Exercise 1. Consider a measurable space $(\Omega, \mathcal{F})$ with finite elementary event set $\Omega = \{1, \cdots, n\}$,[1] the corresponding $\mathcal{F} = 2^{\Omega}$, and the "Lebesgue measure" $\nu_i = 1$, $1 \leq i \leq n$. In discrete time, a deterministic first-order "dynamics" in the Ω has a one step map $S : \Omega \to \Omega$. A stochastic Markov (chain) dynamics, X_k, has one step transitions in terms of a set of conditional probabilities $p^{(\boldsymbol{\nu})}(i,j) = \Pr\{X_{k+1} = j | X_k = i\}$.

(a) Since a deterministic first-order dynamics is just a special, singular case of a Markov dynamics, express the transition probability $p(i,j)$ in terms of the map S.
(b) If a Markov chain with $p(i,j)$ has a unique invariant probability $\boldsymbol{\pi} = \{\pi_1, \cdots, \pi_n\} \neq 0$, express the transition probability "density" $p^{(\boldsymbol{\pi})}(i,j)$ under the measure $\boldsymbol{\pi}$ in terms of the transition probabilty $p^{(\boldsymbol{\nu})}(i,j)$.
(c) Show that

$$\boldsymbol{\pi} P^{(\boldsymbol{\pi})} = \mathbf{1},$$

and

$$P^{(\boldsymbol{\pi})} \boldsymbol{\pi}^T = \mathbf{1}^T,$$

where $P^{(\boldsymbol{\pi})}$ is the transition probability density matrix w.r.t. $\boldsymbol{\pi}$ and $\mathbf{1} = (1, \cdots, 1)$. Please explain these two equations.
(d) The *reversibility* of a Markov chain is introduced in §4.5 of MLN. What is the $P^{(\boldsymbol{\pi})}$ of a reversible Markov chain?
(e) Now return to a deterministic map $S : \Omega \to \Omega$. Show that its has a stationary probability $\boldsymbol{\pi} = (\frac{1}{n}, \cdots, \frac{1}{n})$ if and only if the map S is one to one. Within the context of a deterministic S, discuss the notion of *irreducibility* defined in §4.3 of MLN.

Solution 1. (a) Thinking of the S in terms of a Markov chain, we can write the transition matrix as

$$p(i,j) = \mathbb{P}(X_{k+1} = j \mid X_k = i) = \begin{cases} 1, & \text{if } S(i) = j \\ 0, & \text{otherwise.} \end{cases}$$

(b) Suppose that $\boldsymbol{\pi}$ is the stationary distribution to $p^{\boldsymbol{\nu}}(i,j)$, we then gave that

$$\boldsymbol{\pi} = \boldsymbol{\pi} P^{(\boldsymbol{\nu})}$$

Representing $p^{\boldsymbol{\pi}}$ as the Radon-Nikodym derivative of $p^{(\boldsymbol{\nu})(i,j)}$ with respect to $\boldsymbol{\pi}$, we have that

$$p^{(\boldsymbol{\pi})}(i,j) = p^{(\boldsymbol{\nu})}(i,j)/\pi_j.$$

(c) With this definition, the first equation is clear as

$$\sum_{j \in \Omega} \pi_j p^{(\boldsymbol{\pi})}(i,j) = \sum_{j \in \Omega} \pi_j p^{(\boldsymbol{\nu})}(i,j)/\pi_j = \sum_{j \in \Omega} p^{(\boldsymbol{\nu})}(i,j) = 1,$$

[1]If we let $\Omega = \{A, B, \cdots\}$ be a set of finite symbols, then nonlinear dynamic systems, chaos and Smale's horseshoe, can be studied as *subshifts of finite type* in $\Omega^{\mathbb{Z}}$, the space of all bi-infinite sequences of elements of Ω, known as symbolic dynamics.

for any $i \in \Omega$. The second equation follows as

$$\sum_{j\in\Omega} p^{(\boldsymbol{\pi})}(j,i)\pi_j = \frac{1}{\pi_i}\sum_{j\in\Omega} p^{(\boldsymbol{\nu})}(j,i)\pi_j = \pi_i/\pi_i = 1.$$

(d) When the chain is reversible, we have that

$$\pi_i p^{(\boldsymbol{\nu})}(i,j) = \pi_j p^{(\boldsymbol{\nu})}(j,i) \quad \text{MLN (Defn. 4.5.3.)} \ .$$

In this case, we have that

$$p^{(\boldsymbol{\pi})}(i,j) = p^{(\boldsymbol{\nu})}(i,j)/\pi_j = p^{(\boldsymbol{\nu})}(j,i)/\pi_i.$$

(e) In the case that we're working with a deterministic map S, the transition matrix is given by

$$p(i,j) = \mathbb{P}(X_{k+1} = j \mid X_k = i) = \begin{cases} 1, & \text{if } S(i) = j \\ 0, & \text{otherwise.} \end{cases}$$

Suppose that this Markov chain has stationary distribution $\boldsymbol{\pi} = (1/n, \dots, 1/n)$. Then it follows that

$$\frac{1}{n}\sum_{j\in\Omega} p(i,j) = \frac{1}{n}, \text{ for all } i.$$

This can only hold if each column has percisely one 1 in it due to our definition of $p(i,j)$. This means that the matrix P is invertible as it is similar to the identity matrix and the underlying map S must be a bijection.

Supposing that S is a bijection, we have that for every i there is exactly one j such that $S(i) = j$ and for each i there is exactly one k so that $S(k) = i$. Therefore, in each column and row there is exactly one 1. Therefore, we have that

$$\frac{1}{n}\sum_{j\in\Omega} p(i,j) = \frac{1}{n},$$

so $\boldsymbol{\pi}$ is a stationary distribution. Irreducibility in the sense of the deterministic map is the same as saying there are no cycles $S^k(i)$ with $k < n$ for any $i \in \mathbb{N}$. That is, there are no smaller subcyles of period $k < n$.

Exercise 2. Let $W(t)$ be a standard Brownian motion. Introducing a function of the Brown motion

$$\tilde{W}(s) = (1-s)W\left(\frac{s}{1-s}\right), \quad 0 < s < 1.$$

Compute its expected value, variance, and covariance function

$$\mathrm{Cov}\left[\tilde{W}(s_1), \tilde{W}(s_2)\right], \quad 0 < s_1 < s_2 < 1.$$

$\tilde{W}(s)$ is known as a *Brownian bridge*.

Solution 2. Begining with the expectation, we have that

$$\begin{aligned}\mathbb{E}[\tilde{W}(s)] &= (1-s)\mathbb{E}\left[W\left(\frac{s}{1-s}\right)\right] \\ &= (1-s)\cdot 0 = 0,\end{aligned}$$

since W is a standard Brownian motion. Next up is variance

$$\begin{aligned}\mathrm{Var}[\tilde{W}(s)] &= (1-s)^2\mathrm{Var}\left[W\left(\frac{s}{1-s}\right)\right] \\ &= (1-s)^2\left(\frac{s}{1-s}\right) \\ &= s(1-s),\end{aligned}$$

where we've used that W is a standard Brownian motion. Now to compute the covariance. Assuming $0 < s_1 < s_2 < 1$, we have that

$$\begin{aligned}\mathrm{Cov}\left[\tilde{W}(s_1), \tilde{W}(s_2)\right] &= (1-s_1)(1-s_2)\mathrm{Cov}\left[W\left(\frac{s_1}{1-s_1}\right), W\left(\frac{s_2}{1-s_2}\right)\right] \\ &= (1-s_1)(1-s_2)\frac{s_1}{1-s_1} \\ &= s_1(1-s_2),\end{aligned}$$

where we've used that W is a standard Brownian motion and that $s_1/(1-s_1) < s_2/(1-s_2)$.

Exercise 3. $W(t)$ is a standard Brownian motion.
(a) Let $c > 0$ a constant. Show that the procss defined by $B(t) = cW(t/c^2)$ is a standard Brownian motion.
(b) For $t = n = 0, 1, \cdots$, show that $W^2(n) - n$ is a discrete time martingale.

Solution 3. (a) First, we have that $B(0) = cW(0) = 0$. Taking $0 \leq r < s < t < u < \infty$, we have that

$$(B(u) - B(t)) \text{ independent } (B(s) - B(r))$$

as this is just a constant scaling of a standard Brownian motion which satisfies this independent increments property. Next for $0 \leq r < s$,

$$\begin{aligned} B(s) - B(r) &= c(W(s/c^2) - W(r/c^2)) \\ &\sim \text{Normal}(0, s - r), \end{aligned}$$

where we've used that for a standard Brownian motion W, $W(s/c^2) - W(r/c^2)$ has distribution $\text{Normal}(0, (s-r)/c^2)$. The map $t \mapsto B(t)$ is continuous for all ω since it is just composotion with continuous functions $t \mapsto t/c^2$ and $x \mapsto cx$ with the individual $B(t)(\omega)$ which are continuous in t themselves.
(b) Let $M_n = W^2(n) - n$ and $\mathcal{F}_n$ be the filtration up to time n. We then have that

$$\begin{aligned} \mathbb{E}[M_{n+1} - M_n \mid \mathcal{F}_n] &= \mathbb{E}[W^2(n+1) - W^2(n-1) \mid \mathcal{F}_n] - 1 \\ &= \mathbb{E}[(W(n) + W(n+1) - W(n))^2 - W^2(n) \mid \mathcal{F}_n] - 1 \\ &= \mathbb{E}[2W(n)[W(n+1) - W(n)] + (W(n+1) - W(n))^2 \mid \mathcal{F}_n] - 1. \end{aligned}$$

We'll now simplify this using the choice of filtration, so that

$$\begin{aligned} \mathbb{E}[M_{n+1} - M_n \mid \mathcal{F}_n] &= 2W(n)\mathbb{E}[W(n+1) - W(n)] + \mathbb{E}[(W(n+1) - W(n))^2] - 1 \\ &= 0 + 1 - 1 \\ &= 0, \end{aligned}$$

where we've used that $W(n+1) - W(n) \sim \text{Normal}(0, 1)$.

Exercise 4. $W(t)$ is a standard Brownian motion. What is the characteristic function of $W(N_t)$ where N_t is a Poisson process with intensity λ, and the Brownian motion $W(t)$ is independent of the Poisson process N_t.

Solution 4. Write as sum of normals up to N_t? We'll probably use theorem 3.1.9. Writing that $Z_t = W(N_t)$, we have that the characteristic function of Z_t can be written as

$$\varphi_{Z_t}(u) = \mathbb{E}[e^{iuW(N_t)}] = \mathbb{E}\left[\sum_{n=0}^{\infty} e^{iuW(m)} 1_{N_t=n}\right].$$

By the independence of N_t and $W(t)$, we have that

$$\begin{aligned}
\varphi_{Z_t}(u) &= \sum_{n=0}^{\infty} \mathbb{E}\left[e^{iuW(n)}\right] \mathbb{P}(N_t = n) \\
&= \sum_{n=0}^{\infty} e^{-nu^2/2} \cdot \mathbb{P}(N_t = n) \\
&= \sum_{n=0}^{\infty} e^{-nu^2/2} \cdot \frac{(\lambda t)^n}{n!} e^{-\lambda t}.
\end{aligned}$$

where above we've used that $W(n) \sim \text{Normal}(0, n)$ and its corresponding characteristic function. Additionally, we've used that $N_t \sim \text{Pois}(\lambda t)$. We can try to simplify this as

$$\begin{aligned}
\varphi_{Z_t}(u) &= \sum_{n=0}^{\infty} e^{-nu^2/2} \cdot \frac{(\lambda t)^n}{n!} e^{-\lambda t} \\
&= e^{-\lambda t} \sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} e^{-nu^2/2}.
\end{aligned}$$

I don't know if the sum can be evaluated to a closed form or to the characteristic function of a familiar distribution, so I did not proceed further.

Exercise 5. The n^{th} variation of a function f, over the interval $[0, T]$ is defined as

$$V_T(n, f) := \lim_{\|\Pi\| \to 0} \sum_{j=0}^{m-1} \left| f(t_{j+1}) - f(t_j) \right|^n,$$

in which $\Pi = \{0 = t_0, t_1, \cdots, t_n = T\}$ is a *partition* of the $[0, T]$, and

$$\|\Pi\| = \max_{0 \le j \le n-1} \big(t_{j+1} - t_j\big).$$

Show that $V_T(1, W) = \infty$ and $V_T(3, W) = 0$, where W is a realization of the Brownian motion.

Solution 5. Suppose that the first variation of B over $[0, T]$ is a finite number C i.e.

$$V_T(1, W) = C.$$

We can then write that

$$\sum_{j-0}^{m-1} [W(t_{j+1}) - W(t_j)]^2 \le \max_{0 \le j \le m-1} |W(t_{j+1}) - W(t_j)| \sum_{j=0}^{m-1} |W(t_{j+1}) - W(t_j)| \,.$$

Notice that the last term is similar to the first variation for a fixed partition Π. Since W is contious on a compact interval, we know that

$$\max_{0 \le j \le m-1} |W(t_{j+1}) - W(t_j)| \xrightarrow{||\Pi|| \to 0} 0.$$

Since the first variation is finite in the limit as $||\Pi|| \to 0$, we then have that

$$V_T(2, W) = \lim_{||\Pi|| \to 0} \sum_{j-0}^{m-1} [W(t_{j+1}) - W(t_j)]^2 = 0.$$

This is in contradiction with thereom 7.3.3. (MLN) which states that $V_T(2, W) = T$ almost surely, so we have that $V_T(1, W)$ must be infinite (almost surely).

Exercise 6. (a) Show the transition probability density function for standard Brownian motion $W(t)$:

$$\frac{1}{dx}\Pr\left\{x < W(t+s) \leq x + dx \middle| W(s) = y\right\} = \frac{1}{\sqrt{2\pi t}} e^{-\frac{(x-y)^2}{2t}} = p(x;t|y),$$

in which $t, s > 0$.

(b) Show that $p(x;t|y)$ satisfies the following two linear partial differential equations:

$$\frac{\partial p(x;t|y)}{\partial t} = \frac{1}{2}\left(\frac{\partial^2 p(x;t|y)}{\partial x^2}\right) \quad \text{and} \quad \frac{\partial p(x;t|y)}{\partial t} = \frac{1}{2}\left(\frac{\partial^2 p(x;t|y)}{\partial y^2}\right).$$

Solution 6. (a) We have that

$$W(t+s) - W(s) \sim \text{Normal}(0,t)$$

as W is a standard Brownian motion. We can then write for some $\Delta x \neq 0$,

$$\begin{aligned}\mathbb{P}(x < W(t+s) \leq x + \Delta x \mid W(s) = y) &= \mathbb{P}(x - y < W(t+s) - W(s) < x + \Delta x - y) \\ &= \mathbb{P}(x - y < Z < x - y + \Delta x),\end{aligned}$$

where Z is Normal$(0,t)$. Diving this by Δx and taking the limit as $\Delta x \to 0$, we have that

$$\frac{1}{dx}\mathbb{P}\left\{x < W(t+s) \leq x + dx \middle| W(s) = y\right\} = f_Z(x-y),$$

where f_Z is the density of Z i.e.

$$\frac{1}{dx}\mathbb{P}\left\{x < W(t+s) \leq x + dx \middle| W(s) = y\right\} = \frac{1}{\sqrt{2\pi t}}\exp\left(-\frac{(x-y)^2}{2t}\right) = p(x, t \mid y).$$

(b) We start by computing the partial derivative with respect to t.

$$\begin{aligned}\frac{\partial}{\partial t}\left(p(x,t \mid y)\right) &= \frac{\partial}{\partial t}\left(\frac{1}{\sqrt{2\pi t}}\right) \cdot e^{-\frac{(x-y)^2}{2t}} + \frac{1}{\sqrt{2\pi t}}\frac{\partial}{\partial t}\left(e^{-\frac{(x-y)^2}{2t}}\right) \\ &= -\frac{1}{2t\sqrt{2\pi t}}e^{-\frac{(x-y)^2}{2t}} + \frac{(x-y)^2}{2t^2\sqrt{2\pi t}}e^{-\frac{(x-y)^2}{2t}}.\end{aligned}$$

Next computing the first partial derivative with respect to x, we have

$$\frac{\partial}{\partial x}\left(p(x,t \mid y)\right) = -\frac{(x-y)}{t\sqrt{2\pi t}}e^{-\frac{(x-y)^2}{2t}}.$$

Differentiating this again with respect to x and using the product rule, we see

$$\begin{aligned}\frac{\partial^2}{\partial x^2}\left(p(x,t \mid y)\right) &= -\frac{1}{t\sqrt{2\pi t}}\left(-\frac{(x-y)^2}{t}e^{-\frac{(x-y)^2}{2t}} + e^{-\frac{(x-y)^2}{2t}}\right) \\ &= -\frac{1}{t\sqrt{2\pi t}}e^{-\frac{(x-y)^2}{2t}} + \frac{(x-y)^2}{t^2\sqrt{2\pi t}}e^{-\frac{(x-y)^2}{2t}}.\end{aligned}$$

From our computations, it is clear that

$$\frac{\partial}{\partial t}\left(p(x,t \mid y)\right) = \frac{1}{2}\frac{\partial^2}{\partial x^2}\left(p(x,t \mid y)\right).$$

The second desired equation follows from the fact that x and y are interchangable in the formula for $p(x,t \mid y)$. Therefore,

$$\begin{aligned}
\frac{\partial^2}{\partial x^2}\left(p(x,t \mid y)\right) &= \frac{\partial^2}{\partial y^2}\left(p(x,t \mid y)\right)\\
\frac{\partial}{\partial t}\left(p(x,t \mid y)\right) &= \frac{1}{2}\frac{\partial^2}{\partial y^2}\left(p(x,t \mid y)\right).
\end{aligned}$$