C is closed if $x \in C$ and $\rho_{xy} > 0 \Rightarrow y \in C$. (Note if C is closed and $x \in C$, then $\mathbb{P}_x(X_n \in C) = 1 \ \forall n$.) D is irreducible if $x, y \in D$ implies $\rho_{xy} > 0$ ($\rho_{yx} > 0$).

<u>Theorem:</u> Let C be a finite closed set. Then, C contains a recurrent set. If C is irreducible, then all states in C are recurrent.

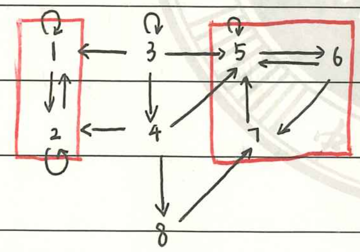
<u>Pf:</u> In view of the previous theorem, we only need to show the first claim.

If otherwise, $\forall y \in C$, $\rho_{yy} < 1$ and $\mathbb{E}_x N(y) = \frac{\rho_{xy}}{1-\rho_{yy}}$. But,

$$\infty > \sum_{y \in C} \mathbb{E}_x N(y) = \sum_{y \in C} \sum_{n=1}^{\infty} p^n(x,y) = \sum_{n=1}^{\infty} \underbrace{\sum_{y \in C} p^n(x,y) =}_{C \text{ is closed}} \sum_{n=1}^{\infty} 1 .$$

$\square$

<u>Example:</u>



	1	2	3	4	5	6	7	8
1	0.4	0.6	0	0	0	0	0	0
2	0.2	0.8	0	0	0	0	0	0
3	0.1	0	0.2	0.3	0.4	0	0	0
4	0	0.6	0	0	0.2	0	0	0.2
5	0	0	0	0	0.3	0.7	0	0
6	0	0	0	0	0.2	0	0.8	0
7	0	0	0	0	0	0	1	0
8	0	0	0	0	0	0	0.1	0.9

(i) $\rho_{31} > 0$ and $\rho_{13} = 0$, so 3 must be transient. Similarly, $\rho_{42} > 0$ and $\rho_{24} = 0$, so 4 must be transient; $\rho_{87} > 0$ and $\rho_{78} = 0$, so 8 is transient.

(ii) $\{1, 2\}$ and $\{5, 6, 7\}$ are irreducible closed sets, so they are recurrent.

Memo No. ______
Date / /

This reasoning can be used to classify transient & recurrent states when S is finite.

(i) There is a y with $\rho_{xy}>0$ and $\rho_{yx}=0$, then x must be ~~recurrent~~ transient.

(ii) $\rho_{xy}>0$ implies $\rho_{yx}>0$. Then, let $C_x=\{y: \rho_{xy}>0\}$. We claim that C is an irreducible closed set, so all states in C_x is recurrent.

[If $y, z\in C_x$, then $\rho_{yz}\geq \rho_{yx}\rho_{xz}>0$. Hence, C_x is irreducible. If $\rho_{yw}>0$ for $y\in C_x$, then $\rho_{xw}\geq \rho_{xy}\rho_{yw}>0$, so $w\in C_x$. Hence, C_x is closed.]

Theorem (Decomposition theorem) Let $R=\{x: \rho_{xx}=1\}$ be the set of recurrent states of a Markov chain. Then, $R=\bigcup_i R_i$, where each R_i is closed and irreducible.

Rmk: For the study of a recurrent state x, we can consider a single irreducible closed set containing x.

Pf: $\forall x\in R$, let $C_x=\{y: \rho_{xy}>0\}$. We have shown that C_x is a irreducible closed set.

If $C_y\cap C_x\neq\phi$, then we have $C_x=C_y$. (Let $z\in C_x\cap C_y$, $\forall y'\in C_y$, then $\rho_{xy'}\geq\rho_{xz}\rho_{zy'}>0 \Rightarrow y'\in C_x$

$\Rightarrow C_y\subseteq C_x$. Similarly, $C_x\subseteq C_y$.)

Hence, we have either $C_x\cap C_y=\phi$ or $C_x=C_y$. This decomposes R into disjoint unions of equivalent classes. □

Example: (Branching process) Suppose the probability of no children is positive (i.e., $p_0>0$). Then, all states $k\geq 1$ are transient: $\rho_{k0}>0$ but $\rho_{0k}=0$ $\forall k\geq 1$. The state 0 is recurrent since $p(0,0)=1$. 0 is an absorbing state.

Example: (Birth-Death chain) State space $S=\{0,1,2,\cdots\}$. Let $p(i,i+1)=p_i>0$,

Memo No. ______
Date / /

$T_a=\inf\{n\geq 1: X_n=a\}$, $T_b=\inf\{n\geq 1: X_n=b\}$

$p(i,i-1)=q_i>0$ if $i\neq 0$ and $p(i,i)=r_i$. $p_i+q_i+r_i=1$. Let $q_0=0$. $N=\inf\{n: X_n=0\}$.

~~We now define a function φ so that $\varphi(X_{N\wedge n})$ is a martingale. First, we have~~

~~For any $0<x<b$, let $\varphi(x)=\mathbb{P}_x(T_0<T_b)$. Then, using the Markov property,~~

~~we get: $\varphi(x)=p_x\mathbb{P}_{x+1}(T_a<T_b)+q_x\mathbb{P}_{x-1}(T_a<T_b)=p_x\varphi(x+1)+q_x\varphi(x-1)$~~

~~Moreover, $\varphi(a-1)=1$, $\varphi(b+1)=0$, $+r_x\mathbb{P}_x(T_a<T_b)=p_x\varphi(x+1)+q_x\varphi(x-1)+r_x\varphi(x)$.~~

~~$\Rightarrow q_x(\varphi(x)-\varphi(x-1))=p_x(\varphi(x+1)-\varphi(x))$~~ $\Rightarrow \varphi(x+1)-\varphi(x)=\frac{q_x}{p_x}(\varphi(x)-\varphi(x-1))$.

~~for~~ $1<x<b-1$.

~~For $x=1$, $\varphi(1)=q_1+p_1\underbrace{\mathbb{P}_2(T_0<T_b)}_{\varphi(2)}+r_1\varphi(1)$~~

~~$\Rightarrow \varphi(2)=[(p_1+q_1)\varphi(1)-q_1]/p_1$~~

We now define a function φ such that $\varphi(X_n)$ is a martingle.

~~Let $\varphi(0)=0$ and $\varphi(1)$~~ For the martingale property to hold, we need

$$\varphi(x)=\mathbb{E}_x\varphi(X_1)=p_x\varphi(x+1)+r_x\varphi(x)+q_x\varphi(x-1).$$

$$\Rightarrow p_x(\varphi(x+1)-\varphi(x))=q_x(\varphi(x)-\varphi(x-1))$$

$\Rightarrow \varphi(k+1)-\varphi(k)=\prod_{j=1}^{k}\frac{q_j}{p_j}$ for $k\geq 1$. $\Rightarrow \varphi(n)=\sum_{k=0}^{n}\prod_{j=1}^{k}\frac{q_j}{p_j}$ for $n\geq 1$.

Let $T_a=\inf\{n\geq 1: X_n=a\}$. For any $a<x<b$, we let $T=T_a\wedge T_b$.

Then, $\varphi(X_{n\wedge T})$ is a bdd martingale. Thus, we have

$$\varphi(x)=\mathbb{E}_x\varphi(X_{n\wedge T}) \xRightarrow{n\to+\infty} \varphi(x)=\mathbb{E}_x\varphi(X_T)=\varphi(a)\mathbb{P}_x(T_a<T_b)+\varphi(b)\mathbb{P}_x(T_b<T_a)$$

$\Rightarrow \varphi(x)=\varphi(a)\mathbb{P}_x(T_a<T_b)+\varphi(b)[1-\mathbb{P}_x(T_a<T_b)]$

$\Rightarrow \mathbb{P}_x(T_a<T_b)=\frac{\varphi(b)-\varphi(x)}{\varphi(b)-\varphi(a)}$. In particular, let $a=0$ and $b=M$, we get

$$\mathbb{P}_x(T_0>T_M)=\frac{\varphi(x)-\varphi(0)}{\varphi(b)-\varphi(0)}=\frac{\varphi(x)}{\varphi(M)}.$$

Letting $M\to+\infty$, we have $\mathbb{P}_x(T_0>T_M)\to\mathbb{P}_x(T_0=\infty)=\lim_{M\to\infty}\frac{\varphi(x)}{\varphi(M)}$.

Thus, we get that 0 is recurrent if and only if $\varphi(M)\to\infty$ as $M\to\infty$, i.e.,

Memo No.
Date / /

$$\varphi(\infty) = \sum_{m=0}^{\infty} \prod_{j=1}^{m} \frac{q_j}{p_j} = \infty. \qquad \text{If } \varphi(\infty)<\infty, \text{ then } \mathbb{P}_x(T_0=\infty) = \frac{\varphi(x)}{\varphi(\infty)}.$$

If 0 is recurrent / transient, then all states are recurrent / transient since the chain is irreducible.

Countable space S

3. Stationary Measure: Given a transition probability p, a measure μ is said ~~A measure~~ to be stationary if $\mu p = \mu$, i.e., $\sum_x \mu(x) p(x,y) = \mu(y)$. (*)

[$p^n(x, A) = \mathbb{P}_x(X_n \in A)$. $\mu p^n(A) = \int \mu(dx)\, p^n(x,A)$, regard as a "matrix multiplication".]

Given (*), we have $\mathbb{P}_\mu(X_1 = y) = \sum_x \mu(x) p(x,y) = \mu(y)$. Using the Markov property and induction in n, we get $\mathbb{P}_\mu(X_n = y) = \mu(y)$ $\forall n \ge 1$.

If μ is a ~~stationary~~ prob. distribution, we say μ is a stationary distribution. It represents a possible equilibrium for the chain.

There is a stronger notion of stationary measure: μ is said to be a reversible measure if it satisfies the detailed balance condition if $\mu(x) p(x,y) = \mu(y) p(y,x)$.

Summing over x, we get $\sum_x \mu(x) p(x,y) = \mu(y)$, i.e., μ is stationary.

"The amount of mass that moves from x to y in 1 jump is the same as the amount that moves from y to x."

Example: Simple RW on graph $G = (V, E)$.

We claim that $\mu(x) = \deg(x)$ defines a reversible measure: for $x \sim y$, $\mu(x) p(x,y) = \deg(x) \frac{1}{\deg(x)} = 1 = \mu(y) p(y,x) = \deg(y) \frac{1}{\deg(y)}$.

Given a finite graph, we have $\sum_x \mu(x) = \sum_x \deg(x) = 2|E| < +\infty$.

Then, we get a stationary distribution by dividing $\mu(S)$: $\tilde{\mu}(x) = \frac{\deg(x)}{2|E|}$.

In particular, for the SRW on $\mathbb{Z}^d$, $\mu(x) \equiv 1$ is a stationary measure.

$\boxed{\deg(x) = \sum_y c(x,y) \text{ is a stationary measure}}$

Memo No.
Date / /

~~Example~~: Asymmetric RW (~~Gambler's ruin~~) $p(i, i+1) = p$, $p(i, i-1) = q = 1-p$, ~~$p(0,0) = 1$, $p(N,N) = 1$~~.

Let μ be a stationary measure, then it satisfies the equations

$$\sum_x \mu(x) p(x,y) = p\,\mu(y-1) + q\,\mu(y+1) = \mu(y) \Rightarrow \mu(y+1) - \mu(y) = \frac{p}{q}(\mu(y) - \mu(y-1)).$$

We get $\mu(y+1) - \mu(y) = \left(\frac{p}{q}\right)^y (\mu(1) - \mu(0))$.

$y \ge 1$

$$\Rightarrow \mu(y) = \mu(1) + \frac{p}{q}(\mu(1) - \mu(0)) + \cdots + \left(\frac{p}{q}\right)^{y-1} (\mu(1) - \mu(0))$$

$$= \mu(0) + \frac{1 - \left(\frac{p}{q}\right)^y}{1 - \frac{p}{q}} (\mu(1) - \mu(0)).$$

Let $\mu(0) = 1$, $\mu(1) = \frac{p}{q}$, we get $\mu(y) = \left(\frac{p}{q}\right)^y$, $y \ge 0$. Similarly, it holds also for $y < 0$.

Example (Ehrenfest chain) $S = \{0, 1, \cdots, r\}$. $p(k, k+1) = \frac{r-k}{r}$, $p(k, k-1) = \frac{k}{r}$.

We claim that $\mu(k) = 2^{-r} \binom{r}{k}$ is a stationary distribution. [μ corresponds to tossing r coins to determine which urn each ball is to be placed in, and the transitions of the chain correspond to picking a coin at random and turning it over.] We can check directly:

$$\mu(k+1)\, p(k+1, k) = 2^{-r} \binom{r}{k+1} \frac{k+1}{r} = 2^{-r} \binom{r-1}{k}$$

$$\mu(k)\, p(k, k+1) = 2^{-r} \binom{r}{k} \frac{r-k}{r} = 2^{-r} \binom{r-1}{k}. \quad \mu \text{ is reversible.}$$

Example (Birth-Death chain) We claim that $\mu(x) = \prod_{k=1}^{x} \frac{p_{k-1}}{q_k}$, $x \ge 0$, is reversible.

In fact, $\mu(x) p(x, x+1) = \prod_{k=1}^{x} \frac{p_{k-1}}{q_k} \cdot p_x = \frac{\prod_{k=0}^{x} p_k}{\prod_{k=1}^{x} q_k} = q_{x+1} \frac{\prod_{k=0}^{x} p_k}{\prod_{k=1}^{x+1} q_k} = q_{x+1} \prod_{k=1}^{x+1} \frac{p_{k-1}}{q_k}$

$= \mu(x+1) p(x+1, x)$.

$\boxed{\text{MC that have a reversible measure}}$ is called a reversible MC

Theorem (Time reversal for Markov chains) Let μ be a stationary measure and suppose X_0 has distribution μ. Then, $Y_m = X_{n-m}$, $0 \le m \le n$, is a MC with initial measure μ and transition probability: $q(x,y) = \mu(y) p(y,x) / \mu(x)$.

q is called "dual transition probability". $q = p$ if μ is reversible.

Memo No. ____________
Date / /

Pf: $\mathbb{P}_\mu(Y_{m+1}=y \mid Y_m=x, Y_{m-1}=x_{m-1}, \cdots, Y_0=x_0)$

$$= \frac{\mathbb{P}_\mu(Y_{m+1}=y, Y_m=x, Y_{m-1}=x_{m-1}, \cdots, Y_0=x_0)}{\mathbb{P}_\mu(Y_m=x, Y_{m-1}=x_{m-1}, \cdots, Y_0=x_0)}$$

$$= \frac{\mathbb{P}_\mu(X_n=x_0, \cdots, X_{n-m+1}=x_{m-1}, X_{n-m}=x, X_{n-m-1}=y)}{\mathbb{P}_\mu(X_n=x_0, \cdots, X_{n-m+1}=x_{m-1}, X_{n-m}=x)}$$

$$= \frac{\mathbb{P}_\mu(X_n=x_0, \cdots, X_{n-m+1}=x_{m-1}, X_{n-m}=x \mid X_{n-m-1}=y)\,\mathbb{P}_\mu(X_{n-m-1}=y)}{\mathbb{P}_\mu(X_n=x_0, \cdots, X_{n-m}=x)}$$

$$= \frac{\mu(y)\,p(y,x)\,p(x,x_{m-1})\cdots p(x_1,x_0)}{\mu(x)\,p(x,x_{m-1})\cdots p(x_1,x_0)} = \frac{\mu(y)p(y,x)}{\mu(x)}. \quad \square$$

~~Not Only~~ We now give a necessary and sufficient condition for a MC to be reversible.

Thm (Kolmogorov's cycle condition) Suppose p is irreducible. A necessary and sufficient condition for the existence of a reversible measure is: (i) $p(x,y)>0$ implies $p(y,x)>0$; (ii) for any loop $x_0, x_1, x_2, \cdots, x_n=x_0$ with $\prod_{1\le i\le n} p(x_i, x_{i-1})>0$, $\prod_{i=1}^{n} \frac{p(x_{i-1},x_i)}{p(x_i,x_{i-1})}=1$.

Pf: $\boxed{\text{"}\Rightarrow\text{"}}$ ~~$\forall x,y\in S$, $\exists n\in\mathbb{N}$ s.t. $p^n(x,y)>0$, $p^n(y,x)>0$.~~

Irreducibility implies that any stationary measure has $\mu(y)>0$ $\forall y\in S$. Otherwise,

$0=\mu(y)=\sum_x \mu(x)p(x,y) \Rightarrow \mu(x)=0$ $\forall x\in S$ s.t. $p(x,y)>0 \Rightarrow \mu(x)=0$ $\forall x\in S$ s.t. $p^2(x,y)>0$

$\Rightarrow \cdots \Rightarrow \mu(x)=0$ $\forall x\in S$. Then, the condition $\mu(x)p(x,y)=\mu(y)p(y,x)$ implies (i) holds. For (ii), we notice that

$$\prod_{i=1}^{n}\frac{p(x_{i-1},x_i)}{p(x_i,x_{i-1})} = \prod_{i=1}^{n}\frac{\mu(x_i)}{\mu(x_{i-1})} = 1.$$

For the sufficiency, fix $a\in S$, set $\mu(a)=1$. If $x_0=a, x_1, x_2, \dots, x_n=x$ is a sequence with $\prod_{i=1}^{n} p(x_{i-1},x_i)>0$ (such a sequence exists), we let

$$\mu(x)=\prod_{i=1}^{n}\frac{p(x_{i-1},x_i)}{p(x_i,x_{i-1})}.$$

$\boxed{\prod_{i=1}^{n} p(x_i,x_{i-1})>0}$

The cycle condition guarantees that this def. does not depend on the path.

Memo No. ____________
Date / /

To check the reversibility condition, we ~~add~~ note that if $p(x,y)>0$ (so $p(y,x)>0$), then adding $x_{n+1}=y$ to the end of a path to x, we have

$$\mu(y)=\mu(x)\frac{p(x,y)}{p(y,x)} \Rightarrow \mu(y)p(y,x)=\mu(x)p(x,y). \quad \square$$

Theorem: If S is finite, then any MC has a stationary ~~measure~~ distribution. In addition, if p is irreducible, the stationary distribution π is unique and statisfies $\pi(x)>0$.

Pf: An elementary proof: let $|S|=n$. Note the rows of p add to 1, so $p-I$ has rank $\le n-1$. So there is a vector ν so that $\nu p=\nu$. ~~Let $q=(I+p)/2$ be the transition matrix~~

By the Perron-Frobenius theorem, we can choose the signs of ν such that all entries of ν are non-negative.

Let $q=(I+p)/2$ be the transition matrix of the lazy chain. We claim that ~~when~~ if p is irreducible, then $q^{n-1}(x,y)>0$ $\forall x,y\in S$. In fact, $\forall x\ne y$, there exists a minimum k s.t. $p^k(x,y)>0$. Since k is minimum, the shortest path does not visit any state more than once, so $k\le n-1$. It follows that $r(x,y):=q^{n-1}(x,y)>0$. Then, we normalize ν to get a stationary distribution π with $\sum_x \pi(x)=1$. From $\pi(y)=\sum_x \pi(x)r(x,y)$, we see that $\pi(y)>0$ $\forall y\in S$. Now, suppose π' is another stationary distribution. We also have $\pi'(y)=\sum_x \pi'(x)r(x,y)$ and $\pi'(y)>0$ $\forall y\in S$. Let $y_0\in S$ s.t. $\frac{\pi(y_0)}{\pi'(y_0)}=\min_{x\in S}\frac{\pi(x)}{\pi'(x)}$. Since $\sum_x \pi(x)=\sum_x \pi'(x)=1$, we have $\frac{\pi(y_0)}{\pi'(y_0)}=:a\le 1 \Rightarrow \frac{\sum_x \pi(x)r(x,y_0)}{\sum_x \pi'(x)r(x,y_0)}=a\le 1$.

$a\le$

$\Downarrow$

$\pi(x)=a\pi'(x)$ $\forall x$

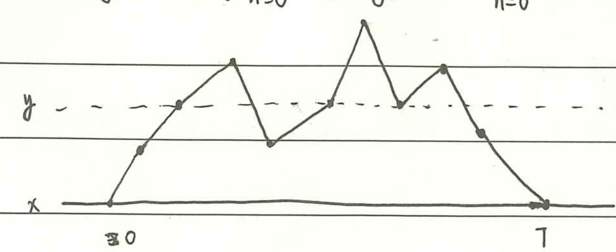
$\Downarrow$

$a=1$. $\square$

Theorem (Existence of stationary measure) Let x be a recurrent state. Let $T=\inf\{n\ge 1: X_n=x\}$. Then, μ_x defines a stationary measure if we let

$$\mu_x(y)=\mathbb{E}_x\left(\sum_{n=0}^{T-1}\mathbb{1}_{\{X_n=y\}}\right)=\sum_{n=0}^{\infty}\mathbb{P}_x(X_n=y, T>n)$$

Pf:



$\mu_x(y)$ is the expected # of visits ~~of~~ to y in $\{0, 1, \dots, T-1\}$, and $(\mu_x p)(y)$ $= \sum_z \mu_x(z) p(z, y)$ is the expected # of visits to y in $\{1, \dots, T\}$, which is $= \mu_x(y)$ since $X_T = X_0 = x$. To give a rigorous proof:

$$\sum_y \mu_x(y) p(y,z) = \sum_y \sum_{n=0}^{\infty} \mathbb{P}_x(X_n = y, T>n) p(y,z) = \sum_{n=0}^{\infty} \sum_y \mathbb{P}_x(X_n=y, T>n) p(y,z).$$

① If $z \neq x$, then

$$\sum_y \mathbb{P}_x(X_n=y, T>n) p(y,z) = \sum_y \mathbb{P}_x(X_n=y, X_{n+1}=z, T>n)$$
$$= \mathbb{P}_x(X_{n+1}=z, T>n+1)$$

Then, $\sum_y \mu_x(y) p(y,z) = \sum_{n=0}^{\infty} \mathbb{P}_x(X_{n+1}=z, T>n+1) = \sum_{n=1}^{\infty} \mathbb{P}_x(X_n=z, T>n) = \sum_{n=0}^{\infty} \mathbb{P}_x(X_n=z, T>n)$ $= \mu_x(z)$.

② If $z = x$, then

$$\sum_y \mathbb{P}_x(X_n=y, T>n) p(y,x) = \sum_y \mathbb{P}_x(X_n=y, X_{n+1}=x, T>n) = \mathbb{P}(X_{n+1}=x, T>n) = \mathbb{P}_x(T=n+1).$$

Thus, $\sum_{n=0}^{\infty} \sum_y \mathbb{P}_x(X_n=y, T>n) p(y,x) = \sum_{n=0}^{\infty} \mathbb{P}_x(T=n+1) = 1 = \mu_x(x)$ since $\mathbb{P}_x(T=0)=0$ by def.

$[\mu_x(x) = \mathbb{E}_x \sum_{n=0}^{T-1} \mathbb{1}_{\{X_n=x\}} = 1]$ □

Rmk: If x is transient, we have $\mu_x p(z) = \mu_x(z)$ $\forall z \neq x$. For $z = x$, we have $\sum_{n=0}^{\infty} \mathbb{P}_x(T=n+1) < 1$, so $\mu_x p(x) < \mu_x(x)$.

Rmk: To show that μ_x defines a measure, we still need to show that $\mu_x(y) < \infty$ for all y. Observe that $\mu_x p = \mu_x$ implies $\mu_x p^n = \mu_x$ $\forall n \geq 1$, and $\mu_x(x) = 1$. ~~Thus~~ $\sum_y \mu_x(y) p^n(y,x) = \mu_x(x) = 1 \Rightarrow$ if $p^n(y,x) > 0$, then $\mu_x(y) < +\infty$. This is true for all n, so we see that $\rho_{yx} > 0 \Rightarrow \mu_x(y) < +\infty$. When x is recurrent, $\rho_{xy} > 0$ implies $\rho_{yx} > 0$, and it follows that $\mu_x(y) < +\infty$. On the other hand, if $\rho_{xy} = 0$, then $\mu_x(y) = 0$ by def.

The above theorem allows us to construct a stationary measure for each closed set of recurrent states. Conversely:

Thm: If p is irreducible and recurrent, then the stationary measure is unique up to constant multiples.

Pf: Let ν be an arbitrary stationary measure and let $a \in S$,

$$\nu(z) = \sum_y \nu(y) p(y,z) = \nu(a) p(a,z) + \sum_{y \neq a} \nu(y) p(y,z)$$
$$= \nu(a) p(a,z) + \sum_{y \neq a} \nu(a) p(a,y) p(y,z) + \sum_{x \neq a} \sum_{y \neq a} \nu(x) p(x,y) p(y,z)$$
$$= \nu(a) \mathbb{P}_a(X_1 = z) + \nu(a) \mathbb{P}_a(X_1 \neq a, X_2 = z) + \mathbb{P}_\nu(X_0 \neq a, X_1 \neq a, X_2 = z)$$

Continuing this expansion, we get

$$(*) \quad \nu(z) = \nu(a) \sum_{m=1}^{n} \mathbb{P}_a(X_k \neq a, 1 \leq k \leq m-1, X_m = z) + \mathbb{P}_\nu(X_j \neq a, 0 \leq j \leq n-1, X_n = z).$$

This converges to 0 if ν is a finite measure using a is recurrent. But ν may be a ∞ measure.

Let $n \to \infty$, $\nu(z) \geq \nu(a) \mu_a(z)$, where $\mu_a(z) = \sum_{n=0}^{\infty} \mathbb{P}_a(X_n = z, T>n)$. Recall that μ_a is a stationary measure with $\mu_a(a) = 1$. Then, we observe that

$$\nu(a) = \sum_x \nu(x) p^n(x,a) \geq \nu(a) \sum_x \mu_a(x) p^n(x,a) = \nu(a) \mu_a(a) = \nu(a).$$

Thus, $\geq$ must be $=$, and we must have $\nu(x) = \nu(a) \mu_a(x)$ whenever $p^n(x,a) > 0$. Since p is irreducible, it follows that $\nu(x) = \nu(a) \mu_a(x)$ $\forall x \in S$. □

We now turn attention to stationary distributions. Stationary measures may exist for transient chains (e.g., SRW in $d \geq 3$). But:

Theorem: If there is a stationary distribution π, then all states y with $\pi(y) > 0$ are recurrent.

Pf: Using $\pi p^n = \pi$, we get $\sum_x \pi(x) \sum_{n=1}^{\infty} p^n(x,y) = \sum_{n=1}^{\infty} \pi(y) = \infty$, when $\pi(y) > 0$.

On the other hand, we have seen that $\sum_{n=1}^{\infty} p^n(x,y) = \frac{\rho_{xy}}{1-\rho_{yy}}$. Thus,

$$\infty = \sum_x \pi(x) \frac{\rho_{xy}}{1-\rho_{yy}} \leq \frac{1}{1-\rho_{yy}} \Rightarrow \rho_{yy} = 1.$$ □

Memo No.

Date / /

<u>Theorem</u>: If p is irreducible and has stationary distribution π, then $\pi(x) = \frac{1}{\mathbb{E}_x T_x}$.

<u>Pf</u>: ~~Irreducibility implies that π is equal to μ_x up constant multiples. Since $\mu_x(x) = 1 > 0$, we have $\pi(x) > 0$. Then, by the previous theorem, all states are recurrent.~~ Irreducibility implies that $\pi(x) > 0$ (otherwise, $\sum_y \pi(y) p^n(y,x) = \pi(x) = 0$ $\Rightarrow \pi(y) = 0 \ \forall y \in S$). Then, the previous theorem implies that all states are recurrent.

Now, μ_x defines a stationary measure with $\mu_x(x) = 1$ and

$$\mu_x(y) = \sum_{n=0}^{\infty} \mathbb{P}_x(X_n = y, T_x > n)$$

Using Fubini's thm, we get $\sum_y \mu_x(y) = \sum_{n=0}^{\infty} \mathbb{P}_x(T_x > n) = \mathbb{E}_x T_x$. Since the stationary measure is unique up to constant multiples, we must have

$$\pi(x) = \frac{\mu_x(x)}{\sum_y \mu_x(y)} = \frac{1}{\mathbb{E}_x T_x}.$$

□

<u>Rmk</u>: Another consequence is the following identity: $\sum_x \frac{1}{\mathbb{E}_x T_x} p(x,y) = \frac{1}{\mathbb{E}_y T_y}$.

A state is said to be <u>positive recurrent</u> if $\mathbb{E}_x T_x < +\infty$. A <u>recurrent</u> state with $\mathbb{E}_x T_x = \infty$ is said to be <u>null recurrent</u>.

Note: positive recurrence implies recurrent: if $\mathbb{E}_x T_x < +\infty$, then $\mathbb{P}_x(T_x = \infty) = 0$.

<u>Theorem</u>: If p is irreducible, then TFAE:

(i) Some $x \in S$ is positive recurrent.

(ii) There is a stationary distribution.

(iii) All states a positive recurrent.

In other words, <u>being positive recurrent</u> is a <u>class property</u>. If it holds for one state in an irreducible set, then it is true for all.

<u>Pf</u>: (i) $\Rightarrow$ (ii). If x is positive recurrent, then

Memo No.

Date / /

$$\pi(y) = \frac{\mu_x(y)}{\mathbb{E}_x T_x}$$ defines a stationary distribution.

(ii) $\Rightarrow$ (iii) Irreducibility tells that $\pi(y) > 0 \ \forall y \in S$. Then, by previous thm, $\mathbb{E}_y T_y = \frac{1}{\pi(y)} < +\infty$. Hence, $\forall y \in S$ is positive recurrent.

(iii) $\Rightarrow$ (i) Trivial. □

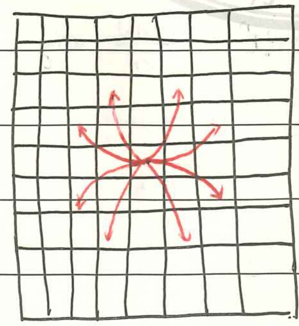
<u>Example</u> (Birth-Death chain) We have seen that $\mu(x) = \prod_{k=1}^{x} \frac{p_{k-1}}{q_k}$ is a reversible measure. Hence, it has a stationary distribution if and only if $\sum_x \underbrace{\prod_{k=1}^{x} \frac{p_{k-1}}{q_k}}_{a(x)} < +\infty$. On the other hand, we have seen that the chain is recurrent if and only if $\sum_{x=0}^{\infty} \underbrace{\prod_{j=1}^{x} \frac{q_j}{p_j}}_{b(x)} = +\infty$.

$$\boxed{b(x) = \frac{p_0}{p_x} \frac{1}{a(x)} \geq \frac{p_0}{a(x)}}$$

$G = (V, E)$

<u>Example</u>: We have seen that for the SRW on a connected graph, there is a reversible measure $\mu(x) = \deg(x)$. Moreover, the SRW is irreducible. When the graph is finite, there is a stationary distribution $\pi(x) = \frac{\deg(x)}{2|E|}$. Hence, $\forall x \in V$, we have $\mathbb{E}_x T_x = \frac{2|E|}{\deg(x)}$

<u>Example</u>: (Random walk of a knight on a chess board) A chess board is an 8x8 grid of squares. We regard it as a simple RW on a graph with 64 vertices.



A knight randomly moves to another grid starting from the lower right corner.

On ~~avergerate~~ average, how many does it take the knight to return to this corner?

The degrees of the vertices are given by

$$\begin{matrix} 2 & 3 & 4 & 4 \\ 3 & 4 & 6 & 6 \\ 4 & 6 & 8 & 8 \\ 4 & 6 & 8 & 8 \end{matrix}$$

Total degree $= 336$

The degree of the lower right corner x is $\deg(x)=2$. Hence, $\pi(x) = \frac{2}{336} = 168 \Rightarrow \mathbb{E}_x T_x = \frac{1}{\pi(x)} = 168$. □

4. Recurrence ~~of~~ or Transience of SRW on $\mathbb{Z}^d$

Theorem: Simple random walk in recurrent in $d \le 2$ and transient in $d \ge 3$.

SRW: $\mathbb{P}(X_i = e_j) = \mathbb{P}(X_i = -e_j) = \frac{1}{2d}$. $S_n = X_1 + X_2 + \cdots + X_n$.

Let $\tau_0 = 0$ and $\tau_n = \inf\{m > \tau_{n-1} : S_m = 0\}$ be the time of the n-th return to 0.

By the Strong MP, we have $\mathbb{P}(\tau_n < \infty) = \mathbb{P}(\tau_1 < \infty)^n$.

[Theorem: TFAE: (i) $\mathbb{P}(\tau_1 < \infty) = 1$. (ii) $\mathbb{P}(S_m = 0 \text{ i.o.}) = 1$. (iii) $\sum_{m=0}^{\infty} \mathbb{P}(S_m = 0) = \infty$]

"A drunk man will eventually find his way home, but a drunk bird will get lost forever."

In $d=1$, we have the classical result.

$$\mathbb{P}(S_{2n} = 0) = \frac{1+o(1)}{\sqrt{\pi n}} \quad \text{as } n \to \infty.$$

In fact, $\mathbb{P}(S_{2n}=0) = \binom{2n}{n}\left(\frac{1}{2}\right)^{2n} = \frac{(2n)!}{n!\,n!}\left(\frac{1}{2}\right)^n = \frac{1+o(1)}{\sqrt{\pi n}}$ using Stirling's formula:

$$n! \sim n^n e^{-n} \sqrt{2\pi n} \quad \text{as } n \to +\infty.$$

Hence, we have $\sum_{m=0}^{\infty} \mathbb{P}(S_m = 0) = \infty$, showing that 0 is recurrent. By irreducibility, all states are recurrent for the SRW on $\mathbb{Z}$.

$\boxed{\text{Let } p_d(n) = \mathbb{P}(S_n = 0) \text{ in } \mathbb{Z}^d}$

Next, we show 2D SRW is recurrent. In order for $S_{2n}=0$, we must have that for some $0 \le m \le n$, there are m up steps, m down steps, $n-m$ to the left, and $n-m$

to the right. Hence,

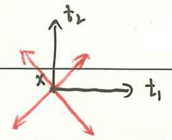
$$\mathbb{P}(S_{2n}=0) = 4^{-2n} \sum_{m=0}^{n} \frac{(2n)!}{m!\,m!\,(n-m)!\,(n-m)!} = 4^{-2n}\binom{2n}{n}\sum_{m=0}^{n}\binom{n}{m}\binom{n}{n-m}$$

$$= 4^{-2n}\binom{2n}{n}^2 = \left[2^{-2n}\binom{2n}{n}\right]^2 = [p_1(2n)]^2 = \frac{1+o(1)}{\pi n}.$$

Consider choosing n balls from a box with n black balls and n white balls

$\Rightarrow \sum_{m=0}^{\infty} \mathbb{P}(S_m = 0) = \infty$, so the SRW is recurrent.

There is a direct proof of $p_2(2n) = [p_1(2n)]^2$: let T_n^1 and T_n^2 be two ind. 1D SRWs. Then, $T_n = (T_n^1, T_n^2)$ jumps from x to $x + (\pm 1, \pm 1)$ with prob. $\frac{1}{4}$.



Then, rotating T_n by $\pi/4$ and dividing by $\sqrt{2}$ gives S_n in $\mathbb{Z}^2$.

Finally, we consider the SRW in 3D. Intuitively, $p_3(2n)$ is of order $O(n^{-3/2})$, which is summable, so the SRW is transient. We can calculate that

$$p_3(2n) = 6^{-2n}\sum_{j,k}\frac{(2n)!}{[j!\,k!\,(n-j-k)!]^2} = 2^{-2n}\binom{2n}{n}\sum_{j,k}\left[3^{-n}\frac{n!}{j!\,k!\,(n-j-k)!}\right]^2$$

$$\le p_1(2n)\max_{j,k} 3^{-n}\frac{n!}{j!\,k!\,(n-j-k)!}$$

where we used $\sum_{j,k} 3^{-n}\frac{n!}{j!\,k!\,(n-j-k)!} = 1$. Next, we claim that

$$\max_{j,k} 3^{-n}\frac{n!}{j!\,k!\,(n-j-k)!} \le \frac{C}{n} \quad \text{for a constant } C > 1.$$

Note the maximum is achieved when all three numbers are as closed as possible to $n/3$ (the difference between them are at most 1). In fact, suppose one of $j, k, n-j-k$ is $< \lfloor n/3 \rfloor$. Then, there is another one that is $> \lfloor n/3 \rfloor$.

WLOG, suppose $j < \lfloor n/3 \rfloor$ and $k > \lfloor n/3 \rfloor$, then $(j+1)!\,(k-1)! \le j!\,k!$, so that $3^{-n}\frac{n!}{j!\,k!\,(n-j-k)!} \le 3^{-n}\frac{n!}{(j+1)!\,(k-1)!\,(n-j-k)!}$.

Then, using Stirlings formula, we get

$$3^{-n}\frac{n!}{j!\,k!\,(n-j-k)!} = [1+o(1)]\frac{3^{-n}n^n}{j^j\,k^k\,(n-j-k)^{n-j-k}}\sqrt{\frac{n}{j\,k\,(n-j-k)}}\,\frac{1}{2\pi}$$

$$= [1+o(1)]\left[1+\frac{3j-n}{n}\right]^j\left[1+\frac{3k-n}{n}\right]^k\left[1+\frac{2n-3j-3k}{n}\right]^{n-j-k}\frac{3^{\frac{3}{2}}}{2\pi n}$$

$$\le C/n.$$

Hence, we get $p_3(2n) \le \frac{C}{n^{3/2}}$, which is summable.

This shows that the 3D SRW is transient.

In dimensions $d\ge 4$: $S_n = (S_n^1, S_n^2, S_n^3, S_n^4, \cdots, S_n^d)$. Let $T_n = (S_n^1, S_n^2, S_n^3)$.

Then, we have $\mathbb{P}(S_n=0) \le \mathbb{P}(T_n=0) \Rightarrow p_d(2n) \le p_3(n) \Rightarrow p_d(n)$ is also summable $\Rightarrow$ Then SRW on $\mathbb{Z}^d$, $d\ge 4$, is transient.

Note that the SRW on $\mathbb{Z}^d$ has a stationary measure (even if SRW on $\mathbb{Z}^d$, $d\ge 3$, is recurrent): the uniform measure $\mu(x) = 1$.

In $d=1,2$, we know the SRW has a unique stationary measure. If there exists a stationary distribution π, then by uniqueness, $\pi(x) = a > 0$ for all $x\in\mathbb{Z}^d$.

Moreover, $1 = \sum_x \pi(x) = \sum_x a = \infty$, a contradiction. Hence, the SRW on $\mathbb{Z}^d$ admits no stationary distribution. As a consequence, 1D and 2D SRW are <u>null recurrent</u>.

5. Asymptotic behavior of Markov chains, Periodicity

We now study the asymptotic behavior of the Markov chain, i.e., $p^n(x,y)$ $\forall x,y\in S$, as $n\to+\infty$. Note $p^n(x,y)\to 0$ as $n\to+\infty$ if y is transient since $\sum_n p^n(x,y) < +\infty$.

Memo No.
Date / /

[For recurrent states y, our goal is to show that $p^n(x,y)\to\pi(y)$ asymptotically under certain conditions if there exists a stationary distribution]

Let $N_n(y) = \sum_{m=1}^{n}\mathbb{1}(X_m=y)$: # of visits to y up to time n.

Thm: Suppose y is recurrent. For any $x\in S$, as $n\to+\infty$, we have

$$\frac{N_n(y)}{n} \to \frac{1}{\mathbb{E}_y T_y}\mathbb{1}(T_y<+\infty) \quad \mathbb{P}_x\text{-a.s.}$$

When $\mathbb{E}_y T_y = \infty$, we let $\frac{1}{\infty} = 0$.

Pf: First, suppose $x=y$ (i.e., we start at y). Let $R(k) = \min\{n\ge 1: N_n(y) = k\}$ – the time of the k-th return to y. Let $t_k = R(k) - R(k-1)$, where $R(0)=0$.

When $X(0)=y$, $t_1, t_2, t_3, \cdots$ are i.i.d. by the strong MP and the strong LLN implies

$$R(k)/k = \frac{1}{k}\sum_{i=1}^{k} t_k \to \mathbb{E}_y T_y \quad \mathbb{P}_y\text{-a.s.}$$

Note that $R(N_n(y)) \le n \le R(N_n(y)+1)$, we have

$$\frac{R(N_n(y))}{N_n(y)} \le \frac{n}{N_n(y)} \le \frac{R(N_n(y)+1)}{N_n(y)+1}\cdot\frac{N_n(y)+1}{N_n(y)}$$

Let $n\to\infty$. We have $N_n(y)\to\infty$ a.s. since y is recurrent. Then, we get

$$\frac{n}{N_n(y)} \to \mathbb{E}_y T_y \quad \mathbb{P}_y\text{-a.s.}$$

Next, we extend it to $x\neq y$. Observe that if $T_y=\infty$, then $N_n(y)=0$ for all n and hence $N_n(y)/n \to 0$ on $\{T_y=\infty\}$.

Conditional on the event $\{T_y<+\infty\}$, the strong MP implies that $t_2, t_3, \cdots$ are i.i.d., so

$\cancel{\mathbb{P}_x(t_k=n) = \mathbb{P}_y(T_y=n)}$ $R(k)/k = t_1/k + (t_2+t_3+\cdots+t_k)/k \to 0 + \mathbb{E}_y T_y$ $\mathbb{P}_x$-a.s.

Repeating the proof for the case $x=y$ shows that $N_n(y)/n \to 1/\mathbb{E}_y T_y$ $\mathbb{P}_x$-a.s. on $\{T_y<\infty\}$. □

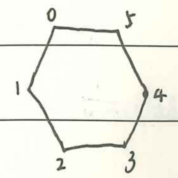
Rmk: This result tells that if y is positive / null recurrent, then the asymptotic fraction of time spent at x is positive and in the second case it is 0.

By the bounded convergence theorem, we have $\mathbb{E}_x \frac{N_n(y)}{n} \to \mathbb{E}_x\left(\frac{1}{\mathbb{E}_y T_y} \mathbb{1}(T_y<+\infty)\right)$, so

$$\frac{1}{n}\sum_{m=1}^{n} p^m(x,y) \to \rho_{xy} / \mathbb{E}_y T_y .$$

This also holds for transient y, since $\mathbb{E}_y T_y = \infty$ and $\sum_{m=1}^{\infty} p^m(x,y) < +\infty$. This shows that $p^n(x,y)$ converges in the Cesaro sense. But this is weaker than the convergence $p^n(x,y)$ itself.

<u>Example</u>: Consider the SRW on the $2n$-cycle:



<u>$2n=6$ case</u>

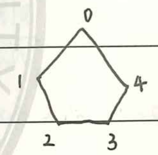
Note $\begin{cases} p^{2k}(0,i)=0 & \text{if } i \text{ is even} \\ p^{2k+1}(0,i)=0 & \text{if } i \text{ is odd} \end{cases}$

Hence, $p^{2k}(0,1) + p^{2k}(0,3) + p^{2k}(0,5) = 0$, but

$$p^{2k+1}(0,1) + p^{2k+1}(0,3) + p^{2k+1}(0,5) = 1.$$

Hence, $p^n(0,x)$ does not converge for this chain as $n \to +\infty$.

On the other hand, if we consider the SRW on the $(2n+1)$-cycle, we do not have the above issue.



<u>$2n+1=5$ case</u>

For example, starting from 0, we can ~~reach~~ come back in both even and odd number (by going through the cycle) of steps.

We will show that in an irreducible MC this "periodicity" issue is the only issue that can prevent the convergence of $p^n(x,y)$.

<u>Def</u>: (Period) Let x be a recurrent state. Let $I_x = \{n \geq 1 : p^n(x,x) > 0\}$. Let d_x be the gcd (greatest common divisor) of I_x. d_x is called the period of x.

<u>Example</u>: For the SRW on the $2n$-cycle, $I_0 = \{0, 2, 4, \cdots\}$. Hence, $d_0 = 2$.

For the SRW on the $(2n+1)$-cycle, we have $2 \in I_0$ and $2n+1 \in I_0$. Hence, $d_0 = 1$.

Period is a "class property".

<u>Lemma</u>: If $\rho_{xy} > 0$, then $d_x = d_y$.

<u>Pf</u>: Let $K, L \in \mathbb{N}$ such that $p^K(x,y) > 0$ and $p^L(y,x) > 0$. (x is recurrent, so $\rho_{yx} > 0$)

Then, $K+L \in I_x \cap I_y$ because $p^{K+L}(y,y) \geq p^L(y,x)p^K(x,y) > 0$, $p^{K+L}(x,x) \geq p^K(x,y)p^L(y,x) > 0$.

Thus, we have $d_y \mid (K+L)$. Let $n \in \mathbb{N}$ be such that $p^n(x,x) > 0$. Then,

$$p^{K+n+L}(y,y) \geq p^L(y,x)\, p^n(x,x)\, p^K(x,y) > 0, \quad \text{so } d_y \mid (K+n+L) \Rightarrow d_y \mid n.$$

Thus, $d_y \mid n \ \forall n \in I_x \Rightarrow d_y \mid d_x$. Similarly, $d_x \mid d_y$. Hence, $d_x = d_y$. □

If a chain is irreducible and $d_x = 1$, then all states in the chain have period 1 and the chain is called <u>aperiodic</u>. Note if there exists a state x with $p(x,x) > 0$, then the chain is aperiodic.

Hence, give an irreducible ~~MC~~ and recurrent MC, people sometimes consider the "~~aperiodic~~ Lazy chain" with transition probability $q = (I+p)/2$. The Lazy chain is always aperiodic.

<u>Lemma</u>: If $d_x = 1$, then $\exists m_0$ s.t. $p^m(x,x) > 0$ for all $m \geq m_0$.

<u>Pf</u>: We first show that I_x contains to consecutive integers $k, k+1$.

A fact from number theory: if the gcd of a set I_x is 1, then there are $i_1, i_2, \cdots, i_m \in I_x$ and positive/negative integer coefficients $c_1, \cdots, c_m \in \mathbb{Z}$ so that

$$c_1 i_1 + c_2 i_2 + \cdots + c_m i_m = 1.$$

Let $a_i = c_i^+$ be the positive coefficients and $b_i = c_i^-$ be the absolute values of the negative ones. The above equation reads $a_1 i_1 + \cdots + a_m i_m = (b_1 i_1 + \cdots + b_m i_m) + 1$, $a_i \in \mathbb{N}$, $b_i \in \mathbb{N}$.

Then, we let $k = b_1 i_1 + \cdots + b_m i_m$.

Given $k, k+1$, we get that $2k$, $2k+1 = k+(k+1)$, $2k+2 = 2(k+2) \in I_x$

$\Rightarrow 3k = 2k+k$, $3k+1$, $3k+2$, $3k+3 = 3(k+3) \in I_x \Rightarrow \cdots \Rightarrow (k-1)k, (k-1)k+1, \cdots (k-1)k+k-1$ $\in I_x$ for k consecutive integers. For any $m \geq (k-1)k+k-1$, we can find $a \in \{(k-1)k, \cdots, (k-1)(k+1)\}$ and $b \in \mathbb{N}$ s.t. $m = a + bk$. Hence, the lemma holds for $m_0 = (k-1)k$. □

Theorem (Convergence theorem) Suppose p is irreducible, aperiodic, and has stationary distribution π. Then, as $n \to +\infty$, $p^n(x,y) \to \pi(y)$.

$\boxed{\text{Rmk}}$: We have shown that the Cesaro limit $\frac{1}{n}\sum_{m=1}^{n} p^m(x,y) \to \frac{\rho_{xy}}{\mathbb{E}_y T_y} = \frac{1}{\mathbb{E}_y T_y} = \pi(y)$.

Hence, if $p^n(x,y)$ has a limit, it must be equal to $\pi(y)$.

More intuitively, if $\frac{1}{n}\sum_{m=1}^{n} p^m(x,y) \to \pi(y)$ for a distribution π, then $\pi p(y) = \lim_{n\to\infty} \frac{1}{n}\sum_{m=1}^{n} p^{m+1}(x,y)$ $= \lim_{n\to\infty} \frac{1}{n}\sum_{m=2}^{n+1} p^m(x,y) = \pi(y)$. Hence, π is stationary.

Pf: The proof use a nice "coupling" argument. We consider two i.i.d. copies of the MC.

That is, we define a MC on $S^2 = S\times S$ with transition probability $\bar{p}((x_1,y_1),(x_2,y_2)) = p(x_1,x_2)\, p(y_1,y_2)$.

Hence, each coordinate moves independently. We first check that $\bar{p}$ is irreducible.

This is the only place where aperiodicity is used.

By irreducibility of p, $\exists\, K, L \in \mathbb{N}$ s.t. $p^K(x_1,x_2)>0$ and $p^L(y_1,y_2)>0$. By previous lemma, $\exists\, m_0$ large s.t. $p^m(x_2,x_2)>0$ and $p^m(y_2,y_2)>0$ for $m \ge m_0$. Then, let $M > m_0$, we have $p^{K+L+M}(x_1,x_2) \ge p^K(x_1,x_2)\, p^{L+M}(x_2,x_2) > 0$, $p^{K+L+M}(y_1,y_2) \ge p^L(y_1,y_2)\, p^{K+M}(y_2,y_2) > 0$.

Thus, $\bar{p}^{K+L+M}((x_1,y_1),(x_2,y_2)) > 0$.

Next, ~~also~~ observe that $\bar{\pi}(x,y) = \pi(x)\pi(y)$ is a stationary distribution for $\bar{p}$, so all states for $\bar{p}$ are positive recurrent.

Now, let (X_n, Y_n) denote the chain on $S\times S$, and let T be the first time that this chain hits the diagonal $\{(y,y): y\in S\}$. Let $T_{(x,x)}$ be the hitting time of (x,x): $T_{(x,x)} = \inf\{n\ge 0: X_n = x, Y_n = x\}$. Since $\bar{p}$ is irreducible and recurrent, $T_{(x,x)} < +\infty$ a.s. so $T<\infty$ a.s. Finally, observe that X_n and Y_n have the same distribution on $\{T\le n\}$. In fact,

$$\mathbb{P}(X_n=y, T\le n) = \sum_{m=1}^{n}\sum_x \mathbb{P}(T=m, X_m=x, X_n=y)$$

$$= \sum_{m=1}^{n}\sum_x \mathbb{P}(T=m, X_m=x)\,\mathbb{P}(X_n=y \mid X_m=x) = \sum_{m=1}^{n}\sum_x \mathbb{P}(T=m, Y_m=x)\,\mathbb{P}(Y_n=y\mid Y_m=x)$$

$$= \mathbb{P}(Y_n=y, T\le n).$$

Hence, we have $\mathbb{P}(X_n=y) = \mathbb{P}(Y_n=y, T\le n) + \mathbb{P}(X_n=y, T>n)$

$\le \mathbb{P}(Y_n=y) + \mathbb{P}(T>n, X_n=y)$.

Similarly, we have $\mathbb{P}(Y_n=y) \le \mathbb{P}(X_n=y) + \mathbb{P}(T>n, Y_n=y)$.

So, $|\mathbb{P}(X_n=y) - \mathbb{P}(Y_n=y)| \le \mathbb{P}(T>n, X_n=y) + \mathbb{P}(T>n, Y_n=y)$

$\Rightarrow \sum_y |\mathbb{P}(X_n=y) - \mathbb{P}(Y_n=y)| \le 2\mathbb{P}(T>n)$

Let $X_0 = x$ and Y_0 have the stationary distribution π, then Y_n has distribution π and it follows that: $\underbrace{\sum_y |p^n(x,y) - \pi(y)|}_{2\|p^n-\pi\|_{TV}} \le 2\mathbb{P}(T>n) \to 0$ as $n\to+\infty$.

So, we in fact get a stronger result: $\|p^n - \pi\|_{TV} \to 0$ as $n\to+\infty$. □

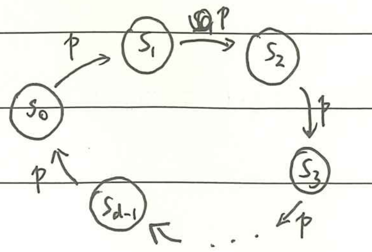
Next, we study the asymptotic behavior for periodic MC.

Lemma: Suppose p is irreducible, recurrent, and all states have period d. Fix $x\in S$, and for each $y\in S$, let $K_y = \{n\ge 1: p^n(x,y)>0\}$.

(i) $\exists\, r_y \in \{0,1,\dots,d-1\}$ so that if $n\in K_y$, then $n = r_y \bmod d$, i.e., the difference $n - r_y$ is a multiple of d.

(ii) Let $S_r = \{y: r_y = r\}$ for $r\in\{0,1,\dots,d-1\}$. If $y\in S_i$, $z\in S_j$, and $p^n(y,z)>0$, then $n = (j-i) \bmod d$.

(iii) $S_0, S_1, \dots, S_{d-1}$ are irreducible classes for p^d, and all states have ~~period~~ period 1.



Pf: (i) Let $m(y)$ be such that $p^{m(y)}(y,x)>0$.

If $n\in K_y$, then $p^{n+m(y)}(x,x) \ge p^n(x,y)\, p^{m(y)}(y,x) > 0$.

$\Rightarrow n+m(y) \in I_x \Rightarrow d \mid (n+m(y))$.

Then, we let $r_y = (-m(y)) \bmod d$.

(ii) Let n, m be such that $p^n(y,z), p^m(x,y) > 0$. Since $p^{m+n}(x,z) > 0$, by (i), we have $n+m \equiv j \mod d$. Since $m \equiv i \mod d$, we get $n \equiv j-i \mod d$.

(iii) $\forall x, y \in S_i$, $\exists n$ s.t. $p^n(x,y) > 0 \Rightarrow n \equiv 0 \mod d$ by (ii)

$\Rightarrow (p^d)^k(x,y) > 0$ for some $k \in \mathbb{N}$. This gives irreducibility. $\square$

A partition $S_0, S_1, \cdots, S_{d-1}$ satisfying the above lemma is called a <u>cyclic decomposition</u> of the state space. Up to the labels, ~~these sets~~ this decomposition is <u>unique</u>.

<u>Theorem</u>: (Convergence theorem, periodic case) Suppose p is irreducible, has a stationary distribution π, and all states have period d. Let $x \in S$, and let $S_0, S_1, \cdots, S_{d-1}$ be the cyclic decomposition of the state space with $x \in S_0$. If $y \in S_r$, then

$$\lim_{m \to +\infty} p^{md+r}(x,y) = \pi(y) d.$$

<u>Pf</u>: If $y \in S_0$, then by previous lemma (iii) and MC convergence theorem, we have that $\lim_{m \to +\infty} p^{md}(x,y)$ exists. To find the limit, we notice that

$$\frac{1}{n}\sum_{m=1}^{n} p^m(x,y) \to \pi(y).$$

But, we know $p^m(x,y) = 0$ unless $d \mid m$ by the previous Lemma (ii), so we get that $\frac{1}{nd}\sum_{m=1}^{n} p^{md}(x,y) \to \pi(y) \Rightarrow p^{md}(x,y)$ must converge to $\pi(y)d$.

If $y \in S_r$ with $r \in \{1, 2, \cdots, d-1\}$, then

$$p^{md+r}(x,y) = \sum_{z \in S_r} p^r(x,z)\, p^{md}(z,y).$$

By the previous case, $p^{md}(z,y) \to \pi(y)d$ as $m \to +\infty$. Moreover, $p^{md}(z,y) \le 1$ and $\sum_z p^r(x,z) = 1$, so by DCT, we have

$$\lim_{m \to \infty} p^{md+r}(x,y) = \sum_{z \in S_r} p^r(x,z)\, \pi(y) d = \pi(y) d. \quad \square$$

Chapter 3 Brownian Motion

Brownian motion is a <u>continuous time</u> <u>Gaussian</u> <u>Markov process</u> with <u>stationary independent increments</u>. It is also a <u>continuous time martingale</u>.

1. Definition & Construction of BM

<u>Def</u>: A one-dimensional Brownian motion (BM) is a real-valued process $B_t, t \ge 0$, with the following properties.

(a) (Independent increments) If $t_0 < t_1 < \cdots < t_n$, then $B(t_0), B(t_1) - B(t_0), \cdots, B(t_n) - B(t_{n-1})$ are independent.

(b) (Stationary Gaussian increments) If $s, t \ge 0$, then $\mathbb{P}(B(s+t) - B(s) \in A) = \int_A \frac{1}{\sqrt{2\pi t}} \exp(-\frac{x^2}{2t})\, dx$, i.e., $B(s+t) - B(s) \sim N(0,t)$.

(c) $t \mapsto B_t$ is continuous a.s.

<u>Rmk</u>: Discuss why these three properties make sense physically.

<u>Rmk</u>: We say the BM is standard if $B_0 = 0$.

(Translation invariance)

<u>Propostion</u>: $\forall t_0 \ge 0$, $\{B_{t+t_0} - B_{t_0}\}_{t \ge 0}$ is a standard BM independent of B_{t_0}.

<u>Pf</u>: Let $\mathcal{A}_1 = \sigma(B_{t_0})$ and $\mathcal{A}_2$ be the events of the form $\{B(t_1) - B(t_0) \in A_1, \cdots, B(t_n) - B(t_{n-1}) \in A_n\}$.

Then, $\mathcal{A}_2$ is a π-system generating $\sigma(\{B_{t+t_0} - B_{t_0}\}_{t \ge 0})$. By π-λ theorem, $\{B_{t+t_0} - B_{t_0}\}_{t \ge 0}$ is independent of B_{t_0}. Then, we can show that $B_{t+t_0} - B_{t_0}$ statisfies the properties (a)-(c) for standard BM. $\square$

<u>Proposition</u> (Brownian scaling relation): Let B_t be a standard BM. Then, $\forall t > 0$, $\{t^{-\frac{1}{2}} B_{st} : s \ge 0\}$ is a standard BM (with time s).

<u>Pf</u>: We check $t^{-\frac{1}{2}} B_{st}$ satisfies the properties (a)-(c) above. $\square$

An equivalent definition of BM with $B_0=0$: B_t, $t\geq 0$, is a real-valued process satisfying

(a') $B(t)$ is a Gaussian process (all its finite dimensional marginals are multivariate normal).

(b') $\mathbb{E}B_t=0$ and $\mathbb{E}(B_sB_t)=s\wedge t$

(c') $t\mapsto B_t$ is continuous a.s.

Pf: (a) + (b) $\Rightarrow$ (a'): for $t_0<t_1<t_2<\cdots<t_n$, the distribution of $(B_{t_0}, B_{t_1}, \cdots, B_{t_n})$ is determined by a Gaussian vector of independent entries $(B_{t_0}, B_{t_1}-B_{t_0}, B_{t_2}-B_{t_1}, \cdots, B_{t_n}-B_{t_{n-1}})$.

(a)+(b) $\Rightarrow$ (b'): if $s<t$, then $\mathbb{E}(B_sB_t)=\mathbb{E}(B_s(B_s+B_t-B_s)) = \mathbb{E}B_s^2+\mathbb{E}(B_s)\mathbb{E}(B_s-B_t)=s$.

(a')+(b') $\Rightarrow$ (a)+(b): $B(s+t)-B(s)$ is Gaussian distributed and $\mathbb{E}(B(s+t)-B(s))=0$,

$\mathbb{E}[(B(s+t)-B(s))^2] = \mathbb{E}[B(s+t)^2] - 2\mathbb{E}[B(s+t)B(s)] + \mathbb{E}[B(s)^2] = s+t-2s+s=t$.

$\mathbb{E}(B(t_0)(B(t_1)-B(t_0)))=0$, $\mathbb{E}[(B(t)-B(s))(B(s)-B(r))]=0$ $\forall$ $r\leq s\leq t$

$\Rightarrow$ Independent increments $\square$

Q: Does BM exist?

The idea is straightforward: we first construct the finite dimensional distributions using ~~the~~ multivariate normal distribution and then extend it ~~to~~ using the Kolmogorov's extension thm. However, Kolmogorov's extension thm only gives a measure on a countable sequence of RVs, while the index set for BM is uncountable. To deal with this issue, we adopt the usual strategy, that is, we ~~first~~ first construct a measure on a countable dense subset of $\mathbb{R}_+$ and then extend it to the whole $\mathbb{R}_+$ using the continuity requirement on BM.

Step 1: (Construct finite-dimensional measures)

Fix $x\in\mathbb{R}$ and $0<t_1<t_2<\cdots<t_n$, we define a measure on $\mathbb{R}^n$ as

$$\mu_{x,t_1,t_2,\cdots,t_n}(A_1\times A_2\times\cdots\times A_n) = \int_{A_1}dx_1\int_{A_2}dx_2\cdots\int_{A_n}dx_n\prod_{m=1}^{n}p_{t_m-t_{m-1}}(x_{m-1},x_m)$$

where A_i are measurable, $x_0=x$, $t_0=t$ and p_t is the heat kernel:

$$p_t(a,b) = (2\pi t)^{-\frac{1}{2}}\exp(-(b-a)^2/2t).$$

We observe that for each fixed x, the family μ is a consistent set of finite dimensional distributions, that is, $\forall$ $j\in\{1,\cdots,n\}$,

$$(*)\quad \mu_{x,t_1,\cdots,t_{j-1},t_{j+1},\cdots,t_n}(A_1\times\cdots\times A_{j-1}\times A_{j+1}\times\cdots\times A_n) = \mu_{x,t_1,\cdots,t_n}(A_1\times\cdots\times A_{j-1}\times\mathbb{R}\times A_{j+1}\times\cdots\times A_n)$$

This is trivial when $j=n$, because $p_{t_n-t_{n-1}}$ is a PDF.

The key point is that $(*)$ also holds for $1\leq j<n$. We need to show that

$$\int_{\mathbb{R}} p_{t_j-t_{j-1}}(x,y)\,p_{t_{j+1}-t_j}(y,z)\,dy = p_{t_{j+1}-t_{j-1}}(x,z).$$

The LHS is the PDF for $\mathcal{N}(x, t_j-t_{j-1}) + \mathcal{N}(0, t_{j+1}-t_j)$, independent sum.

The RHS is the PDF for $\mathcal{N}(x, t_{j+1}-t_{j-1})$.

Step 2: (Kolmogorov's extension) Let $Q_2=\{m2^{-n}: m, n\geq 0\}$ be the dyadic rationals.

Let $\Omega=\{\omega: Q_2\to\mathbb{R}\}$ and $\mathcal{F}_q$ is the σ-field generated by finite dimensional sets:

$$\{\omega_{t_1}\in A_1, \cdots, \omega_{t_n}\in A_n: t_1, t_2, \cdots, t_n\in Q_2\}.$$

Enumerating the dyadic rationals in Q_2, we have constructed a sequence of consistent probability measures indexed by them. Then, applying the Kolmogorov's extension thm, we can construct a probability measure ν_x on $(\Omega_q, \mathcal{F}_q)$ so that $\nu_x\{\omega: \omega(0)=x\}=1$ and

$$\nu_x\{\omega: \omega(t_i)\in A_i,\ i=1,\cdots,n\} = \mu_{x,t_1,\cdots,t_n}(A_1\times\cdots\times A_n).$$

Step 3: (Extension to a process defined on $\mathbb{R}_+=[0,\infty)$)

Theorem: Let $T<\infty$ and $x\in\mathbb{R}$. Then, ν_x assigns probability 1 to paths $\omega: Q_2\to\mathbb{R}$ that are uniformly continuous on $Q_2\cap[0,T]$.

Before proving this key theorem, we first discuss a key consequence of ~~this result~~ it.

Memo No.
Date / /

We can move ν_x to $(C^0, \mathcal{C})$ why $C^0 = \{\omega \text{ continuous}: [0,\infty) \to \mathbb{R}\}$ and $\mathcal{C}$ is the σ-field generated by the coordinate maps $t \mapsto \omega(t)$.

Let $\psi: \Omega_q \to C^0$ be the map that takes a uniformly continuous path/sample in Ω_q to its unique continuous extension, i.e., let $\omega \in \Omega_q$ be a uniformly continuous function. Then, $\psi(\omega): \mathbb{R}_+ \to \mathbb{R}$ is defined as $\psi(\omega)(x) = \lim_{n\to\infty} \omega(x_n)$, where x_n is a sequence of points in $\mathbb{Q}_2$ such that $x_n \to x$. We can show that ψ is measurable. Then, we let $\mathbb{P}_x = \nu_x \circ \psi^{-1}$ be the prob. measure on C^0.

Our construction guarantees that $B_t(\omega) = \omega_t$ has the right finite-dimensional distributions for $t \in \mathbb{Q}_2$. Then, using continuity of paths and a simple limiting argument, we find that properties (a) and (b) hold for $t \in [0,\infty)$.

Like the MC, we now have one set of RVs $B_t(\omega) = \omega_t$, and a family of probability measures $\mathbb{P}_x$, $x \in \mathbb{R}$, so that under $\mathbb{P}_x$, B_t is a BM with $\mathbb{P}_x(B_0 = x) = 1$.

Pf of the theorem: By translation invariance and scaling relation, we can take $T=1$ and $x=0$. Using the definition, we get $\mathbb{E}_0[|B_t - B_s|^n] = \mathbb{E}_0[|B_{t-s}|^n] \le C_n |t-s|^{n/2}$ $\forall$ fixed $n \ge 2$, $n \in \mathbb{N}$. Then, the following result due to Kolmogorov gives the Hölder continuity of the paths of BM.

Theorem † Suppose $\mathbb{E}|X_s - X_t|^\beta \le K|t-s|^{1+\alpha}$ where $\alpha, \beta > 0$, $t,s \in \mathbb{Q}_2$. If $\gamma < \alpha/\beta$, then with probability 1, there is a constant $C(\omega)$ so that

$$|X_q - X_r| \le C|q-r|^\gamma \quad \text{for all } q,r \in \mathbb{Q}_2 \cap [0,1].$$

Proof of Thm †: Let $G_n = \{|X(i2^{-n}) - X((i-1)2^{-n})| \le 2^{-\gamma n} \text{ for all } 0 < i \le 2^n\}$

By Chebyshev's inequality, $\mathbb{P}(G_n^c) \le 2^n \cdot (2^{-\gamma n})^{-\beta}\, \mathbb{E}|X(i2^{-n}) - X((i-1)2^{-n})|^\beta$

$$\le 2^{-n} \cdot 2^{\gamma n\beta} \cdot K 2^{-n(1+\alpha)} = K 2^{n(\gamma\beta - \alpha)} = K 2^{-n\lambda},$$

$\underbrace{\lambda = \alpha - \gamma\beta > 0.}$

Memo No.
Date / /

Lemma: On $H_N = \bigcap_{n=N}^{\infty} G_n$, we have $|X(q) - X(r)| \le \frac{2}{1-2^{-\gamma}}|q-r|^\gamma$ $\forall q,r \in \mathbb{Q}_2 \cap [0,1]$ with $\boxed{|q-r| \le 2^{-N}}$

Pf: Let $I_i^k = [(i-1)2^{-k}, i2^{-k})$ and let m be the smallest k for which q, r are in different intervals I_i^k. Since they were in the same interval on level $k-1$, $q \in I_i^m$ and $r \in I_{i+1}^m$ for some $0 \le i < 2^m$, we can write

$$r = i2^{-m} + 2^{-r(1)} + \cdots + 2^{-r(l)}$$

$$q = i2^{-m} - 2^{-q(1)} - \cdots - 2^{-q(k)}$$

for some $m < r(1) < \cdots < r(l)$ and $m < q(1) < \cdots < q(k)$. Then, on H_N, we have

$$|X(q) - X(i2^{-m})| \le \sum_{h=1}^{k} (2^{-q(h)})^\gamma \le \sum_{h=m}^{\infty} (2^{-\gamma})^h = \frac{2^{-\gamma m}}{1-2^{-\gamma}}$$

$$|X(r) - X(i2^{-m})| \le \frac{2^{-\gamma m}}{1-2^{-\gamma}}.$$

Then, $|X(q) - X(r)| \le \frac{2^{-\gamma m+1}}{1-2^{-\gamma}}$. We also have $|q-r| \ge 2^{-m}$. This concludes the proof. □

Now, to ~~conclude~~ conclude the proof of Theorem †, we have

$$\mathbb{P}(H_N^c) \le \sum_{n=N}^{\infty} \mathbb{P}(G_n^c) \le K \sum_{n=N}^{\infty} 2^{-\lambda n} = \frac{K 2^{-N\lambda}}{1-2^{-\lambda}}.$$

Thus, $\sum_{N=1}^{\infty} \mathbb{P}(H_N^c) < +\infty$, so by Borel–Cantelli, we get that for almost every ω, $\exists\, \delta(\omega) > 0$ s.t.

$$|X(q) - X(r)| \le A|q-r|^\gamma \quad \text{for } q,r \in \mathbb{Q}_2 \text{ with } |q-r| < \delta(\omega).$$

It remains to extend this to $q, r \in \mathbb{Q}_2 \cap [0,1]$. Let $s_0 = q < s_1 < s_2 < \cdots < s_n = r$ with $|s_i - s_{i-1}| < \delta(\omega)$ and $s_i \in \mathbb{Q}_2$. Using triangle inequality, we get that

$$|X(q) - X(r)| \le \sum_{i=0}^{n-1} |X(s_{i+1}) - X(s_i)| \le A \sum_{i=0}^{n-1} |s_{i+1} - s_i|^\gamma \le \frac{A}{\delta(\omega)^{1-\gamma}} \sum_{i=1}^{n} |s_{i+1} - s_i|$$

$$\le \frac{A}{\delta(\omega)^{1-\gamma}} |q-r| \le \frac{A}{\delta(\omega)^{1-\gamma}} |q-r|^\gamma. \quad \square$$

By the scaling relation, $\mathbb{E}|B_t - B_s|^{2m} = C_m |t-s|^m$, $C_m = \mathbb{E}|B_1|^{2m}$.

Thus, the above thm shows that a.s., ω is γ-Hölder continuous $\forall \gamma < \frac{m-1}{2m}$.

Letting $m \to +\infty$, we get:

Thm: (Wiener 1923) Brownian paths are Hölder continuous $\forall$ exponent $\gamma < 1/2$.

Memo No.
Date / /

Theorem With prob. one, Brownian paths are not Lipschitz continuous (and hence not differentiable) at any point.

Pf: Fix a constant $C>0$. Let $A_n = \{w:$ there is an $s\in[0,1]$ so that $|B_t - B_s| \le C|t-s|$ when $|t-s| \le \frac{3}{n}\}$. For $1\le k\le n-2$, let

$$Y_{k,n} = \max\left\{ \left| B\left(\tfrac{k+j}{n}\right) - B\left(\tfrac{k+j-1}{n}\right)\right| : j=0,1,2 \right\}.$$

$$B_n = \{ \text{at least one } Y_{k,n} \le 5C/n \}.$$ By triangle ineq, we have $A_n \subset B_n$

The worst case is $s=1$. Picking $k=n-2$, we observe that

$$\left|B\left(\tfrac{n-3}{n}\right) - B\left(\tfrac{n-2}{n}\right)\right| \le \left|B\left(\tfrac{n-3}{n}\right) - B(1)\right| + \left|B(1) - B\left(\tfrac{n-2}{n}\right)\right| \le \frac{5C}{n}.$$

Using $A_n \subset B_n$, we get $\mathbb{P}(A_n) \le \mathbb{P}(B_n) \le n\,\mathbb{P}(|B(1/n)| \le 5C/n)^3$

$$= n\,\mathbb{P}(|B(1)| \le 5C/\sqrt{n})^3 \le Cn^{-\frac{1}{2}}.$$

Letting $n\to\infty$, we get $\mathbb{P}(A_n)\to 0$. But $\overset{n}{A} \to A_n$ is increasing. Hence, we must have $\mathbb{P}(A_n)=0$, which completes the proof. □

Multi-dimensional BM: With 1D BM we just constructed, we can define d-dimensional BM with $d\ge 2$. A d-dim BM starting at $\vec{x}\in\mathbb{R}^d$ is defined as $\vec{B}_t = (B_t^1, \cdots, B_t^d)$, where $B_t^1, \cdots, B_t^d$ are d independent BMs with $B_0^i = x^i$. In this case, the transition prob is the heat kernel in $\mathbb{R}^d$: $p_t(x,y) = (2\pi t)^{-d/2}\exp\left(-\frac{(x-y)^2}{2t}\right)$, i.e., for $0<t_1<\cdots<t_n$,

$$\mathbb{P}_{\vec{x}}(\vec{B}_{t_0}\in A_0, \dots, \vec{B}_{t_n}\in A_n) = \int_{A_0} d\vec{x}_0 \int_{A_1} d\vec{x}_1 \cdots \int_{A_n} d\vec{x}_n \prod_{m=1}^{n} p_{t_m - t_{m-1}}(\vec{x}_{m-1}, \vec{x}_m),$$

with $\vec{x}_0 = \vec{x}$ and $t_0 = 0$.

2. Markov Property

MP says that "$B(t+s) - B(s)$ is a BM that is independent of what happened before time s". To make this statement rigorous, we first need to make sense of "what happened before time s". The most natural choice is $\mathcal{F}_s^0 = \sigma(\{B_r : r\le s\})$. However, it is more convenient to consider a slightly larger σ-field:

$\mathcal{F}_s^+ = \bigcap_{t>s} \mathcal{F}_t^0 \supseteq \mathcal{F}_s^0$. This field is nicer because it is right continuous:

$$\bigcap_{t>s}\mathcal{F}_t^+ = \bigcap_{t>s}\left(\bigcap_{u>t}\mathcal{F}_u^0\right) = \bigcap_{u>s}\mathcal{F}_u^0 = \mathcal{F}_s^+.$$

Intuitively, $\mathcal{F}_s^+$ allows for an "infinitesimal peek at the future", i.e., $A\in\mathcal{F}_s^+$ if it is in $\mathcal{F}_{s+\varepsilon}^0$ $\forall\,\varepsilon>0$. If $f(u)>0$ $\forall\,u>0$, then $\limsup_{t\downarrow s}\frac{B_t - B_s}{f(t-s)}$ is measurable w.r.t. $\mathcal{F}_s^+$ but not $\mathcal{F}_s^0$. We will see below that $\mathcal{F}_s^+$ and $\mathcal{F}_s^0$ are the same up to null sets.

We have defined a family of measures $\mathbb{P}_x$, $x\in\mathbb{R}^d$, on $(C(\mathbb{R}^d), \mathcal{C})$ so that under $\mathbb{P}_x$, $B_t(w) = w(t)$ is a BM starting at x. For $s\ge 0$, we define the shift transformation $\theta_s: C(\mathbb{R}^d)\to C(\mathbb{R}^d)$ by $(\theta_s w)(t) = w(t+s)$ for $t\ge 0$.

Then, we define the MP similar to what we have done for MCs.

Theorem (MP for BM) If $s\ge 0$ and Y is bounded and $\mathcal{C}$-measurable, then $\forall\,x\in\mathbb{R}^d$,

$$\mathbb{E}_x(Y\circ\theta_s \mid \mathcal{F}_s^+) = \mathbb{E}_{B_s} Y,$$

where the RHS is the function $\varphi(x) = \mathbb{E}_x Y$ evaluated at $x = B_s$.

Pf: We need to show that $\mathbb{E}_x(Y\circ\theta_s; A) = \mathbb{E}(\mathbb{E}_{B_s}Y; A)$ for all $A\in\mathcal{F}_s^+$.

We prove this equation for a carefully chosen special case and then use the monotone class thm to conclude the general case. Suppose $Y(w) = \prod_{m=1}^{n} f_m(w(t_m))$, where $0<t_1<t_2<\cdots<t_n$ and the functions f_m are bounded and measurable.

Let $0<h<t_1$, $0<s_1<s_2<\cdots<s_k\le s+h$, and $A = \{w: w(s_j)\in A_j,\ 1\le j\le k\}$, A_j Borel measurable.

By def. of BM,

$$\mathbb{E}_x(Y\circ\theta_s; A) = \mathbb{E}\left(\prod_{m=1}^{n} f_m(w(t_m+s)); A\right) = \int_{A_1} dx_1\, p_{s_1}(x, x_1)\int_{A_2} dx_2\, p_{s_2-s_1}(x_1, x_2)\cdots$$

$$\int_{A_k} dx_k\, p_{s_k - s_{k-1}}(x_{k-1}, x_k)\int dy\, p_{s+h-s_k}(x_k, y)\cdot\varphi(y, h),$$

Memo No.
Date / /

where $\varphi(y,h) = \int dy_1\, p_{t_1-h}(y,y_1)\, f_1(y_1) \int dy_2\, p_{t_2-t_1}(y_1,y_2)\, f_2(y_2) \cdots \int dy_n\, p_{t_n-t_{n-1}}(y_{n-1},y_n)\, f_n(y_n)$.

We notice that the RHS can be written as $\mathbb{E}_x(\varphi(B_{s+h},h);A)$.

Thus, we get $\mathbb{E}_x[Y\circ\theta_s;A] = \mathbb{E}_x[\varphi(B_{s+h},h);A]$. This holds for all finite dimensional sets A, so the $\pi-\lambda$ thm implies that this identity holds for all $A\in\mathcal{F}^o_{s+h}\supset\mathcal{F}^+_s$.

Denote $\psi(y_1) = f_1(y_1)\int dy_2\, p_{t_2-t_1}(y_1,y_2)\, f_2(y_2)\cdots\int dy_n\, p_{t_n-t_{n-1}}(y_{n-1},y_n)\, f_n(y_n)$. It is easy to check that ψ is bdd measurable. Let $\varphi(x,h) = \int dy_1\, p_{t_1-h}(x,y_1)\,\psi(y_1)$.

Letting $h\to 0$ and using DCT, we get that if $x_h\to x$, then $\varphi(x_h,h)\to\varphi(x,0)$ as $h\to 0_+$.

Then ~~taking~~ we take $h\to 0$ in $\mathbb{E}_x(Y\circ\theta_s;A) = \mathbb{E}_x(\varphi(B_{s+h},h);A)$. Using DCT, we get that

$$\mathbb{E}_x(Y\circ\theta_s;A) = \mathbb{E}_x(\underbrace{\varphi(B_s,0)}_{\mathbb{E}_{B_s}Y};A)\quad\forall A\in\mathcal{F}^+_s.$$

$\overset{(++)}{\Rightarrow}$ $\mathbb{E}_x(Y\circ\theta_s;A) = \mathbb{E}_x(\mathbb{E}_{B_s}Y;A)$ $\forall A\in\mathcal{F}^+_s$, for $Y=\prod_{m=1}^n f_m(w(t_m))$ with f_m being bdd measurable functions.

Next, we apply the MCT: $\mathcal{H} = \{$bounded measurable fcns for which (++) holds$\}$. $\mathcal{H}$ satisfies (ii) and (iii). Let $\mathcal{A} = \{$sets of the form $\{w: w(t_j)\in A_j\}$, A_j – Borel measurable$\}$. The special case treated above shows that $\mathbb{1}_A\in\mathcal{H}$ $\forall A\in\mathcal{A}$. The conclusion follows. □

Note that $\mathbb{E}_x(Y\circ\theta_s|\mathcal{F}^+_s) = \mathbb{E}_{B(s)}Y\in\mathcal{F}^o_s \Rightarrow \mathbb{E}_x(Y\circ\theta_s|\mathcal{F}^+_s) = \mathbb{E}(Y\circ\theta_s|\mathcal{F}^o_s)$. Then:

<u>Theorem</u>: If $Z\in\mathcal{C}$ is bounded, then $\forall s\geq 0$ and $x\in\mathbb{R}^d$,

$$\mathbb{E}_x(Z|\mathcal{F}^+_s) = \mathbb{E}_x(Z|\mathcal{F}^o_s).$$

Pf: Similar to the pf of the previous thm, it suffices to prove the result when $Z=\prod_{m=1}^n f_m(B(t_m))$.

Then, Z can be written as $Z=(Y\circ\theta_s)\cdot(X)$, where X is $\mathcal{F}^o_s$ measurable and $Y\in\mathcal{C}$.

Thus, $\mathbb{E}_x(Z|\mathcal{F}^+_s) = X\,\mathbb{E}_x(Y\circ\theta_s|\mathcal{F}^+_s) = X\,\mathbb{E}_{B_s}Y\in\mathcal{F}^o_s$. □

As a special case, if $Z\in\mathcal{F}^+_s$, then this thm implies $Z=\mathbb{E}_x(Z|\mathcal{F}^+_s) = \mathbb{E}_x(Z|\mathcal{F}^o_s)\in\mathcal{F}^o_s$.

Memo No.
Date / /

Hence, $\mathcal{F}^o_s$ and $\mathcal{F}^+_s$ are the same up to null sets (sets with measure zero).

<u>Thm</u>: (Blumenthal's 0-1 law) If $A\in\mathcal{F}^+_0$, then $\forall x\in\mathbb{R}^d$, $\mathbb{P}_x(A)\in\{0,1\}$.

Pf: $\forall A\in\mathcal{F}^+_0$, we have $\mathbb{1}_A = \mathbb{E}_x(\mathbb{1}_A|\mathcal{F}^+_0) = \mathbb{E}_x(\mathbb{1}_A|\mathcal{F}^o_0) = \mathbb{P}_x(A)$ $\mathbb{P}_x$-a.s. where we use that $\mathcal{F}^o_0$ is trivial under $\mathbb{P}_x$. □

$\mathcal{F}^+_0$: the germ field. This says that $\mathcal{F}^+_0$ is trivial. It is useful in studying the local behavior of Brownian paths.

<u>Thm</u>: For 1D BM, $\mathbb{P}_0(\tau=0)=1$ where $\tau := \inf\{t\geq 0: B_t>0\}$.

Pf: We have $\mathbb{P}_0(\tau\leq t) \geq \mathbb{P}_0(B_t>0) = 1/2$. Letting $t\downarrow 0$, we get $\mathbb{P}_0(\tau=0) = \lim_{t\downarrow 0}\mathbb{P}_0(\tau\leq t)\geq\frac{1}{2}$. Since $\{\tau=0\}$ is in the germ field $\mathcal{F}^+_0$, we have $\mathbb{P}_0(\tau=0)=1$. □

By symmetry, BM must also hit $(-\infty,0)$ immediately, i.e., $\mathbb{P}(\tau_-=0)=1$ for $\tau_-=\inf\{t\geq 0: B_t<0\}$.

Since B_t is continuous in t, we get:

~~Thm~~ <u>Thm</u>: Let $T_0 = \inf\{t>0: B_t=0\}$. Then, $\mathbb{P}_0(T_0=0)=1$.

Using the following time inversion trick, we can use the above results to get information about the behavior as $t\to+\infty$.

<u>Thm</u>: If B_t is a BM starting at 0, then so is the process defined by $X_0=0$ and $X_t = tB(1/t)$ for $t>0$.

Pf: We check the second definition of BM. (i) If $0<t_1<t_2<\cdots$, then $(X(t_1),\cdots,X(t_n))$ has multivariate normal distribution with mean 0. (ii) $\mathbb{E}X_s=0$ and if $s<t$, then $\mathbb{E}(X_tX_s) = ts\,\mathbb{E}(B(1/t)B(1/s)) = ts\cdot\frac{1}{t}=s$. (iii) X_t is continuous at $t\neq 0$. We still need to check the continuity at $t=0$. We need to show that $B(t)/t\to 0$ a.s. as $t\to+\infty$.

By the SLLN, we have that $B_n/n\to 0$ a.s. as $n\to\infty$ through integers. To handle values in between integers, we use Kolmogorov's maximal inequality to conclude that

$$\mathbb{P}\Big(\sup_{0\le k\le 2^m} |B(n+k2^{-m}) - B(n)| > n^{2/3}\Big) \le n^{-\frac{4}{3}}\, \mathrm{Var}(B_{n+1}-B_n) = n^{-\frac{4}{3}}.$$

Letting $m\to\infty$, we have $\mathbb{P}\big(\sup_{u\in[n,n+1]} |B_u - B_n| > n^{2/3}\big) \le n^{-\frac{4}{3}}$

Since $\sum_n n^{-\frac{4}{3}} < +\infty$, by B-C, we get that $B_u/u \to 0$ a.s. as $u\to+\infty$. □

The above thm allows us to relate the behavior of B_t as $t\to\infty$ and as $t\to 0$.

Together with Blumenthal's 0-1 law, it gives some interesting results.

Let $\mathcal{F}_t' = \sigma(B_s : s\ge t)$ = the future at time t, $\mathcal{T} = \bigcap_{t\ge 0} \mathcal{F}_t'$ = the tail σ-field.

Thm: If $A\in\mathcal{T}$, then ~~either $\mathbb{P}_x(A)\in\{0,1\}\ \forall x\in\mathbb{R}$~~ either $\mathbb{P}_x(A)\equiv 0$ or $\mathbb{P}_x(A)\equiv 1$ $\forall x\in\mathbb{R}$.

Rmk: This form is stronger that Blumenthal's 0-1 law. In that case, $\mathbb{P}_x(A)$ may depend on x.

Pf: The tail σ-field of B_t is the same as the germ σ-field of X_t. Hence, we have $\mathbb{P}_x(A)\in\{0,1\}$. To improve this, observe that $A\in\mathcal{F}_1'$, so $\mathbb{1}_A = \mathbb{1}_D\circ\theta_1$ for some event D. Applying the Markov property, we get

$$\mathbb{P}_x(A) = \mathbb{E}_x(\mathbb{1}_D\circ\theta_1) = \mathbb{E}_x\big(\mathbb{E}_x(\mathbb{1}_D\circ\theta_1 \mid \mathcal{F}_1^+)\big) = \mathbb{E}_x(\mathbb{E}_{B_1}\mathbb{1}_D)$$
$$= \frac{1}{\sqrt{2\pi}}\int \mathbb{P}_y(D)\, e^{-\frac{(y-x)^2}{2}}\, dy.$$

Taking $x=0$, we observe that if $\mathbb{P}_0(A)=0$, then $\mathbb{P}_y(D)=0$ for a.e. y w.r.t. to the Lebesgue measure. Using the above formula again, we get $\mathbb{P}_x(A)=0$ for all x.

If $\mathbb{P}_0(A)=1$, then $A^c\in\mathcal{T}$ with $\mathbb{P}_0(A^c)=0$. The above argument shows that $\mathbb{P}_x(A^c)=0$ for all x. □

As an application of the above result:

Thm: Let B_t be a 1D BM starting at 0. Then with prob. 1,
$$\limsup_{t\to\infty} B_t/\sqrt{t} = \infty, \quad \liminf_{t\to\infty} B_t/\sqrt{t} = -\infty.$$

Pf: Let $0<K<+\infty$. We have
$$\mathbb{P}_0(B_n/\sqrt{n} > K \text{ i.o.}) \ge \limsup_{n\to+\infty} \mathbb{P}(B_n > K\sqrt{n}) = \mathbb{P}_0(B_1 > \sqrt{K}) > 0.$$

So the 0-1 law tells that $\mathbb{P}_0(B_n > K\sqrt{n} \text{ i.o.}) = 1$. Since K is arbitrary, we get the first result. □

Thm: Let B_t be a 1D BM and let $A = \bigcap_n \{B_t = 0 \text{ for some } t\ge n\}$. Then, $\mathbb{P}_x(A)=1$ for all x.

Pf: This follows from the previous thm, translation invariance, and continuity of Brownian paths. □

In other words, this says that 1D BM is "recurrent". For any starting point x, it will return to 0 i.o., i.e., there is a sequence of times $t_n\uparrow\infty$ so that $B_{t_n}=0$.

The Blumenthal's 0-1 law tells that $\mathcal{F}_s^+$ and $\mathcal{F}_s^0$ are identical up to null sets.

Let $\mathcal{N}_x = \{A : A\subseteq D \text{ with } \mathbb{P}_x(D)=0\}$, $\mathcal{F}_s^x = \sigma(\mathcal{F}_s^+\cup\mathcal{N}_x)$, $\mathcal{F}_s = \bigcap_x \mathcal{F}_s^x$.

$\mathcal{N}_x$ are the null sets and $\mathcal{F}_s^x$ are the completed σ-fields for $\mathbb{P}_x$. Since we do not want the filtration to depend on the initial state, we take the intersection of all σ-fields.

By def, $\mathcal{F}_s$ are right-continuous, i.e., $\mathcal{F}_s = \bigcap_{t>s}\mathcal{F}_t$.

3. Strong Markov Property

A RV T taking values in $[0,\infty]$ is called a stopping time if $\forall t\ge 0$, $\{T<t\}\in\mathcal{F}_t$. ~~An equivalent~~ Another def is to require that $\{T\le t\}\in\mathcal{F}_t$. These two defs are equivalent in continuous time for a right continuous filtration $\mathcal{F}_t$:

(1) If $\{T\le t\}\in\mathcal{F}_t$, then $\{T<t\} = \bigcup_n\{T\le t-\frac{1}{n}\}\in\mathcal{F}_t$.

(2) If $\{T<t\}\in\mathcal{F}_t$, then $\{T\le t\} = \bigcap_n\{T<t+\frac{1}{n}\}\in\mathcal{F}_t$ (by using that $\mathcal{F}_t$ is right continuous.)

Thm: (Hitting time) If G is an open set and $T=\inf\{t\ge 0: B_t\in G\}$, then T is a stopping time.

Pf: Since G is open and $t\mapsto B_t$ is continuous, we have $\{T<t\} = \bigcup_{q<t,\, q\in\mathbb{Q}}\{B_q\in G\}\in\mathcal{F}_t$. □

Thm: If T_n is a sequence of stopping times and $T_n \downarrow T$, then T is a stopping time.

Pf: $\{T<t\} = \bigcup_n \{T_n < t\}$. □

Thm: If T_n is a sequence of stopping times and $T_n \uparrow T$, then T is a stopping time.

Pf: $\{T\le t\} = \bigcap_n \{T_n \le t\}$. □

Thm: (Hitting time 2) If K is a closed set and $T = \inf\{t\ge 0: B_t \in K\}$, then T is a stopping time.

Pf: Let $B(x,r) = \{y: |y-x| \le r\}$ be the ball of radius r around x. Let $G_n = \bigcup_{x\in K} B(x, 1/n)$ and $T_n := \inf\{t\ge 0: B_t \in G_n\}$. Note G_n is open, so T_n is a stopping time. We claim that $T_n \uparrow T$. First, notice that $T_n \le T$ $\forall n$, so $\lim_{n\to\infty} T_n \le T$. On the other hand, suppose $T_n \uparrow S < \infty$. Since $B(T_n) \in \bar{G}_n$ for all n and $B(T_n) \to B(S)$, we get that $B(S) \in K$ (since K is closed) and $T \le S$. □

Combining the above results, we see that the hitting time of a countable union / countable intersection of closed / open sets are stopping times.

In fact, the hitting time of any Borel set is a stopping time, but proving that is difficult.

Thm: Let S be a stopping time and let $S_n = ([2^n S]+1)/2^n$ ($[\cdot]$: integer part). Then, S_n are stopping times.

[$S_n = (m+1)2^{-n}$ if $m2^{-n} \le S < (m+1)2^{-n}$, i.e., we stop at the first time of the form $k2^{-n}$ after S (i.e., $>S$)]

Pf: If $m2^{-n} < t \le (m+1)2^{-n}$, then $\{S_n < t\} = \{S \le m2^{-n}\} \in \mathcal{F}_{m2^{-n}} \in \mathcal{F}_t$. □

Given a non-negative RV $S(\omega)$, we define the random shift operator θ_S, which "cuts off the part of ω before $S(\omega)$ and shift the time axis to the right by $S(\omega)$".

In formulas, we write

$$(\theta_S \omega)(t) = \begin{cases} \omega(S(\omega)+t) & \text{on } \{S<+\infty\} \\ \Delta & \text{on } \{S=\infty\} \end{cases}$$

where Δ is an extra point added to $C(\mathbb{R}_+)$. But, we will always explicitly restrict our attension to $\{S<\infty\}$.

Next, we define the σ-field $\mathcal{F}_S$ — "the information known up to time S":

$$\mathcal{F}_S = \{A: A\cap\{S\le t\} \in \mathcal{F}_t \text{ for all } t\ge 0\}$$

When $\mathcal{F}_t$ is right-continuous, this def. is unchanged if we replace $\{S\le t\}$ by $\{S<t\}$.

Thm: If $S \le T$ are stopping times, the $\mathcal{F}_S \subseteq \mathcal{F}_T$.

Pf: If $A\in\mathcal{F}_S$, then $A\cap\{T\le t\} = (A\cap\{S\le t\})\cap\{T\le t\} \in \mathcal{F}_t$. □

Thm: If $T_n \downarrow T$ are stopping times, then $\mathcal{F}_T = \bigcap_n \mathcal{F}_{T_n}$.

Pf: The previous thm tells that $\mathcal{F}_T \subset \mathcal{F}_{T_n}$ $\forall n \Rightarrow \mathcal{F}_T \subseteq \bigcap_n \mathcal{F}_{T_n}$. On the other hand, let $A\in \bigcap_n \mathcal{F}_{T_n}$. Since $A\cap\{T_n<t\}\in\mathcal{F}_t$ and $T_n\downarrow T$, $A\cap\{T<t\} = \bigcup_n (A\cap\{T_n<t\}) \in \mathcal{F}_t$. □

Thm: Let S be a stopping time. Then, $B_S \in \mathcal{F}_S$, i.e., the value B_S is measurable w.r.t. the information known up to time S.

Pf: Let $S_n = ([2^n S]+1)/2^n$ be stopping times defined before. If A is a Borel set, we have $\{B(S_n)\in A\} = \bigcup_{m=1}^{\infty} \{S_n = m/2^n, B(m/2^n)\in A\} \in \mathcal{F}_{S_n}$.

Let $n\to\infty$, we have $S_n \downarrow S$ and $B(S) = \lim_{n\to\infty} B(S_n) \in \bigcap_n \mathcal{F}_{S_n} = \mathcal{F}_S$. □

Now, we state the strong MP, which says that MP holds at stopping times.

☆ Thm (Strong MP) Let $(s,\omega) \to Y_s(\omega)$ be bounded and $\mathcal{R}\times\mathcal{C}$ measurable. If S is a stopping time, then $\forall x\in\mathbb{R}^d$, $\mathbb{E}_x(Y_S\circ\theta_S \mid \mathcal{F}_S) = \mathbb{E}_{B_S} Y_S$ on $\{S<+\infty\}$, where the RHS is the function $\varphi(x,t) = \mathbb{E}_x Y_t$ evaluated at $x = B_S$ and $t = S$.

Rmk: The proof only uses the following two properties of BM: (i) it is a Markov process; (ii) (the Feller property) if f is a bounded continuous function, then $x \mapsto \mathbb{E}_x f(B_t)$ is continuous. Hence, the SMP holds for general Markov processes satisfying the Feller property.

Pf. ~~We first~~ ~~"discrete~~ The basic idea is to "discretize" the stopping time.

We first consider a simplified assumption that there is a sequence of times $t_n \uparrow \infty$, so that $\mathbb{P}_x(S<\infty) = \sum_x \mathbb{P}_x(S=t_n)$. Then, the SMP follows directly from the MP.

Let $Z_n = Y_{t_n}(\omega)$ and $A\in\mathcal{F}_S$, then $\mathbb{E}_x(Y_S\circ\theta_S; A\cap\{S<\infty\}) = \sum_{n=1}^{\infty} \mathbb{E}_x(Z_n\circ\theta_{t_n}; A\cap\{S=t_n\})$ (*)

If $A\in\mathcal{F}_S$, then $A\cap\{S=t_n\} = A\cap\{S\le t_n\} - A\cap\{S<t_n\}\in\mathcal{F}_{t_n}$. Hence, using the MP, we get that

$(*) = \sum_{n=1}^{\infty} \mathbb{E}_x(\mathbb{E}_{B(t_n)} Z_n; A\cap\{S=t_n\}) = \mathbb{E}_x(\mathbb{E}_{B(S)}Y_S; A\cap\{S<+\infty\})$.

To prove the result in general, we let $S_n = ([2^nS]+1)/2^n$ be the stopping time defined before.

To let $n\to+\infty$, we restrict our attention to Y of the form $Y_s(\omega) = f_0(s)\prod_{m=1}^{n} f_m(\omega(t_m))$, where $0<t_1<t_2<\cdots<t_n$ and $f_0,\dots,f_n$ are bounded continuous. Then, we can check that

$\varphi(x,s) = \mathbb{E}_x Y_s = f_0(s)\int dy_1\, p_{t_1}(x,y_1)f_1(y_1)\cdots\int dy_n\, p_{t_n-t_{n-1}}(y_{n-1},y_n)f_n(y_n)$ is bdd continuous.

Next, we complete the proof. Let $A\in\mathcal{F}_S$. Since $S\le S_n$, we have $A\in\mathcal{F}_S\subset\mathcal{F}_{S_n}$. Applying the special case proved above to S_n and observing that $\{S<\infty\}=\{S_n<\infty\}$, we get that

$$\mathbb{E}_x(Y_{S_n}\circ\theta_{S_n}; A\cap\{S<\infty\}) = \mathbb{E}_x(\varphi(B(S_n),S_n); A\cap\{S<\infty\})$$

As $n\to\infty$, $S_n\downarrow S$, $B(S_n)\to B(S)$, $\varphi(B(S_n),S_n)\to\varphi(B(S),S)$ and $Y_{S_n}\circ\theta_{S_n}\to Y_S\circ\theta_S$.

Then, by the DCT, we get that $\mathbb{E}_x(Y_S\circ\theta_S; A\cap\{S<\infty\}) = \mathbb{E}_x(\varphi(B(S),S); A\cap\{S<\infty\})$ holds when Y takes the form $Y_s(\omega)=f_0(s)\prod_{m=1}^{n} f_m(\omega(t_m))$.

Finally, we apply the Monotone class theorem. Let $\mathcal{H}$ = collection of Y for which

$\mathbb{E}_x(Y_S\circ\theta_S; A) = \mathbb{E}_x(\mathbb{E}_{B(S)}Y_S; A)\ \forall A\in\mathcal{F}_S$.

It is easy to see that
(ii) If $f,g\in\mathcal{H}$, then $f+g\in\mathcal{H}$ and $cf\in\mathcal{H}\ \forall c\in\mathbb{R}$.
(iii) If $f_n\in\mathcal{H}$ are monotone and $f_n\uparrow f$, then $f\in\mathcal{H}$.

To show (i) for some $\mathcal{A}$ that generates the σ-field $\mathcal{R}\times\mathcal{C}$, we let $\mathcal{A}$ to be the collection of sets of the form $A = G_0\times\{\omega: \omega(s_j)\in G_j, 1\le j\le k\}$, where

$s_1<s_2<\cdots$ and G_j are open sets. We let $K_j = G_j^c$ and $f_j^n(x) = 1\wedge n\rho(x,K_j)$, where $\rho(x,K)=\inf\{|x-y|: y\in K\}$. Then, f_j^n are bounded continuous functions with $f_j^n\uparrow 1_{G_j}$.

But, we have shown that $Y_s^n(\omega) = f_0^n(s)\prod_{j=1}^{k} f_j^n(\omega(s_j))\in\mathcal{H}$ and (iii) holds. This implies that $1_A\in\mathcal{H}$, and verifies (i) in the MCT. □

4. Path Properties

We will use the SMP to derive some path properties for 1D BM.

① Zeros of BM

Let $R_t = \inf\{u>t: B_u=0\}$ and let $T_0=\inf\{u>0: B_u=0\}$. We already know that $\mathbb{P}_x(R_t<\infty)=1$, so $B(R_t)=0$ a.s. and the SMP implies that

$$\mathbb{P}_x(T_0\circ\theta_{R_t}>0 \mid \mathcal{F}_{R_t}) = \mathbb{P}_0(T_0>0)=0 \Rightarrow \mathbb{P}_x(T_0\circ\theta_{R_t}>0 \text{ for some rationl } t)=0 \quad (*)$$

Let $Z(\omega)=\{t: B_t(\omega)=0\}$ be the set of zeros of ω.

Then, $Z(\omega)$ is a closed set. Moreover, it has no isolated point (i.e., $Z(\omega)$ is a perfect set).

To see this, we notice that (*) implies: if a point $u\in Z(\omega)$ is isolated frome the left (i.e., there is a rational $t<u$ so that $(t,u)\cap Z(\omega)=\phi$), then a.s., it is a decreasing limit of points in $Z(\omega)$.

~~Topo~~ We know from the point topology that a perfect set is uncountable. Hence, $Z(\omega)$ is uncountable.

Moreover, by Fubini's thm, $\forall T>0$,

$$\mathbb{E}_x\big(|Z(\omega)\cap[0,T]|\big) = \int_0^T \mathbb{P}_x(B_t=0)\,dt = 0$$

So, $Z(\omega)$ is also a measure zero set. (In fact, Z is a fractal set of dimension 1/2.)

② Hitting times: $T_a=\inf\{t>0: B_t=a\}$. Our goal is to derive the distribution $\forall a$.

By translation symmetry, it suffices to study this problem for $a=0$.

Memo No.
Date / /

Thm: Under $\mathbb{P}_0$, $\{T_a : a \geq 0\}$ has stationary independent increments.

Pf: Notice that if $0 < a < b$, then $T_a < T_b$ and $T_b \circ \theta_{T_a} = T_b - T_a$. Hence, $\forall$ bdd measurable function f, using the SMP and translation invariance, we get

$$\mathbb{E}_0 (f(T_b - T_a) \mid \mathcal{F}_{T_a}) = \mathbb{E}_0 (f(T_b) \circ \theta_{T_a} \mid \mathcal{F}_{T_a}) = \mathbb{E}_a f(T_b) = \mathbb{E}_0 f(T_{b-a}).$$

This gives stationarity of increments. It remains to prove independence.

Let $a_0 < a_1 < \cdots < a_n$, and let $f_1, f_2, \cdots, f_n$ be bounded measurable. Define $F_i = f_i(T_{a_i} - T_{a_{i-1}})$.

Conditioning on $\mathcal{F}_{T_{a_{n-1}}}$, we get

$$\mathbb{E}_0 f_n(T_{a_n - a_{n-1}}) = \mathbb{E}_0 f_n(T_{a_n} - T_{a_{n-1}}) = \mathbb{E}_0 F_n$$

$$\mathbb{E}_0 \Big(\prod_{i=1}^{n} F_i \Big) = \mathbb{E}_0 \Big(\prod_{i=1}^{n} F_i \,\Big|\, \mathcal{F}_{T_{a_{n-1}}} \Big) = \mathbb{E}_0 \Big(\prod_{i=1}^{n-1} F_i \cdot \mathbb{E}_0 (F_n \mid \mathcal{F}_{T_{a_{n-1}}}) \Big)$$

$$= \mathbb{E}_0 \Big(\prod_{i=1}^{n-1} F_i \Big) \mathbb{E}_0 F_n.$$

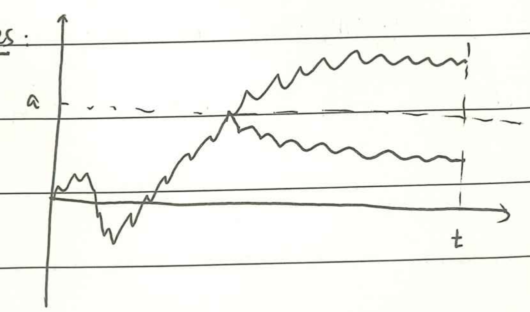
By induction, we get $\mathbb{E}_0 \prod_{i=1}^{n} F_i = \prod_{i=1}^{n} \mathbb{E}_0 F_i$, which implies the independence. $\square$

~~The~~ Let $t_k = T_k - T_{k-1}$. The above results implies that $t_1, t_2, t_3, \cdots$ are i.i.d. and each of them has the same distribution as T_1. On the other hand, by scaling relation, we have $T_n \overset{d}{=} n^2 T_1$. Hence, we have $\frac{t_1 + t_2 + \cdots + t_n}{n^2} \to T_1$ in law. This shows that T_1 has stable law (and hence any T_a has stable law). Using the above scaling ~~that~~ and the fact that $T_a \geq 0$, we can determine the law of T_1 uniquely. Here we give a direct calculation of the distribution of T_a using the <u>reflection principle</u>.

Theorem: $\forall a > 0$, we have $\mathbb{P}_0 (T_a < t) = 2 \mathbb{P}_0 (B_t > a)$.

Rmk: Since BM is continuous, in contrast to the reflection principle for MC, we have "=" here.

Pf by pictures:



If B_s hits a at some time $s < t$, then SMP implies that $B_t - B_{T_a}$ is indep. of what happened before T_a. Using symmetry of normal distribution and $\mathbb{P}_a (B_u = a) = 0$ $\forall u > 0$,

Memo No.
Date / /

we get $\mathbb{P}_0 (T_a < t, B_t > a) = \frac{1}{2} \mathbb{P}_0 (T_a < t) \Rightarrow \mathbb{P}_0 (T_a < t) = 2 \mathbb{P}_0 (B_t > a)$

$\qquad\qquad\qquad \| $

$\qquad \mathbb{P}_0 (B_t > a)$

Mathematical proof: We let $Y_s(\omega) = \begin{cases} 1, & \text{if } s < t \text{ and } \omega(t-s) \geq a \\ 0, & \text{otherwise} \end{cases}$

Define $S = \inf\{ s < t : B_s = a \}$ with $\inf \phi = \infty$. Then, we have $Y_S \circ \theta_S(\omega) = \begin{cases} 1, & \text{if } S < t,\ B_t(\omega) \geq a \\ 0, & \text{otherwise} \end{cases}$

Then, using the SMP, we get

$$\mathbb{E}_0 (Y_S \circ \theta_S \mid \mathcal{F}_S) = \mathbb{E}_{B_S} Y_S \quad \text{on} \quad \{ S < \infty \} = \{ T_a < t \}.$$

$$= \mathbb{E}_a Y_S = \frac{1}{2} \quad \text{on} \quad \{ S < \infty \} = \{ T_a < t \}$$

~~$\Rightarrow \mathbb{E}_0 (\mathbb{E}_0 (Y_S \circ \theta_S \mid \mathcal{F}_S)) = \mathbb{E}_0 (Y_S \circ \theta_S)$~~ Taking expectation over $\{ S < \infty \}$ of both sides:

$$\text{LHS} = \mathbb{E}_0 (\mathbb{E}_0 (Y_S \circ \theta_S \mid \mathcal{F}_S) \mathbb{1}_{S < +\infty}) = \mathbb{E}_0 (Y_S \circ \theta_S ; S < +\infty) = \mathbb{P}_0 (T_a < t, B_t \geq a)$$

$$\text{RHS} = \frac{1}{2} \mathbb{E}_0 (\mathbb{1}_{S < +\infty}) = \frac{1}{2} \mathbb{P}_0 (T_a < t). \quad \square$$

We have calculated that $\mathbb{P}(T_a \leq t) = 2 \mathbb{P}_0 (B_t > a) = \int_a^\infty \frac{2}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}} dx$

$$= \int_{a/\sqrt{t}}^\infty \sqrt{\frac{2}{\pi}} e^{-s^2/2} ds = \int_0^{\sqrt{t}/a} \frac{1}{s^2} \sqrt{\frac{2}{\pi}} e^{-\frac{1}{2s^2}} ds$$

$$\overset{x = (as)^2}{=} \int_0^t \frac{a}{\sqrt{2\pi x^3}} \exp\Big(-\frac{a^2}{2x}\Big) dx.$$ Hence, the PDF is $\frac{a}{\sqrt{2\pi t^3}} \exp\Big(-\frac{a^2}{2t}\Big)$.

Example: $L = \sup\{ t \leq 1 : B_t = 0 \}$ the last visit time of 0.

$$\mathbb{P}_0 (L \leq s) = \int_{-\infty}^\infty p_s(0, x)\, \mathbb{P}_x (T_0 > 1-s)\, ds = \int_{-\infty}^{+\infty} p_s(0,x)\, \underbrace{\mathbb{P}_0 (T_{-x} > 1-s)}_{= \mathbb{P}_0 (T_x > 1-s)}\, ds$$

$$= 2 \int_0^\infty \frac{1}{\sqrt{2\pi s}} \exp\Big(-\frac{x^2}{2s}\Big) \int_{1-s}^\infty \frac{x}{\sqrt{2\pi r^3}} \exp\Big(-\frac{x^2}{2r}\Big) dr\, dx$$

$$= \frac{1}{\pi} \int_{1-s}^\infty (s r^3)^{-\frac{1}{2}} \int_0^\infty x \exp\Big(-\frac{x^2}{2}\Big(\frac{1}{r} + \frac{1}{s}\Big)\Big) dx\, dr$$

$$= \frac{1}{\pi} \int_{1-s}^\infty (s r^3)^{-\frac{1}{2}} \frac{rs}{r+s} dr = \frac{1}{\pi} \int_{1-s}^\infty \Big(\frac{s}{r}\Big)^{\frac{1}{2}} \frac{1}{r+s} dr \qquad t = \frac{s}{r+s}$$

$$= \frac{1}{\pi} \int_0^s \frac{1}{\sqrt{t(1-t)}} dt = \frac{2}{\pi} \int_0^{\sqrt{s}} \frac{dt}{\sqrt{1-t^2}} = \frac{2}{\pi} \arcsin(\sqrt{s}). \quad \square$$

$\rightarrow$ <u>The arcsin law.</u>

Memo No.

Date / /

5. Martingales

X_t is a continuous time martingale if (i) $\mathbb{E}|X_t|<\infty$; (ii) X_t is adapted to $\mathcal{F}_t$; (iii) $\mathbb{E}(X_t \mid \mathcal{F}_s) = X_s \quad \forall\, s<t$.

Thm: Suppose X_t is a right-continuous martingale adapted a right-continuous filtration. If T is a bounded stopping time, then $\mathbb{E}X_T = \mathbb{E}X_0$.

Pf: Let n be an integer s.t. $T\le n-1$ a.s. We again discretize T as $T_m = ([2^m T]+1)/2^m$. ~~Then,~~ Note $Y_k^m = X(k2^{-m})$ is a martingale w.r.t. $\mathcal{F}_k^m = \mathcal{F}(k2^{-m})$ and $S_m = 2^m T_m$ is a stopping time for $(Y_k^m, \mathcal{F}_k^m)$. Hence,

$$X(T_m) = Y^m_{S_m} = \mathbb{E}(Y^m_{n2^m} \mid \mathcal{F}^m_{S_m}) = \mathbb{E}(X_n \mid \mathcal{F}(T_m)).$$

As $m\uparrow\infty$, $X(T_m)\to X(T)$ and $\mathcal{F}(T_m)\downarrow \mathcal{F}(T)$ by right-continuity of the path and filtration. Thus, we get $X(T) = \lim_{m\to\infty} \mathbb{E}(X_n \mid \mathcal{F}(T_m)) = \mathbb{E}(X_n \mid \mathcal{F}(T))$. Taking expectation gives

$\mathbb{E}X_T = \mathbb{E}X_n = \mathbb{E}X_0$ since X is a martingale.

Thm: B_t is a martingale w.r.t. the filtration $\mathcal{F}_t$.

Pf: By MP, $\mathbb{E}_x(B_t \mid \mathcal{F}_s) = \mathbb{E}_{B_s}(B_{t-s}) = B_s$ a.s., where we used $\mathbb{E}_y B_u = y \quad \forall u\ge 0$ □

Thm: If $a<x<b$, then $\mathbb{P}_x(T_a<T_b) = \frac{b-x}{b-a}$

Pf: Let $T = T_a\wedge T_b$. we have seen that $T<\infty$ a.s. Then $\forall t>0$, $T\wedge t$ is a bdd stopping time, so we have $x = \mathbb{E}_x B_{T\wedge t}$. Letting $t\to\infty$ and using the DCT (since $a\le B_{T\wedge t}\le b$), we get $x = \mathbb{E}_x B_T \Rightarrow x = a\,\mathbb{P}_x(T_a<T_b) + b\,\mathbb{P}_x(T_a>T_b) = a\,\mathbb{P}_x(T_a<T_b) + b(1-\mathbb{P}_x(T_a<T_b))$ □

Thm: $B_t^2 - t$ is a martingale.

Pf: $\mathbb{E}_x(B_t^2 \mid \mathcal{F}_s) = \mathbb{E}_x[B_s^2 + 2(B_t-B_s)B_s + (B_t-B_s)^2 \mid \mathcal{F}_s] = B_s^2 + B_s\,\mathbb{E}_x(B_t-B_s) + \mathbb{E}_x[(B_t-B_s)^2]$

$$= B_s^2 + t - s$$ □

Thm: If $a<x<b$, let $T = \inf\{t : B_t\notin (a,b)\}$. Then, $\mathbb{E}_x T = (b-x)(x-a)$.

Pf: $\forall t>0$, $T\wedge t$ is a bounded stopping time. So, $\mathbb{E}_x(B^2_{T\wedge t} - T\wedge t) = x^2$

Memo No.

Date / /

Letting $t\to\infty$, using the MCT, we get $\mathbb{E}_x(T\wedge t)\to\mathbb{E}T$; using the DCT, we get $\mathbb{E}B^2_{T\wedge t}\to\mathbb{E}B_T^2$. Thus, we get $x^2 = \mathbb{E}_x(B_T^2 - T) = a^2\,\mathbb{P}(T_a<T_b) + b^2\,\mathbb{P}(T_b<T_a) - \mathbb{E}_x T$

$\Rightarrow \mathbb{E}_x T = a^2\frac{b-x}{b-a} + b^2\frac{x-a}{b-a} - x^2 = -ab + (b+a)x - x^2 = (b-x)(x-a)$. □

Thm. $\exp(\theta B_t - \frac{\theta^2}{2}t)$ is a martingale.

Pf: $\mathbb{E}_x(\exp(\theta B_t - \frac{\theta^2}{2}t) \mid \mathcal{F}_s) = \exp(\theta B_s)\,\mathbb{E}_x[\exp(\theta(B_t-B_s)) \mid \mathcal{F}_s]$

$$= \exp(\theta B_s)\,\mathbb{E}_x\exp(\theta(B_t-B_s)) = \exp(\theta B_s)\exp(\tfrac{\theta^2}{2}(t-s)).$$ □

Thm: $\forall\lambda\ge 0$, we have $\mathbb{E}_0\exp(-\lambda T_a) = e^{-a\sqrt{2\lambda}}$.

Pf: $\forall t>0$, we have $1 = \mathbb{E}_0\exp(\theta B(T_a\wedge t) - \frac{\theta^2}{2}(T_a\wedge t))$. Taking $t\to+\infty$ and using DCT, we get $1 = \mathbb{E}_0\exp(\theta B(T_a) - \frac{\theta^2}{2}T_a)$. Taking $\theta = \sqrt{2\lambda}$, we get $\mathbb{E}_0 e^{-\lambda T_a} = e^{-a\sqrt{2\lambda}}$. □

Rmk: Note $\mathbb{E}_0\exp(-\lambda T_a)$ is the Laplace transform of T_a. Taking the inverse Laplace ~~tranforma,~~ transform, we can calaulate the distribution of T_a.

The following result gives a general way to construct martingales for 1D BM. It can also be derived from Itô's formula.

Thm: If $u(t,x)$ is a polynomial in t and x satisfying the heat equation:

$$\frac{\partial u}{\partial t} + \frac{1}{2}\frac{\partial^2 u}{\partial x^2} = 0,$$ then $u(t,B_t)$ is a martingale.

Pf: Recall the heat kernel $p_t(x,y) = \frac{1}{\sqrt{2\pi t}}\exp(-\frac{(y-x)^2}{2t})$. we can check directly that it satisfies the heat equation $\forall t\ge 0$: $\frac{\partial p_t}{\partial t} + \frac{1}{2}\frac{\partial^2 p_t}{\partial y^2} = 0$. Then,

$$\frac{\partial}{\partial t}\mathbb{E}_x u(t,B_t) = \int \frac{\partial}{\partial t}(p_t(x,y)\,u(t,y))\,dy = \frac{1}{2}\int \frac{\partial^2}{\partial y^2}p_t(x,y)\cdot u(t,y)\,dy$$
$$+ \int p_t(x,y)\frac{\partial}{\partial t}u_t(x,y)\,dy \quad \text{(Integrating by parts twice)}$$
$$= \int p_t(x,y)\left(\frac{1}{2}\frac{\partial^2}{\partial y^2}u(t,y) + \frac{\partial}{\partial t}u(t,y)\right)dy = 0.$$

(Since $u(t,y)$ is polynomial, there are no boundary terms in the integration by parts argument.)

So far, we have seen that $\mathbb{E}_x u(t,B_t)$ is constant in t. For the mg property,

Memo No.
Date / /

We need to prove that $\mathbb{E}_x(u(t,B_t)\mid \mathcal{F}_s) = u(s,B_s)$, $\forall\ 0<s<t$.

We use the MP to get $\mathbb{E}_x[u(t,B_t)\mid \mathcal{F}_s] = \mathbb{E}_x[v(t-s,B_{t-s})\circ\theta_s\mid \mathcal{F}_s]$

$= \underbrace{\mathbb{E}_{B(s)}v(t-s,B_{t-s})}$ $= v(0,B_s) = u(s,B_s)$, where $v(r,x) = u(r+s,x)$ also satisfies

$\to$ it is constant in t; so we take $t=s$. the heat equation w.r.t. r and x. $\square$

Examples of polynomials: x, x^2-t, x^3-3tx, $x^4-6x^2t+3t^2$, ...

The result can also be extended to $\exp(\theta x - \frac{\theta^2}{2}t)$, since it ~~decays~~ increases only exponentially.

Example: Let $T=\inf\{t: B_t\notin(-a,a)\}$. Then, $B_{T\wedge t}^4 - 6B_{T\wedge t}^2(T\wedge t) + 3(T\wedge t)^2$ is a martingale,

So we have $\mathbb{E}_x B_{T\wedge t}^4 - 6\mathbb{E}_x[(T\wedge t)B_{T\wedge t}^2] + 3\mathbb{E}_x(T\wedge t)^2 = 0$. We know $T<\infty$ a.s.

Taking $t\to+\infty$, we get $\underset{\overset{\shortparallel}{a^4}}{\mathbb{E}_x B_T^4} - 6\mathbb{E}_x(B_T^2 T) + 3\mathbb{E}_x T^2 = 0$

$a^4 - 6a^2\underset{\overset{\shortparallel}{a^2}}{\mathbb{E}T} + 3\mathbb{E}_x T^2 = 0 \Rightarrow \mathbb{E}_x T^2 = \frac{5}{3}a^4$. $\square$

6. Donsker's Theorem (RW converges to BM)

Thm (Skorokhod's representation thm) If $\mathbb{E}X=0$, $\mathbb{E}X^2<\infty$, then there is a stopping time T for BM so that $B_T \overset{d}{=} X$ and $\mathbb{E}T = \mathbb{E}X^2$.

Pf: Suppose X is supported on $\{a,b\}$ with $a<0<b$. Since $\mathbb{E}X=0$, we have

$$\mathbb{P}(X=a) = \frac{b}{b-a},\quad \mathbb{P}(X=b) = \frac{-a}{b-a}$$

We let $T=T_{a,b}=\inf\{t: B_t\notin(a,b)\}$. The previous thm implies $B_T\overset{d}{=}X$ and $\mathbb{E}_0 T = -ab = \mathbb{E}B_T^2$. For the general case, we write $F(x)=\mathbb{P}(X\le x)$ as a mixture of two point distributions with mean 0. Let $c = -\int_{-\infty}^0 u\,dF(u) = \int_0^\infty v\,dF(v)$.

Let φ be a bounded function with $\varphi(0)=0$. Then,

$$c\int\varphi(x)\,dF(x) = \left(\int_0^\infty \varphi(v)\,dF(v)\right)\left(\int_{-\infty}^0(-u)\,dF(u)\right)$$
$$+\left(\int_{-\infty}^0\varphi(u)\,dF(u)\right)\left(\int_0^\infty v\,dF(v)\right)$$
$$=\int_0^\infty dF(v)\int_{-\infty}^0 dF(u)\,\left(v\varphi(u)-u\varphi(v)\right).$$

Thus, we have $\int \varphi(x)\, dF(x) = C^{-1}\int_0^\infty dF(v) \int_{-\infty}^0 dF(u)\, (v-u) \underbrace{\left(\frac{v}{v-u}\varphi(u) + \frac{-u}{v-u}\varphi(v)\right)}_{\text{A two-point distribution}}$ (*)

We let $(U, V) \in \mathbb{R}^2$ such that $\mathbb{P}((U,V) = (0,0)) = F(\{0\})$,

$$(+) \quad \mathbb{P}((U,V)\in A) = C^{-1} \iint_{(u,v)\in A} dF(u)\, dF(v)\, (v-u), \quad A \subset (-\infty, 0) \times (0, \infty).$$

Then, define a family of two-point probability measures: $\mu_{0,0}(\{0\}) = 1$ and

$$\mu_{u,v}(\{u\}) = \frac{v}{v-u}, \quad \mu_{u,v}(\{v\}) = \frac{-u}{v-u} \quad \text{for } u < 0 < v.$$

Now, we write (*) as $\int \varphi(x)\, dF(x) = \mathbb{E} \int \varphi(x)\, d\mu_{U,V}(x)$.

We can show that the above holds without requiring $\varphi(0) = 0$.

(Taking $\varphi \equiv 1$ shows that (+) defines a prob. measure.)

From the above calculations, it follows that if we have (U, V) with distributions given in (+) and an independent BM defined on the same space, then $B(T_{U,V}) \overset{d}{=} X$.

To make $T_{U,V}$ a stopping time, we let $\mathcal{F}_0 = \sigma(X_0, U, V)$.

Next, we check that: $\mathbb{E}(T_{U,V}) = \mathbb{E}\{\mathbb{E}[T_{U,V} \mid U, V]\} = -\mathbb{E}(UV)$.

On the other hand, with (+) we get that

$$\mathbb{E}(-UV) = C^{-1} \int_0^\infty dF(v) \int_{-\infty}^0 dF(u)\, (-uv)(v-u)$$

$$= \int_{-\infty}^0 dF(u)\, (-u) \left\{-u + C^{-1}\int_0^\infty dF(v)\, v^2\right\}$$

$$= \int_{-\infty}^0 u^2\, dF(u) + \int_0^\infty v^2\, dF(v) = \mathbb{E}X^2. \qquad \square$$

A simple extension of the above result gives:

<u>Thm</u>: Let $X_1, X_2, \cdots$ be i.i.d. with a distribution F, which has mean 0 and variance 1. Let $S_n = X_1 + X_2 + \cdots + X_n$. There is a sequence of stopping times $T_0 = 0, T_1, T_2, \cdots$ such that $S_n \overset{d}{=} B(T_n)$ and $T_n - T_{n-1}$ are i.i.d.

<u>Pf</u>: Let $(U_1, V_1), (U_2, V_2), \cdots, (U_n, V_n)$ be i.i.d. and have distribution given in (+). Let B_t be an independent BM. Let $T_0 = 0$, and, for $n \geq 1$, let

$$T_n = \inf\{t \geq T_{n-1} : B_t - B(T_{n-1}) \notin (U_n, V_n)\}. \qquad \square$$

With this result, we immediately get another way of showing the CLT of S_n.

<u>Thm (CLT) Under the setting of the above thm,</u> $S_n/\sqrt{n} \to N(0,1)$ in law.

Pf: By Brownian scaling, then $S_n/\sqrt{n} \overset{d}{=} B(T_n)/\sqrt{n} \overset{d}{=} B(T_n/n)$.

By WLLN, $T_n/n \to 1$ in probability, which would imply that $B(T_n/n) \to B_1$ in law.

To fill in the details, $\forall \varepsilon > 0$, pick $\delta > 0$ so that $\mathbb{P}(|B_t - B_1| > \varepsilon$ for some $t \in (1-\delta, 1+\delta)) < \varepsilon/2$.

Then, pick N large enough so that $\mathbb{P}(|T_n/n - 1| > \delta) < \varepsilon/2 \quad \forall n \geq N$. The last two estimates give $\forall n \geq N$, $\mathbb{P}(|B(T_n/n) - B(1)| > \varepsilon) < \varepsilon$. Since ε is arbitrary, it follows that $B(T_n/n) \to B(1)$ in prob. □

The main goal is to prove the ~~te~~ celebrated Donsker's thm, showing that RW paths converge to BM paths. Given a RW S_n, let

$$S(u) = \begin{cases} S_k, & \text{if } u = k \in \mathbb{N} \\ \text{linear on } [k, k+1] & \text{for } k \in \mathbb{N} \end{cases}$$

<u>Thm</u>: (Dosker's thm) Under the setting of the above thm, we get $S(n\cdot)/\sqrt{n} \to B(\cdot)$ in law, i.e., the associated measurs on $C([0,1])$ converge weakly.

Pf: For simplicity of notations, let $X_{n,m}$, $1 \leq m \leq n$, be a triangular array of RVs, $S_{n,m} = X_{n,1} + \cdots + X_{n,m}$, and suppose $S_{n,m} = B(\tau^n_m)$. Then, let

$$S_{n,(u)} = \begin{cases} S_{n,m} & \text{if } u = m \in \{0, 1, \cdots, n\}, \\ \text{linear for } u \in [m-1, m] & \text{where } m \in \{1, \cdots, n\}. \end{cases}$$

<u>Lemma</u>: If $\tau^n_{[ns]} \to s$ in probability for each $s \in [0,1]$, then $\|S_{n,(n\cdot)} - B(\cdot)\| \to 0$ in prob.

Pf of Thm

In the original setting, $X_{n,m} = X_m/\sqrt{n}$ and define $\tau^n_1, \cdots, \tau^n_n$ so that $(S_{n,1}, \cdots, S_{n,n}) \overset{d}{=} (B(\tau^n_1), \cdots, B(\tau^n_n))$. By Brownian scaling, we have $\tau^n_m \overset{d}{=} T_m/n$.

By LLN, the hypothesis of the lemma is satisfied.

To conclude the proof, we need to show that: $\forall$ bdd continuous function φ on $C([0,1])$, $\mathbb{E}\varphi(S_{n,(n\cdot)}) \to \mathbb{E}\varphi(B(\cdot))$. In fact,

$\forall \varepsilon > 0$, let $G_\delta := \{\omega \in C([0,1]) :$ if $\|\omega - \omega'\| < \delta$, then $|\varphi(\omega) - \varphi(\omega')| < \varepsilon\}$.

Since φ is continuous, $G_\delta \uparrow C([0,1])$ as $\delta \downarrow 0$. Let $\Delta = \|S_{n,(n\cdot)} - B(\cdot)\|$. Then,

$$|\mathbb{E}\varphi(S_{n,(n\cdot)}) - \mathbb{E}\varphi(B(\cdot))| \leq \varepsilon + (2\sup|\varphi(\omega)|)\{\mathbb{P}(G^c_\delta) + \underbrace{\mathbb{P}(\Delta \geq \delta)}_{\downarrow \text{ as } n\to\infty \ 0}\}$$

for large enough n.

□

中国科学技术大学
University of Science and Technology of China
Memo No.
Date / /

Pf: By unif. continuity of B, if $\varepsilon > 0$ then there is a $\delta > 0$ so that $1/\delta \in \mathbb{N}$ and

(*) $\mathbb{P}(|B_t - B_s| < \varepsilon \quad \forall\, 0 \leq s \leq 1,\ |t-s| < 2\delta) > 1 - \varepsilon$.

By the hypothesis of the lemma, there exists large enough N_δ so that if $n \geq N_\delta$ then

$$\mathbb{P}(\underbrace{|\tau^n_{[nk\delta]} - k\delta| < \delta \text{ for } k = 1, 2, \cdots, 1/\delta}_{E}) \geq 1 - \varepsilon$$

Since $m \to \tau^n_m$ is increasing, we see that if $s \in ((k-1)\delta, k\delta)$,

$$\tau^n_{[ns]} - s \geq \tau^n_{[n(k-1)\delta]} - k\delta > -2\delta \quad \text{on the good event } E,$$

$$\tau^n_{[ns]} - s \leq \tau^n_{[nk\delta]} - (k-1)\delta < 2\delta \quad \text{on } E.$$

Hence, $\forall n \geq N_\delta$, $\mathbb{P}(\sup_{0 \leq s \leq 1} |\tau^n_{[ns]} - s| < 2\delta) \geq 1 - \varepsilon$. (**)

Combining (*) with (**), we get ~~$\mathbb{P}(\sup_{0\leq s\leq 1}|\tau^n_{[ns]} - s| < 2\delta) \geq$~~

(***) $\quad \mathbb{P}(|S_{n,m} - B_{m/n}| < \varepsilon \quad \forall m \leq n) \geq 1 - 2\varepsilon$ for large enough n.

To deal with $t = (m+\theta)/n$ with $0 < \theta < 1$, we observe that

$$|S_{n,(nt)} - B_t| \leq (1-\theta)|S_{n,m} - B_{m/n}| + \theta|S_{n,m+1} - B_{(m+1)/n}|$$
$$+ (1-\theta)|B_{m/n} - B_t| + \theta|B_{(m+1)/n} - B_t|$$

We use (***) to bound the first two terms and (*) to bound the last two. Then, we get if $n \geq N_\delta$ and $1/n < 2\delta$, then $\|S_{n,(n\cdot)} - B(\cdot)\| < 2\varepsilon$ with prob. $\geq 1 - 2\varepsilon$. □

With Donsker's thm, the following simple fact allows us to calculate some probabilities regarding RW using techniques of BMs:

If $\psi : C([0,1]) \to \mathbb{R}$ is continuous $\mathbb{P}_0$-a.s., then $\psi(S(n\cdot)/\sqrt{n}) \to \psi(B(\cdot))$ in law.

<u>Example</u>: Find the law of the maxima: $\max_{0 \leq m \leq n} S_n/\sqrt{n}$.

Let $\psi(\omega) = \max\{\omega(t) : 0 \leq t \leq 1\}$. $\psi : C([0,1]) \to \mathbb{R}$ is continuous. Then, by Donsker's thm, $\max_{0 \leq m \leq n} S_m/\sqrt{n} \Rightarrow M_1 = \max_{0 \leq t \leq 1} B_t$. Recall that

$$\mathbb{P}_0(M_1 \geq a) = \mathbb{P}_0(T_a \leq 1) = 2\mathbb{P}_0(B_1 > a).$$ □