) = 1$ for all $x \neq 0$.

If $x = 0$, for $n \in \mathbb{N}$, by Markov property

$$\mathbb{P}(|X_t| \leq \epsilon \text{ for some } t \geq n) = \int_{\mathbb{R}^2} p(n, x, y)p_\epsilon(y)dy = 1$$

Hence $\mathbb{P}(\{t \geq 0, |X_t| \leq \epsilon\} \text{ is unbounded }) = 1$.

(c) It will suffice to consider case $d = 3$ and to show for all $N \in \mathbb{N}$ that

$$\mathbb{P}(\{t \geq 0 : |X_t| \leq N\} \text{ is unbounded}) = 0$$

We repeat the argument from (b) with $1/|x|$ in place of $\log(|x|)$ - in $d = 3$, $1/|x|$ is the harmonic function on $\mathbb{R}^3\backslash\{0\}$ - to obtain

$$\frac{p}{a} + \frac{1-p}{b} = 1$$

for $p = \mathbb{P}(|X_T| = a)$ and $(X_T)_{t\geq 0} \sim \text{BM}_x(\mathbb{R}^3)$, $|x| = 1$. Let $b \to \infty$ to see

$$\lim_{b\to\infty} p(a,b) = \mathbb{P}(|X_t| \leq a \text{ for some } t \geq 0) = a$$

By scaling, for $|x| = N + 1$,

$$\mathbb{P}(|X_t| \leq N \text{ for some } t \geq 0) = \frac{N}{N+1}$$

Set $T_0 = 0$ and for $k \geq 1$,

$$S_k = \inf\{t \geq T_{k-1} : |X_t| = N + 1\}, \quad T_k = \inf\{t \geq S_k : |X_t| = N\}$$

Then

$$\mathbb{P}(T_k < \infty) \leq \frac{N}{N+1}\mathbb{P}(T_{k-1} < \infty) \leq \left(\frac{N}{N+1}\right)^k \to 0$$

by strong Markov property and induction. Set $K = \sup\{k : T_k < \infty\}$. Then $K < \infty$ a.s. Now for all $k \geq 1$, on $\{T_{k-1} < \infty\}$, has $S_k < \infty$ a.s. Hence $S_{K+1} < \infty$ a.s. But on $\{S_{K+1} \leq \infty\}$, has $|X_t| > N$ for all $t \geq S_{K+1}$ hence $\mathbb{P}(\{t \geq 0 : |X_t| \leq N\}$ is unbounded $) = 0$.

$$\text{(End of proof)} \ \square$$

7.9. Brownian Motion and the Dirichlet Problem

Let D be a connected open set in $\mathbb{R}^d$. Write ∂D for the boundary. Let $c : D \to [0, \infty)$, $f : \partial D \to [0, \infty)$ be measurable functions. For $x \in \overline{D}$, let $(X_t)_{t \geq 0}$ be a Brownian motion starting at x.

Define the **expected total cost** by

$$\phi(x) = \mathbb{E}\left(\int_0^T c(X_t)dt + f(X_T)1_{t < \infty}\right)$$

where $T = \inf\{t \geq 0 : X_t \notin D\}$. By a

By a **solution of the Dirichlet problem (in D with data c and f)**, we mean a function $\psi \in C(\overline{D}) \cap C^2(D)$ such that

$$\begin{cases} -\frac{1}{2}\Delta\psi = c & \text{in } D \\ \psi = f & \text{on } \partial D \end{cases}$$

Accordingly, say ψ is a **super-solution** if $-\frac{1}{2}\Delta\psi \geq c$, $\psi \geq f$, and call ψ a **sub-solution** with inverse inequalities.

Suppose $\psi \in C_b^2(\mathbb{R}^d)$ restrict to a solution of Dirichlet problem on D. Consider the martingale (as in **Theorem 7.5.1**)

$$M_t = \psi(X_t) - \psi(X_0) - \int_0^t \frac{1}{2}\Delta\psi(X_s)ds$$

Suppose we could justify optional stopping theorem at T and suppose $T < \infty$ a.s. Then

$$0 = \mathbb{E}(M_0) = \mathbb{E}(M_T) = \mathbb{E}(\phi(X_T)) - \psi(x) - \mathbb{E}\left(\int_0^T \frac{1}{2}\Delta\psi(X_s)ds\right)$$

so

$$\psi(x) = \mathbb{E}(f(X_T)) + \mathbb{E}\left(\int_0^T c(X_s)ds\right) = \phi(x)$$

We will spend time in exploring the condition in which this holds.

==

(19th November, Monday)

Settings from the last lecture : D a connected open set $\subset \mathbb{R}^d$, $c : D \to [0, \infty)$, $f : \partial D \to [0, \infty)$ measurable functions,

$$\phi(x) = \mathbb{E}\Big[\int_0^T c(X_t)dt + f(X_T)1_{T<\infty} \Big]$$

$(X_t)_{t\geq 0} \sim \mathrm{BM}_x(\mathbb{R}^d)$, $T = \inf\{t \geq 0 : X_t \notin D\}$, times spent in $B \subset D$ before leaving.

Theorem 7.9.2 (a)) Let ψ be a *non-negative super-solution* of the Dirichlet Problem(DP), $\psi \in C(\overline{D}) \cap C^2(D)$,

$$\begin{cases} -\frac{1}{2}\Delta\psi \geq c & \text{in } D \\ \psi \geq f & \text{on } \partial D \end{cases}$$

Then $\psi \geq \phi$.

> **proof)** Clear that $\psi \geq \phi$ on ∂D. Fix $x \in D$ and $N \geq 1$ such that $x \in D_N = \{y \in D : |y| < N, \mathrm{dist}(y, \partial D) > 1/N\}$. There exists $g \in C_b^2(\mathbb{R}^d)$ such that $g = \psi$ on D_N. Consider
>
> $$M_t = g(X_t) - g(X_0) - \int_0^t \frac{1}{2}\Delta g(X_s)ds$$
>
> and $T_N = \inf\{t \geq 0 : X_t \notin D_N\}$. Then M is a martingale. By the optional stopping theorem, has
>
> $$0 = \mathbb{E}(M_0) = \mathbb{E}(M_{T_N \wedge t}) = \mathbb{E}(\psi(X_{T_N \wedge t})) - \psi(x) + \mathbb{E}\Big[\int_0^{T_N \wedge t} \Big(-\frac{1}{2}\Delta\psi\Big)(X_s)ds \Big] \quad \cdots\cdots\cdots (\star)$$
>
> Consider the limit $N \to \infty$ and then $t \to \infty$. Then $T_N \wedge t \nearrow T$, hence
>
> $$\mathbb{E}\int_0^{T_N \wedge t} \Big(-\frac{1}{2}\Delta\psi\Big)(X_s)ds \geq \mathbb{E}\int_0^{T_N \wedge t} c(X_s)ds \to \mathbb{E}\int_0^T c(X_s)ds$$
>
> by monotone convergence. On $\{T < \infty\}$,
>
> $$\psi(X_{T_N \wedge t}) \to \psi(X_T) \quad \text{a.s.}$$
>
> so sine $\psi \geq 0$, by Fatou's lemma,
>
> $$\liminf \mathbb{E}(\psi(X_{T_N \wedge t})) \geq \mathbb{E}(\psi(X_T)1_{T<\infty}) \geq \mathbb{E}(f(X_T)1_{T<\infty})$$
>
> So taking $\liminf$ in $(\star)$, we see $\psi \geq \phi$.
>
> $$\textit{(End of proof)} \ \square$$

Theorem 7.9.2 (b)) Let ψ be a *non-negative bounded solution* of the Dirichlet Problem, $\psi \in C(\overline{D}) \cap C^2(D)$

$$\begin{cases} -\frac{1}{2}\Delta\psi = c & \text{in } D \\ \psi = f & \text{on } \partial D \end{cases}$$

and suppose that $\mathbb{E}(\psi(X_t)1_{t<T}) \to 0$ as $t \to \infty (\cdots (\dagger))$. Then $\psi = \phi$.
Remark about the condition (†) : Since we assume ψ bounded and

$$|\mathbb{E}(\psi|X_t)1_{t<T}| \leq \|\psi\|_\infty \mathbb{P}(t < T)$$

We see condition (†) holds whenever $T < \infty$ a.s. for all starting points x. e.g. we're OK if $(d = 1 \text{ and } \partial D \neq \varnothing)$ or $(d = 2 \text{ and } \mathbb{R}^2 \backslash D \text{ contains an open ball})$ or $(d \geq 3 \text{ and } D \text{ bounded})$.

proof) Clear that $\psi = \phi$ on ∂D. Fix $x \in D$ and $N \geq 1$ such that $x \in D_N = \{y \in D : |y| < N, \text{dist}(y, \partial D) > 1/N\}$. There exists $g \in C_b^2(\mathbb{R}^d)$ such that $g = \psi$ on D_N. Consider

$$M_t = g(X_t) - g(X_0) - \int_0^t \frac{1}{2}\Delta g(X_s)ds$$

and $T_N = \inf\{t \geq 0 : X_t \notin D_N\}$. Then M is a martingale. By the optional stopping theorem, has

$$0 = \mathbb{E}(M_0) = \mathbb{E}(M_{T_N \wedge t}) = \mathbb{E}(\psi(X_{T_N \wedge t})) - \psi(x) + \mathbb{E}\left[\int_0^{T_N \wedge t} \left(-\frac{1}{2}\Delta\psi\right)(X_s)ds\right] \quad \cdots\cdots\cdots(\star)$$

Consider the limit $N \to \infty$ and then $t \to \infty$. Then $T_N \wedge t \nearrow T$, hence

$$\mathbb{E}\int_0^{T_N \wedge t} \left(-\frac{1}{2}\Delta\psi\right)(X_s)ds = \mathbb{E}\int_0^{T_N \wedge t} c(X_s)ds \to \mathbb{E}\int_0^T c(X_s)ds$$

by monotone convergence. Now

$$\mathbb{E}(\psi(X_{T_N \wedge t})) = \mathbb{E}(\psi(X_{T_N})1_{T_N \leq t}) + \mathbb{E}(\psi(X_t)1_{t < T_N})$$

We have, sine $\psi \in C(\overline{D})$,

$$\psi(X_{T_N})1_{T_N \leq t} \to \psi(X_T)1_{T < \infty} \quad \text{a.s.}$$

and since ψ is bounded, by dominated convergence,

$$\mathbb{E}\left[\psi(X_{T_N})1_{T_N \leq t}\right] \to \mathbb{E}(\psi(X_T)1_{T < \infty}) = \mathbb{E}(f(X_T)1_{T < \infty})$$

Also by the assumption

$$\lim_{N \to \infty} \mathbb{E}(\psi(X_t)1_{t < T_N}) = \mathbb{E}(\psi(X_t)1_{t < T}) \to 0 \quad \text{as } t \to \infty$$

Take limit in $(\star)$ to see $\psi = \phi$. *(End of proof)* $\square$

This theorem tells us about uniqueness of solution of the Dirichlet problem, but it does not say anything about existence of the solution. We are next going to examine this problem.

Theorem 7.9.2 (c).(I)) Assume $d \geq 3$ and $D = \mathbb{R}^d$ and c has compact support and $c \in C^2(D)$. Then $\phi \in C^2(\mathbb{R}^d)$ and $-\frac{1}{2}\Delta\phi = c$.

proof) Suppose g is continuous and of compact support on $\mathbb{R}^d$. Note

$$P_t g(x) = \mathbb{E}g(x + X_t) = \int_{\mathbb{R}^d} (2\pi t)^{-d/2} e^{-|x-y|^2/2t} g(y)dy$$

where $(X_t)_{t \geq 0} \sim \text{BM}(\mathbb{R}^d)$ so

$$\begin{cases} \|P_t g\|_\infty \leq \|g\|_\infty & \text{and also} \\ \|P_t g\|_\infty \leq (2\pi t)^{-d/2}\text{vol}(\text{supp}(g))\|g\|_\infty \end{cases} \quad \cdots\cdots\cdots(\diamond)$$

So splitting the integral at 1 gives

$$\mathbb{E}\int_0^t g(x + X_t)dt = \int_0^\infty P_t g(x)dt \leq (1 + \text{vol}(\text{supp}(g)))\|g\|_\infty$$

Hence also for $\epsilon > 0$,

$$\mathbb{E} \int_0^\infty \sup_{|x-y|<\epsilon} |g(y + X_t)| dt < \infty$$

(note, $g_\epsilon(x) = \sup_{|y-x|<\epsilon} |g(y)|$ is also continuous of compact support.) Now we can apply this to and its derivative to justify differentiating

$$\phi(x) = \mathbb{E} \int_0^\infty c(x + X_t) dt$$

under the integrals to obtain $\phi \in C^2(D)$ and

$$\Delta\phi(x) = \mathbb{E} \int_0^\infty \Delta c(x + X_t) dt = \left(\int_0^s + \int_s^t + \int_t^\infty \right) \mathbb{E}(\Delta c(x + X_r)) dr$$

for $0 < s < t$. Consider the limits $s \to 0$ and $t \to \infty$. First and the third integrals tend to 0 using the estimates above. Now

$$\frac{1}{2} \int_s^t \mathbb{E}(\Delta c(x + X_r)) dr = \frac{1}{2} \int_s^t \int_{\mathbb{R}^d} \Delta c(y) p(r, x, y) dy dr$$

$$= \frac{1}{2} \int_s^t \int_{\mathbb{R}^d} c(y) \Delta p(t, x, y) dy dr$$

$$= \int_s^t \int_{\mathbb{R}^d} c(y) \dot{p}(r, x, y) dy dr$$

$$= \int_{\mathbb{R}^d} c(y)(p(t, x, y) - p(s, x, y)) dy \quad \text{(FTC)}$$

But $\int c(y) p(t, x, y) dy \to 0$ as $t \to \infty$ and $\int c(y) p(s, x, y) dy = \mathbb{E}(c(x + X_s)) \to c(x)$ as $s \to 0$. Hence $-\frac{1}{2}\Delta\psi = c$ as required.

(End of proof) $\square$

===

(21st November, Wednesday)

(part iii drop-in hours tomorrow 10:30 - 12:30).
(Bring a few different colored pens to class on Friday, if possible)
(Office hours on Monday, time and place TBD)

Theorem 7.9.4) Let D be a connected open set $\subset \mathbb{R}^d$ and let ϕ be a *non-negative* measurable function on D. Suppose

$$\phi(x) = \int_{S(x,p)} \phi(y) \sigma_{x,p}(dy)$$

whenever $B(x, p) \subset D$ where $\sigma_{x,p}$ is the uniform distribution on $S(x, p) = \{y : |x - y| = p\}$. Then *either* $\phi(x) = \infty$ for all x *or* $\phi \in C^\infty(D)$ with $\Delta\phi = 0$.

This would not be examinable, and will not be proved in the lecture - see online notes for proof.

Theorem 7.9.5) *(Blumenthal's zero-one law)* Let $(X_t)_{t\geq 0} \sim \mathrm{BM}_x(\mathbb{R}^d)$. Then for all $A \in \mathcal{F}_{0+}^X = \bigcap_{t>0} \mathcal{F}_t^X$, has $\mathbb{P}(A) \in \{0, 1\}$.

proof) Consider the π-system $\mathcal{A} = \bigcup_{s>0} \sigma(X_t - X_s : t \geq s)$. Then for all $A_0 \in \mathcal{F}^X_{0+}$ and all $A \in \mathcal{A}$, we have $\mathbb{P}(A_0 \cap A) = \mathbb{P}(A_0)\mathbb{P}(A)$. By Dynkin's lemma, this extend to all $A \in \sigma(\mathcal{A})$ (check details).

Now for all $s > 0, t \geq s$, $X_t - X_s$ is $\sigma(\mathcal{A})$-measurable. But $X_s \to 0$ as $s \to 0$, so X_t is $\sigma(\mathcal{A})$-measurable. So $\mathcal{F}^X_\infty = \sigma(\mathcal{A}) \supset \mathcal{F}^X_{0+}$. So for $A \in \mathcal{F}^X_{0+}$,

$$\mathbb{P}(A) = \mathbb{P}(A \cap A) = \mathbb{P}(A)\mathbb{P}(A)$$

and therefore $\mathbb{P}(A) \in \{0, 1\}$.

(End of proof) $\square$

Proposition 7.9.6) *(Brownina motion enters all cones immediately)* Let A be an open subset of S^{d-1} and let $\epsilon > 0$. Set C a cone,

$$C = \{ty : y \in A, t \in (0, \epsilon)\}$$

Let $X \sim \mathrm{BM}_0(\mathbb{R}^d)$, and define $T_C = \inf\{t \geq 0 : X_t \in T_C\}$. Then $\mathbb{P}(T_C = 0) = 1$.

See example sheet.

Some definition (**External cone condition, ECC)**) : Let $D \subset \mathbb{R}^d$ be open and connected. Then D satisfies ECC if for all $x \in \partial D$, there exist A open $\subset S^{d-1}$, $\epsilon > 0$ such that

$$\{x + ty : t \in (0, \epsilon), y \in A\} \cap D = \phi$$

This is weaker than the Lipschitz boundary condition.

Theorem 7.9.2(c)) Let D be a connected open set $\subset \mathbb{R}^d$. Let $c \in C^2(\mathbb{R}^d)$, $f \in C(\partial D)$. Set

$$\phi(x) = \mathbb{E}\Big(\int_0^T c(X_t)dt + f(X_T)1_{T<\infty} \Big) \quad \forall x \in \overline{D}$$

where $X \sim \mathrm{BM}_0(\mathbb{R}^d)$ and $T = \inf\{t \geq 0 : X_t \notin D\}$. Assume D satisfies the *external cone condition(ECC)* and that ϕ is locally bounded. Then $\phi \in C^2(D) \cap C(\overline{D})$ and $-\frac{1}{2}\Delta\phi = c$ in D and $\phi = f$ on ∂D. i.e. ϕ solves the Dirichlet condition.

This is significant - the proof using ECC is hard to achieve just using usual theories of partial differential equations.(recall, we usually prove theorems under Lipschitz boundary condition in PDEs). Probability wins!

Lemma 7.9.3) Let D_0 be a *bounded subset* of D. Define $T_0 = \inf\{t \geq 0 : X_t \notin D_0\}$, $x \in \overline{D}$. Then

$$\phi(x) = \mathbb{E}\Big(\int_0^{T_0} c(X_t)dt + \phi(X_{T_0}) \Big)$$

proof) Set $\tilde{X}_t = X_{T_0+t}$, $\tilde{\mathcal{F}}_t = \mathcal{F}_{T_0+t}$, and $\tilde{T} = \inf\{t \geq 0 : \tilde{X}_t \notin D\}$. By *Strong Markov property*, $(\tilde{X}_t)$ is an $(\tilde{\mathcal{F}}_t)$-BM. Now

$$\int_0^T c(X_t)dt + f(X_T)1_{T<\infty} = \int_0^{T_0} c(X_t)dt + \int_0^{\tilde{T}} c(X_t)dt + f(\tilde{X}_{\tilde{T}})1_{\tilde{T}<\infty} \quad \cdots\cdots\cdots (\dagger)$$

Here we used that $\tilde{T} < \infty$ *iff* $T < \infty$ and $X_T = \tilde{X}_{\tilde{T}}$. By **Prop 7.4.1**,

$$\mathbb{E}\left(\int_0^{\tilde{T}} c(X_t)dt + f(\tilde{X}_{\tilde{T}})1_{\tilde{T}<\infty}|\mathcal{F}_{T_0}\right) = \phi(X_{T_0}) \quad \text{a.s.}$$

So we obtain result on taking $\mathbb{E}$ in ($\dagger$)

(End of proof) $\square$

Recall, we proved **Theorem 7.9.2(c)** in a weaker setting earlier. Here, we prove in a slightly more general setting.

proof of Theorem 7.9.2(c) II. We show in the case $c = 0$ that $\phi \in C^\infty(D)$ with $\Delta\phi = 0$ -then follows from a simple argument using symmetry of a sphere and **Theorem 7.9.4**.

See online notes.

proof of Theorem 7.9.2(c) III. We show for all $y \in \partial D$ that

$$\lim_{x \to y, x \in \overline{D}} \phi(x) = \phi(y)$$

Fix $y \in \partial D$ and choose U bounded open set $\subset \mathbb{R}^d$ with $u \in U$. Set $D_0 = D \cap U$. By *exterior cone condition*, there exist $A \in S^{d-1}$, $\epsilon > 0$ such that $\{y + tz : t \in (0,\epsilon), z \in A\} \cap D = \phi$. Let $(X_t)_{t \geq 0}$ be a $BM_0(\mathbb{R}^d)$. Set $T_0(x) = \inf\{t \geq 0 : x + X_t \notin D\}$. By **Proposition 7.9.6**, for $T_C = \inf\{t \geq 0 : y + X_t \in C\}$ we have $T_C = 0$ a.s. Now on the event $\{T_C = 0\}$, we have $T_0(x) \to 0$ as $x \to y$(x staying in $\overline{D}$) so $X_{T_0(x)} + x \in \partial D$ eventually as $x \to y$ and $X_{T_0(x)} + x \to y$ as $x \to y$. Hence

$$\phi(X_{T_0(x)} + x) = f(X_{T_0(x)} + x) \to f(y) \quad \text{as } x \to y$$

Also by **Prop 7.7.1**, $\mathbb{E}(\sup_{x \in D_0} T_0(x)) < \infty$. Now

$$\phi(x) = \mathbb{E}\left(\int_0^{T_0(x)} c(x + X_t)dt + \phi(X_{T_0(x)} + x)\right)$$

Let $x \to y$, then using dominated convergence to see $\phi(x) \to f(y)$ as $x \to y$.

(End of proof) $\square$

==

(23rd November, Friday)

The following is the last part of proof of **Theorem 7.9.2(c)**

proof of Theorem 7.9.2(c) IV. It will suffice to consider the case where $d \geq 3$, or set $\tilde{D} = D \times \mathbb{R}^2$ and extend c as a constant in new directions.

Consider for now the case D bounded. There exists $\tilde{c} \in C^2(\mathbb{R}^d)$ of compact support with $\tilde{c} = c$ in D. Set

$$\phi_0(x) = \mathbb{E}\int_0^\infty \tilde{c}(X_t)dt$$

By **part I**, $\phi_0 \in C^2(\mathbb{R}^d)$ with $-\frac{1}{2}\Delta\phi_0 = \tilde{c}$. Now

$$\phi_0(x) = \mathbb{E}\Big(\int_0^T \tilde{c}(X_t)dt + \phi_0(X_T)\Big) = \phi(x) + \phi_1(x)$$

(by **Lemma 7.9.3**) where $\phi_1(x) = \mathbb{E}(\phi_0(X_T))$. By **part II**, $\phi_1 \in C^\infty(D)$ and $\Delta\phi_1 = 0$. So $\phi \in C^2(D)$ and $-\frac{1}{2}\Delta\phi = c$.

Turn to the case D unbounded. Choose any bounded open $D_0 \subset D$. Then

$$\phi(x) = \phi_0(x) + \phi_1(x)$$

(new ϕ_0 and ϕ_1) where ϕ_0 and ϕ_1 are defined by

$$\phi_0(x) = \mathbb{E}\Big(\int_0^{T_0} c(X_t)dt\Big), \quad \phi_1(x) = \mathbb{E}(\phi(X_{T_0}))$$

where $T_0 = \inf\{t \geq 0 : X_t \notin D_0\}$. We know $\phi_0 \in C^2(D)$ and $-\frac{1}{2}\Delta\phi_0 = c$ and since ϕ is locally bounded, $\phi_1 \in C^\infty(D_0)$ with $\Delta\phi_1 = 0$. Hence $\phi \in C^2(D_0)$ and $-\frac{1}{2}\Delta\phi = c$.

(End of proof) $\square$

Plan for the rest of the course : we will not cover all the materials(e.g. Donsker's invariance principle or Skorohod embedding for random walks) - do not have enough time. But will discuss about Poisson random measures.

8. Poisson Random Measures

Our ambition is to classify all stationary independent increments - this is what Lévy-Khinchin theorem tells us. To prove this, we need some works on Poisson random measures.

8.1. Construction

For $\lambda \in (0, \infty)$, a random variable X with values in $\mathbb{Z}^+ \cup \{\infty\}$ has distribution $P(\lambda)$ if

$$\mathbb{P}(X = x) = e^{-\lambda}\frac{\lambda^x}{x!}, \quad x \in \mathbb{Z}^+$$

Also conventionally write $X \sim P(0)$ if $X \equiv 0$, and $P \sim P(\infty)$ if $X \equiv \infty$.

Proposition 8.1.1.) *(Addition property)* Let $(N_k : k \in \mathbb{N})$ be a sequence of independent random variables with $N_k \sim P(\lambda_k)$ for all k. Then

$$\sum_k N_k \sim P\Big(\sum_k \lambda_k\Big)$$

Proposition 8.1.2.) *(Splitting property)* Suppose $N \sim P(\lambda)$. Let $(Y_n : n \in \mathbb{N})$ be a sequence of i.i.d. random variables in $\mathbb{N}$. Set $N_k = \sum_{n=1}^N 1_{Y_n=k}$ for $k \geq 1$. Then $(N_k : k \in \mathbb{N})$ is a sequence of independent $P(\lambda_k)$ random variables with $\lambda_k = \lambda\mathbb{P}(Y_1) = k$.

The proofs are just elementary calculations, but please do them!!!

Let $(E, \mathcal{E}, \mu)$ be a σ-*finite* measure space.

- A **Poisson random measure** of intensity μ is a map $M : \Omega \times \mathcal{E} \to \mathbb{Z}^+ \cup \{\infty\}$ (for $A \in \mathcal{E}$, think of $M(A)$ as $M(\{\omega\} \times A)$, a random variable) such that for any disjoint sequence $(A_k : k \in \mathbb{N})$ in $\mathcal{E}$,

  (i) $M(\cup_k A_k) = \sum_k M(A_k)$.

  (ii) $M(A_k)$ is a $P(\mu(A_k))$ random variable for all k.

  (iii) The random variables $M(A_k)$, $k \in \mathbb{N}$ are independent.

- We can construct a PRM(μ) as follows.

  First in case $\mu(E) < \infty$, take $N \sim P(\lambda)$, $\lambda = \mu(E)$ and take independent random variables $(Y_n : n \in \mathbb{N})$ independent of N with distribution $\frac{1}{\lambda}\mu$. Define $M(A) = \sum_{n=1}^{N} 1_{Y_n \in A}$ Then $M \sim \text{PRM}(\mu)$, i.e. $M(A) \sim P(\mu(A))$. (**an exercise**)

  In general, if $\mu(E) = \infty$, as E is σ-finite, there exist disjoint $E_n \in \mathcal{E}$ with $\cup_n E_n = E$ such that $\mu(E_n) < \infty$ for all n. Using preceding construction, we obtain $(M_n)_n$ independent PRM($\mu 1_{E_n}$)'s. Then set $M = \sum_n M_n$. Then $M \sim \text{PRM}(\mu)$. (**also an exercise**)

- The uniqueness of such map also can be shown - refer to the online lecture notes.

8.2. Integrals with respect to Poisson Random Measures

Theorem 8.2.1) Let M be a PRM(μ) and let g be a measurable function on $(E, \mathcal{E})$. Assume $\mu(E) < \infty$. Define

$$M(g) = \begin{cases} \int_E g(y) M(dy) & \text{if } M(E) < \infty \\ 0 & \text{o/w} \end{cases}$$

Then $M(g)$ is a well-defined random variable and

$$\mathbb{E}(e^{inM(g)}) = \exp\left[\int_E (e^{ing(y)} - 1)\mu(dy) \right]$$

and if $g \in L^1(\mu)$ then $M(g) \in L^1(\mathbb{P})$ with $\mathbb{E}(M(g)) = \mu(g)$ and $\text{Var}(M(g)) = \mu(g^2)$.

  proof) Let's see these all true in the case $M(A) = \sum_{n=1}^{N} 1_{Y_n \in A}$. Then $M(g) = \sum_{n=1}^{N} g(Y_n)$ and this is clearly a well-defined measurable function. We have

$$\mathbb{E}(e^{inM(g)} | N = n) = \mathbb{E}(e^{in \sum_{k=1}^{n} g(Y_k)}) = \prod_{k=1}^{n} \mathbb{E}(e^{ing(Y_k)}) = \left(\int_E e^{ing(y)} \frac{\mu(dy)}{\lambda} \right)^n$$

So

$$\mathbb{E}(e^{inM(g)}) = \sum_{n=0}^{\infty} e^{-\lambda} \frac{\lambda^n}{n!} \left(\int_E e^{ing(y)} \frac{\mu(dy)}{\lambda} \right)^n$$

$$= e^{-\lambda} \exp\left(\int_E e^{ing(y)} \mu(dy) \right)$$

so we obtain the average and variance of $M(g)$ from the moment generating function.

  For the general case, since $(E, \mathcal{E}, \mu)$ is σ-finite, let $E = \cup_n E_n$ where $E_n \in \mathcal{E}$, disjoint and $\mu(E_n) < \infty$. Let M_n be the Poisson random measure on each E_n. Then

$1_{M<\infty} \sum_{n=1}^{\infty} M_n \to M 1_{M<\infty}$. The statement holds for each M_n by above, so by dominated convergence(dominated by 1),

$$\mathbb{E}(e^{inM(g)}) = \sum_{n=0}^{\infty} e^{-\lambda} \frac{\lambda^n}{n!} \left(\int_E e^{ing(y)} \frac{\mu(dy)}{\lambda} \right)^n$$
$$= e^{-\lambda} \exp \left(\int_E e^{ing(y)} \mu(dy) \right)$$

also holds.

<div align="right">(End of proof) □</div>

25