CONTENTS FOR APPENDICES

A SAMPLE PATH PROPERTIES: SEPARABILITY AND CONTINUITY

A.1 Separability

Not all processes are separable. Consider the process $X(\omega, t), t \in [0, 1]$ on $(\Omega, \mathcal{A}, P)$ where $\Omega = [0, 1]$, $\mathcal{A} = \mathcal{B}([0, 1])$ is the Borel σ-algebra and P Lebesgue measure. Define

$$X(\omega, t) = \begin{cases} 1 & \omega = t \\ 0 & \text{otherwise} \end{cases} . \tag{A.1}$$

Such a process is not separable. Assume that it is. Let $S \subset [0, 1]$ be a separating set. Then

$$\begin{aligned} \Pr\{\omega : X(\omega, t) = 0, \forall t \in [0, 1]\} &= \Pr\{\omega : X(\omega, t) = 0, \forall t \in S\} \\ &= 1 - \Pr\{\omega : X(\omega, t) = 1, \text{for some } t \in S\} \\ &= 1 - \Pr\{\omega : \omega \in S\} = 1 \end{aligned}$$

since S is countable with $|S| = 0$. But

$$\begin{aligned} \Pr\{\omega : X(\omega, t) = 0, \forall t \in [0, 1]\} &= 1 - \Pr\{\omega : X(\omega, t) = 1, \text{for some } t \in [0, 1]\} \\ &= 1 - \Pr\{\omega : \omega \in [0, 1]\} = 0 . \end{aligned}$$

A.2 Continuity

Now examine continuity of sample paths.

Theorem A.1 (Wong and Hajek, Proposition 4.2) *Let $\{X(t), t \in T\}$ be a separable process and let T be a finite interval. If there exists strictly positive constants α, β, C such that*

$$E|X(t + h) - X(t)|^{\alpha} \leq Ch^{1+\beta}$$

then

$$\sup_{t \in T} |X(t + h) - X(t)| \xrightarrow{a.s} 0$$

Proof Using the Chebyshev inequality and the hypothesis, we have

$$P(|X(t + h) - X(t)| \geq h^{\gamma}) \leq \frac{E\{|X(t + h) - X(t)|^{\alpha}\}}{h^{\gamma\alpha}} \leq \frac{Ch^{1+\beta}}{h^{\gamma\alpha}} = Ch^{1+\beta-\gamma\alpha} = Ch^{1+\delta} \tag{A.2}$$

where we choose $\gamma < \beta/\alpha$ and $\delta = \beta - \gamma\alpha > 0$. Therefore,

$$P\left(\max_{0\le k\le 2^n}\left|X\left(\frac{k+1}{2^n}\right) - X\left(\frac{k}{2^n}\right)\right| \ge \left(\frac{1}{2^n}\right)^{\gamma}\right)$$
$$\le \sum_{k=0}^{2^n} P\left(\left|X\left(\frac{k+1}{2^n}\right) - X\left(\frac{k}{2^n}\right)\right| \ge \left(\frac{1}{2^n}\right)^{\gamma}\right)$$
$$\le \sum_{k=0}^{2^n} C\left(\frac{1}{2^n}\right)^{1+\delta} = C2^{-n\delta}$$

Let $Z_\nu = \sup_{0\le k\le 2^\nu} |X(\omega, k+1/2^\nu) - X(\omega, k/2^\nu)|$. Since $\sum_{\nu=0}^{\infty} 2^{-\delta\nu} < \infty$,

$$\sum_{\nu=0}^{\infty} P(Z_\nu \ge (1/2^\nu)^\gamma) < \infty,$$

and by the Borel–Cantelli lemma $P(\{Z_\nu \ge 1/2^{\nu\delta},\ \ i.o.\}) = 0$. So, $\exists N(\omega)$ almost-surely finite such that

$$Z_\nu(\omega) < \frac{1}{2^{\nu\delta}} \quad \text{for all } \nu \ge N(\omega)$$

and $\lim_{n\to\infty} \sum_{\nu=n+1}^{\infty} Z_\nu(\omega) = 0$ with probability one. Since $|X(t) - X(s)| \le |X(t) - X(k/2^n)| + |X(s) - X(k/2^n)|$ by the triangle inequality and $|X(t) - X(k/2^n)| \le \sum_{\nu=n+1}^{\infty} Z_\nu$,

$$\sup_{t,s\in S, |t-s|<2^{-n}} |X(t) - X(s)| \le 2\sum_{\nu=n+1}^{\infty} Z_\nu \overset{a.s.}{\to} 0.$$

Since S is a separating set,

$$\sup_{t,s\in T, |t-s|<h} |X(t) - X(s)| \overset{a.s.}{\to} 0 \quad \text{as } h \to 0.$$

□

Now we return to our remark 9.32 and establish the property of Brownian motion that for a Brownian motion $(dX(t))^2 \sim O(dt)$.

Theorem A.2 (Wong and Hajek, Proposition 3.4, p. 53) *Let $T = [a, b]$ and $T_n = [a = t_0^{(n)} < t_1^{(n)} < \cdots < t_{N(n)}^{(n)} = b]$, $n = 1, 2, \ldots$, a sequence of partitions such that*

$$\Delta_n = \max_{1\le\nu\le N(n)} (t_\nu^{(n)} - t_{\nu-1}^{(n)}) \to 0 \quad \textit{as } n \to \infty.$$

If $\{X(t), t \ge 0\}$ is a Brownian motion, then

$$\sum_{\nu=1}^{N(n)} (X(t_\nu^{(n)}) - X(t_{\nu-1}^{(n)}))^2 \xrightarrow{q.m.} b - a \quad \textit{as } n \to \infty.$$

If $\sum_{n=1}^{\infty} \Delta_n < \infty$, then the convergence is also almost sure.

Proof First prove the quadratic mean convergence. Let

$$S_n = \sum_{\nu=1}^{N(n)} (X(t_\nu^{(n)}) - X(t_{\nu-1}^{(n)}))^2 - (b - a).$$

Then S_n is a sum of independent random variables with mean given by

$$\begin{aligned}
E\{S_n\} &= \sum_{\nu=1}^{N(n)} E\{(X(t_\nu^{(n)}) - X(t_{\nu-1}^{(n)}))^2\} - (b-a) \\
&= \sum_{\nu=1}^{N(n)} (t_\nu^{(n)} - t_{\nu-1}^{(n)}) - (b-a) \\
&= 0.
\end{aligned}$$

We need to show that the variance of S_n converges to zero as $n \to \infty$.

$$\begin{aligned}
E\{S_n^2\} &= E\left\{\left(\sum_{\nu=1}^{N(n)} [(X(t_\nu^{(n)}) - X(t_{\nu-1}^{(n)}))^2 - (t_\nu^{(n)} - t_{\nu-1}^{(n)})]\right)^2\right\} \\
&= \sum_{\nu=1}^{N(n)} E\{[(X(t_\nu^{(n)}) - X(t_{\nu-1}^{(n)}))^2 - (t_\nu^{(n)} - t_{\nu-1}^{(n)})]^2\},
\end{aligned}$$

since we only have to worry about the diagonal terms. Each term is the fourth moment of a Gaussian random variable. Hence,

$$\begin{aligned}
E\{S_n^2\} &= \sum_{\nu=1}^{N(n)} 2\,(t_\nu^{(n)} - t_{\nu-1}^{(n)})^2 \\
&\le 2\,\Delta_n \sum_{\nu=1}^{N(n)} (t_\nu^{(n)} - t_{\nu-1}^{(n)}) \\
&= 2\,\Delta_n (b-a)
\end{aligned}$$

which converges to zero as $n \to \infty$. So, we have shown the quadratic mean convergence. It remains to show the almost sure convergence.

Using the Chebyshev inequality implies

$$P(|S_n| > \epsilon) \le \frac{E\{S_n^2\}}{\epsilon^2} \le 2\,(b-a)\frac{\Delta_n}{\epsilon^2}.$$

Note that this implies convergence in probability. Also, since by assumption $\sum \Delta_n \langle \infty$, it follows that for every $\epsilon \rangle 0$,

$$\sum_{n=1}^{\infty} P(|S_n| > \epsilon) \le \frac{2\,(b-a)}{\epsilon^2} \sum_{n=1}^{\infty} \Delta_n < \infty.$$

The Borel–Cantelli theorem is satisfied and

$$P(\{|S_n| > \epsilon \ i.o.\}) = 0.$$

Hence, the set of sample points ω for which $S_n(\omega)$ does not converge to 0, $\{|S_n| > \epsilon \ i.o.\}$, has probability zero. □

B SECOND ORDER PROCESSES

B.1 Integration Theorems

We state here the fundamental convergence theorems from integration theory. For proofs see [477], p. 47–53. We assume a measure space $(\Omega, \mathcal{A}, \mu)$ throughout.

Definition B.1 *A measurable function f is said to be* **integrable** *if $\int |f|\, d\mu < \infty$.*

Theorem B.2

1. **The Monotone Convergence Theorem** *Let f_n be a sequence of nonnegative measurable functions such that $f_n \le f_{n+1}$ for all n, and $\lim_{n\to\infty} f_n = f$. Then*
$$\int \lim_{n\to\infty} f_n\, d\mu = \lim_{n\to\infty} \int f_n\, d\mu.$$
2. **Fatou's Lemma** *Let f_n be a sequence of nonnegative measurable functions. Then*
$$\int \liminf_{n\to\infty} f_n\, d\mu \le \liminf_{n\to\infty} \int f_n\, d\mu.$$
3. **The Dominated Convergence Theorem** *Let f_n be a sequence of complex-valued integrable functions such that $\lim_{n\to\infty} f_n = f$ a.s., and for some integrable function g and all n, $|f_n| \le g$ a.s. Then f is integrable and*
$$\int \lim_{n\to\infty} f_n\, d\mu = \lim_{n\to\infty} \int f_n\, d\mu.$$

The following absolute convergence criterion is useful for showing completeness of normed spaces. First some definitions.

Definition B.3 *Let $\{x_n\}$ be a sequence in a normed vector space $\mathcal{X}$ with norm $\|\cdot\|$. The series $\sum_1^\infty x_n$* **converges** *if $\sum_1^N x_n \to x$ for some $x \in \mathcal{X}$. It is* **absolutely convergent** *if $\sum_1^\infty \|x_n\| < \infty$.*

Theorem B.4 *A normed vector space $\mathcal{X}$ is complete iff every absolutely convergent series in $\mathcal{X}$ converges.*

Proof Let $\mathcal{X}$ be a normed vector space and suppose every absolutely convergent series in $\mathcal{X}$ converges. Let $\{x_n\}$ be a Cauchy sequence in $\mathcal{X}$. Choose a subsequence $\{x_{n_k}\}$ such that $\|x_n - x_m\| < 2^{-k}$ for $n, m > n_k$. We write x_{n_k} as a series,

$$x_{n_k} = \sum_{j=1}^{k} x_{n_j} - x_{n_{j-1}},$$

taking $x_{n_0} = 0$ when needed on the right-hand side, and we find

$$\sum_{j=1}^{k} \|x_{n_j} - x_{n_{j-1}}\| = \|x_{n_1}\| + \sum_{j=2}^{\infty} \|x_{n_j} - x_{n_{j-1}}\| \le \|x_{n_1}\| + \sum_{j=1}^{\infty} 2^{-j} = \|x_{n_1}\| + 1 < \infty.$$

So the series is absolutely convergent, and thus $\{x_{n_k}\}$ converges. Let $x = \lim_{k\to\infty} x_{n_k}$, then x_n also converges to x since

$$\|x_n - x\| \le \|x_n - x_{n_k}\| + \|x_{n_k} - x\|,$$

and the two terms on the right can be made arbitrarily small for large n and n_k by the Cauchy criterion and convergence of $\{x_{n_k}\}$. Conversely, suppose $\mathcal{X}$ is complete and $\{x_n\}$ is a sequence in $\mathcal{X}$ such that $\sum_{k=1}^{\infty} x_n$ is absolutely convergent. Let $y_n = \sum_{k=1}^{n} x_n$, then for $n > m$ we have

$$\|y_n - y_m\| \leq \sum_{m+1}^{n} \|x_n\|,$$

where the right-hand side converges to 0 by absolute convergence. Thus the sequence of partial sums $\{y_n\}$ is Cauchy and hence the series $\sum_{k=1}^{\infty} x_n$ converges. □

B.2 The Hilbert Space $L^2(P)$

We show that $L^2(P)$, the space of second order random variables, is indeed a Hilbert space. The main assertion to be proved is completeness.

Theorem B.5 *Let $(\Omega, \mathcal{A}, P)$ be a probability measure space. Then $L^2(P)$ is a Hilbert space with norm and inner product for all $X, Y \in L^2(P)$*

$$\langle X, Y\rangle = EX^*Y, \qquad \|X\|^2 = EX^*X,$$

where $(\cdot)^*$ is complex conjugation for complex variables.

Proof $L^2(P)$ is a vector space if $X, Y \in L^2(P)$ then

$$|X+Y|^2 \leq (2(\max(|X|, |Y|))^2 \leq 2^2(|X|^2 + |Y|^2),$$

and thus $X+Y \in L^2(P)$. Furthermore, $|X^*Y| \leq |X+Y|^2$ so EX^*Y is finite and thus $\langle\cdot,\cdot\rangle$ is well defined. It is then straightforward to verify that $\langle\cdot,\cdot\rangle$ is indeed an inner product (see Definition 9.6).[1] It is left to prove that $L^2(P)$ is complete. We will use Theorem B.4. Let $\{X_k\}$ be a sequence of random variables in $L^2(P)$ such that $\sum_{k=1}^{\infty} \|X_k\| < \infty$. Let $a = \sum_{k=1}^{\infty} \|X_k\|$, $Y_n = \sum_{k=1}^{n} |X_k|$ and $Y = \sum_{k=1}^{\infty} |X_k|$. Then for all n, we have

$$\|Y_n\| \leq \sum_{k=1}^{n} \|X_k\| \leq \sum_{k=1}^{\infty} \|X_k\| = a,$$

and so by the monotone convergence theorem

$$\|Y\|^2 = E|Y|^2 = \lim_{n\to\infty} E|Y_n|^2 = \lim_{n\to\infty} \|Y_n\|^2 \leq a^2.$$

Thus $Y \in L^2(P)$. Also Y is finite a.s., otherwise $P(E) > 0$ where $E = \{\omega : Y(\omega) = \infty\}$ and if we let $E_n = \{\omega : Y(\omega) > n\}$ then $E = \bigcap_1^{\infty} E_n$ and so $P(E_n) > 0$ for all n. Thus we would have $E|Y|^2 \geq n^2 E(1_{E_n}) \geq n^2 E(1_E) = n^2 P(E)$ for all n which implies $E|Y|^2 \geq \infty$, a contradiction. Since $\mathbb{C}$ is complete and Y is finite a.s. then $\sum_{k=1}^{\infty} X_k$ converges a.s. by Theorem B.4. Now let $X = \sum_{k=1}^{\infty} X_k$, then we have $X \in L^2(P)$ and

[1] In order to satisfy $\langle X, Y\rangle = 0$ iff $X = Y$ in the definition, two random variables in $L^2(P)$ are considered equal if they are equal a.s. Therefore $L^2(P)$ is really a space of equivalence classes. Nothing further will be lost in our analysis by ignoring this minor point.

$|X - \sum_{k=1}^{n} X_k| \in L^2(P)$, since $|X| \le Y$ and $|X - \sum_{k=1}^{n} X_k| \le (2Y)^2$. So, by the dominated convergence theorem

$$\lim_{n\to\infty} \left\| X - \sum_{k=1}^{n} X_k \right\|^2 = \lim_{n\to\infty} E\left| X - \sum_{k=1}^{n} X_k \right|^2 = E \lim_{n\to\infty} \left| X - \sum_{k=1}^{n} X_k \right|^2 = 0.$$

So, $\sum_{n=1}^{\infty} X_k$ converges in $L^2(P)$ and hence $L^2(P)$ is complete by Theorem B.4. □

Theorem B.6 *If SOP $\{X(\omega, t), t \in T\}$ is q.m. continuous and $S \subseteq T$ is dense in T then the set of all random variables derived from linear operations on $\{X(t), t \in S\}$ is dense in $\mathcal{H}_X$. That is, $\forall Y(\omega) \in \mathcal{H}_X$:*

$$Y(\omega) \overset{\text{q.m.}}{=} \lim_{n\to\infty} \tilde{Y}_n(\omega) \quad \text{where } \forall n : \tilde{Y}_n(\omega) = \sum_{k=1}^{N_n} \alpha_{n,k} X(\omega, \tilde{t}_{n,k})$$

where all $\tilde{t}_{n,k}$ are in S.

Proof $Y(\omega)$ is in $\mathcal{H}_X$, so it is the limit of a sequence of $Y_n(\omega)$ of the form:

$$Y_n(\omega) = \sum_{k=1}^{N_n} \alpha_{n,k} X(\omega, t_{n,k}).$$

Let $\epsilon > 0$, and fix n large enough so that

$$\|Y(\omega) - Y_n(\omega)\| < \frac{\epsilon}{2} \quad \text{implying } C = \sum_{k=1}^{N_n} |\alpha_{n,k}|^2 > 0.$$

Now, since $\{X(t)\}$ is q.m. continuous, at every $t_{n,k}$ we may find a $\delta_{n,k}$ such that

$$d(t_{n,k}, s) < \delta_{n,k} \Rightarrow \|X(t_{n,k}) - X(s)\| < \frac{\epsilon^2}{4C} 2^{-k}.$$

Since S dense in T, for every k, let $\tilde{t}_{n,k} \in S$ such that $d(t_{n,k}, \tilde{t}_{n,k}) < \delta_{n,k}$.

Now, estimate

$$\|Y_n(\omega) - \tilde{Y}_n(\omega)\|^2 = \left\| \sum_{k=1}^{N_n} \alpha_{n,k} X(\omega, t_{n,k}) - \sum_{k=1}^{N_n} \alpha_{n,k} X(\omega, \tilde{t}_{n,k}) \right\|^2 \tag{B.1}$$

$$= \left\| \sum_{k=1}^{N_n} \alpha_{n,k} (X(\omega, t_{n,k}) - X(\omega, \tilde{t}_{n,k})) \right\|^2 \tag{B.2}$$

$$= E\left\{ \left| \sum_{k=1}^{N_n} \alpha_{n,k} (X(\omega, t_{n,k}) - X(\omega, \tilde{t}_{n,k})) \right|^2 \right\}. \tag{B.3}$$

The Cauchy-Schwarz inequality for $\mathbb{C}^N$ says

$$\left| \sum_{k=1}^{N} x_n y_n^* \right|^2 \le \left(\sum_{k=1}^{N} |x_n|^2 \right) \left(\sum_{k=1}^{N} |y_n|^2 \right), \tag{B.4}$$

so continuing we write

$$\leq E\left\{\left(\sum_{k=1}^{N_n}|\alpha_{n,k}|^2\right)\left(\sum_{k=1}^{N_n}|(X(\omega,t_{n,k})-X(\omega,\tilde{t}_{n,k})|^2\right)\right\} \tag{B.5}$$

$$=\left(\sum_{k=1}^{N_n}|\alpha_{n,k}|^2\right)\left(\sum_{k=1}^{N_n}E\left\{|(X(\omega,t_{n,k})-X(\omega,\tilde{t}_{n,k})|^2\right\}\right) \tag{B.6}$$

$$=C\sum_{k=1}^{N_n}\|(X(\omega,t_{n,k})-X(\omega,\tilde{t}_{n,k}))\|^2<\frac{\epsilon^2}{4}\sum_{k=1}^{N_n}2^{-k}<\left(\frac{\epsilon}{2}\right)^2. \tag{B.7}$$

Then $\|Y_n(\omega)-\tilde{Y}_n(\omega)\|<\epsilon/2$, implying

$$\|Y(\omega)-\tilde{Y}_n(\omega)\|\leq\|Y(\omega)-Y_n(\omega)\|+\|Y_n(\omega)-\tilde{Y}_n(\omega)\|<\frac{\epsilon}{2}+\frac{\epsilon}{2}=\epsilon \tag{B.8}$$

and we are done. □

C COMPACT OPERATORS AND SPECTRAL REPRESENTATION

C.1 Compact Operators

For this follow Reed and Simon's [215] particularly incisive development.

Assume operators $A : \mathcal{F} \to \mathcal{G}$ on Banach spaces[2] $\mathcal{F}$ and $\mathcal{G}$ with norms $\|\cdot\|_{\mathcal{F}}, \|\cdot\|_{\mathcal{G}}$. In almost all cases these are Hilbert spaces denoted by $\mathcal{H}_{\mathcal{F}}, \mathcal{H}_{\mathcal{G}}$, with associated inner products $\langle\cdot,\cdot\rangle_{\mathcal{F}}, \langle\cdot,\cdot\rangle_{\mathcal{G}}$. Then the bounded operators and their eigenfunctions are important.

Definition C.1 *The* **operator norm** $\|A\|$ *is defined to be*

$$\|A\| = \sup_{f\in\mathcal{F}} \frac{\|Af\|_{\mathcal{G}}}{\|f\|_{\mathcal{F}}} = \sup_{f\in\mathcal{F}:\|f\|_{\mathcal{F}}=1} \|Af\|_{\mathcal{G}}; \tag{C.1}$$

the **space of bounded norm, linear operators** *are denoted* $\mathcal{B}(\mathcal{F},\mathcal{G})$.

The subclass of bounded, linear operators, those which are compact play an important role in the process representations. It is the most important property of compact operators that they can essentially be represented via the superposition of a finite rank operator and a remainder which has an arbitrarily small norm.

Definition C.2 *An operator $A \in \mathcal{B}(\mathcal{F},\mathcal{G})$ is compact if A takes bounded subsets of $\mathcal{F}$ into precompact sets in $\mathcal{G}$. Equivalently, A is compact if and only for every bounded sequence $\{f_n\} \subset \mathcal{F}$, $\{Af_n\}$ has a subsequence convergent in $\mathcal{G}$.*

One of the most useful properties of compact operators is that their spectrum is discrete. This supports standard eigen expansion methods and orthogonal process representations.

Theorem C.3

1. *Let A be a finite dimensional operator on L^2, $A : L^2 \to L^2$ according to*

$$Af = \sum_{k=1}^{n} \alpha_k \langle\phi_k, f\rangle\langle\phi_k,\cdot\rangle\ , \tag{C.2}$$

where ϕ_k is a C.O.N. basis and $\alpha_i \in \mathbb{C}$. Then A is compact.

2. *Let A be a continuous operator, $A(x,y), (x,y) \in [0,1]^2$ mapping square integrable functions to square integrable functions, $A : L^2[0,1] \to L^2[0,1]$, according to*

$$Af(t) = \int_0^1 A(t,s)f(s)\, ds. \tag{C.3}$$

Then A is compact.

Proof

1. Define $\beta_k(f) = \alpha_k\langle\phi_k, f\rangle$ and let f_n be a bounded sequence in L^2. Then $\beta_k(f_n)$ is bounded for all $k = 1,\ldots,N$ since $|\beta_k(f_n)| \le |\alpha_k|\|f_n\|$ by Cauchy-Schwartz. Thus, choose a convergent subsequence $\beta_1(f_{n_k})$ of $\beta_1(f_n)$. Similarly, $\beta_2(f_{n_k})$ is

[2] A Banach space is a complete normed space not necessarily having an inner product consistent with the norm (such as is true for a Hilbert space).

bounded and thus choose a convergent subsequence $\beta_2(f_{n_{k_l}})$. Now, $\beta_1(f_{n_{k_l}})$ is also convergent since it is a subsequence of a convergent sequence. We continue this process a total of N times, obtaining a final subsequence f_{n_j} where $\beta_k(f_{n_j})$ converges for all $k = 1, \ldots, N$. Setting $\beta_k = \lim_{j\to\infty} \beta_k(f_{n_j})$ for each $k = 1, \ldots, N$ and $f = \sum_1^N \beta_k \phi_k$ we have

$$\lim_{j\to\infty} \left\| \sum_{k=1}^{N} \beta_k(f_{n_j})\phi_k - f \right\| = \lim_{j\to\infty} \sum_{k=1}^{N} \left| \beta_k(f_{n_j}) - \beta_k \right| = 0 \,. \tag{C.4}$$

Thus, $Af_{n_j} \to f$ in L^2 and we conclude that A is compact.

2. For all $\phi \in L^2$, we have

$$\|A\phi\|^2 = \int |\langle A(t,\cdot), \phi(\cdot)\rangle|^2 dt \le \int \|A(t,\cdot)\|^2 dt = \|A(\cdot,\cdot)\|^2 \|\phi\|^2, \tag{C.5}$$

and also $\|A(\cdot,\cdot)\| < \infty$ since $A(\cdot,\cdot)$ is continuous on a compact set. So A is a bounded (equivalently continuous) operator. Now define $B_K = \{\phi : \|\phi\| \le K\}$. $A(\cdot,\cdot)$ is uniformly continuous on $[0,1]^2$, since continuity extends to uniform continuity on compact sets. Thus for all $\epsilon > 0$ there exists $\delta > 0$ such that for all $s \in [0,1]$

$$|t - t'| < \delta \Rightarrow |A(t,s) - A(t',s)| < \epsilon \,. \tag{C.6}$$

So if $\phi \in B_K$ and $|t - t'| < \delta$ then

$$\begin{aligned} \left|A\phi(t) - A\phi(t')\right| &= \left| \int \Big(A(t,s) - A(t',s)\Big)\phi(s)\, ds \right| \le \sup_{s\in[0,1]} |A(t,s) - A(t',s)| \\ &\quad \times \int |\phi(s)|\, ds \le \epsilon \int |\phi(s)|\, ds \\ &= \epsilon |\langle |\phi|, 1_{[0,1]}\rangle| \le \epsilon \|\phi\| \|1_{[0,1]}\| = \epsilon K. \end{aligned} \tag{C.7}$$

Thus we have established that the family $A[B_K]$ are uniformly equicontinuous.[3] We also have uniform boundedness of $A[B_K]$ since

$$|A\phi(t)| \le \sup_{(t,s)\in[0,1]^2} |A(t,s)| K$$

by the same argument as in Eqn. (C.7). So, by the the Arzela Ascoli theorem [4] for every sequence $\{\phi_n\} \subset B_K$, the sequence $\{A\phi_n\} \subset A[B_K]$ has a convergent subsequence.

Field case: The identical proof holds for the field case with absolute value $|\cdot|$ on elements in $[0,1]$ replaced by $\|\cdot\|_{\mathbb{R}^d}$ for elements in D. □

[3] A family or sequence of functions is said to be **uniformly equicontinuous** if for all $\epsilon > 0$ there exists $\delta > 0$ such that $|t - s| < \delta$ implies $|f(t) - f(s)| < \epsilon$ for all f in the family (sequence).

[4] Let f_n be uniformly bounded and uniformly equicontinuous on a compact interval. Then there exists a uniformly convergent subsequence. See [478] page 312 for a proof.

C.2 Riesz–Schauder and Hilbert–Schmidt

Now it follows that self-adjoint compact operators can be written in the span of their eigenvectors.

Theorem C.4 Riesz–Schauder and Hilbert–Schmidt Theorem *(a) Let A be compact in $\mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$. Then the spectrum $\sigma(A)$ is a discrete set of eigenvalues having no limit points except possibly $\lambda = 0$. Any $\lambda \in \sigma(A)$ is an eigenvalue, with the corresponding space of eigenvectors of at most finite dimension.*

(b) Let A be compact on $\mathcal{B}(\mathcal{H})$, and $\{\phi_n\}$ an orthonormal set in $\mathcal{H}$. Then

$$A = \lim_{N\to\infty} \sum_{n=1}^{N} A\phi_n \langle \phi_n, \cdot \rangle \ . \tag{C.8}$$

(c) Let A be self adjoint and compact on $\mathcal{B}(\mathcal{H})$. Then the set of eigenvectors $\{\phi_n\}$ satisfying $A\phi_n = \lambda_n \phi_n$ are a C.O.N. basis for $\mathcal{H}$ with $\lambda_n \to 0$ as $n \to 0$, and

$$A = \lim_{N\to\infty} \sum_{n=1}^{N} \lambda_n \phi_n \langle \phi_n, \cdot \rangle \ . \tag{C.9}$$

Here the limit is in the operator norm sense $\lim_{N\to\infty} \sup_x \|Ax - \sum_{k=1}^{N} \lambda_k \times \langle \phi_k, x\rangle \phi_k\| = 0$.

Proof Part (a) see Theorem VI.15 p. 203 of [215]. For part (b), the sequence $\{\alpha_n\}$ defined by

$$\alpha_n = \sup_{\psi \in [\phi_1, \phi_2, \ldots, \phi_n]^{\perp}} \|A\psi\| \tag{C.10}$$

is monotone decreasing and therefore converges to a limit $\alpha \geq 0$. To see $\alpha = 0$, choose $\psi_n \in [\phi_1, \ldots, \phi_n]^{\perp}$, $|A\psi_n| \geq \alpha/2$. Since $\psi_n \overset{w}{\to} 0$, $A\psi_n \to 0$ in norm (theorem VI.11 p. 197 of [215]). This implies $\alpha = 0$. We note however,

$$\begin{aligned}
\|A - \sum_{j=1}^{n} \langle \phi_j, \cdot \rangle A\phi_j\| &= \sup_{\|f\|=1} \left\| Af - \sum_{j=1}^{n} \langle \phi_j, f\rangle A\phi_j \right\| \\
&= \sup_{\|f\|=1} \left\| \sum_{j=1}^{\infty} \langle \phi_j, f\rangle A\phi_j - \sum_{j=1}^{n} \langle \phi_j, f\rangle A\phi_j \right\| \\
&= \sup_{\|f\|=1} \left\| \sum_{j=n+1}^{\infty} \langle \phi_j, f\rangle A\phi_j \right\| \\
&= \alpha_n .
\end{aligned} \tag{C.11}$$

This implies $\sum_{j=1}^{n} \langle \phi_j, \cdot \rangle A\phi_j \to A$ in operator norm as $n \to \infty$.

(c) The proof of the first part establishes $\{\phi_k\}$ are a C.O.N for $\mathcal{H}$, with $\lambda_n \to 0$. For this see [215], p. 203. The second part follows from part (b) of Theorem 9.23 using the eigenfunctions as the C.O.N. □

D HILLE-YOSIDA THEORY

We now do all this more carefully. Let B be a Banach space with $sup - norm$, E be the state space, $\{X(t), t \in T\}$ be an E-valued Markov process, $\{H_t, t \geq 0\}$ its semigroup.

Theorem D.1 *Let $f \in B$ (the space of bounded measurable functions on E) and H_t be defined as above, then the semi-group has the following three properties:*

1. *H_t takes an element of B to an element of B:*
$$H_t : B \to B. \tag{D.1}$$
2. *H_t is continuous in f, in other words, if there is a sequence of functions (f_n) converging to f, then the sequence $(H_t f_n)$ also converges and it converges to $H_t f$.*
3. *H_t on B is strongly continuous,i.e.*
$$\lim_{t \to 0} H_t f = f \quad \forall f \in B. \tag{D.2}$$
4. *$H_t f$ is continuous in t for all $f \in B$.*

Proof H_t is defined as

$$H_t = \int P_t(x, dy) f(y).$$

Since $f(y) \in B$, $\|f\|$ is finite implying for all $t > 0$,

$$\|H_t f\| = \left\| \int P_t(x, dy) f(y) \right\| \leq \|f\| \int P_t(x, dy) \leq \|f\|.$$

This is the contraction property of the semi-group.

That $H_t f$ is continuous in f follows from

$$\begin{aligned} \|H_t f - H_t f_n\| &= \|H_t(f - f_n)\|, \text{ linearity} \\ &\leq \|f - f_n\|, \text{ contraction property.} \end{aligned}$$

That H_t is strongly continuous follows from $f \in \bar{C} \subset B$ implies

$$\begin{aligned} \lim_{t \downarrow 0} H_t f(x) &= \lim_{t \downarrow 0} \int P_t(x, dy) f(y) \\ &= \int \lim_{t \downarrow 0} P_t(x, dy) f(y) \\ &= \int \delta_x(dy) f(y) = f(x). \end{aligned}$$

Having proved it for bounded, continuous functions we now prove it for bounded functions only. Let $f \in B$, and (f_n) be a sequence of bounded continuous functions converging to f, then

$$\begin{aligned} \|H_t f - f\| &= \|H_t f - H_t f_n + H_t f_n - f\| \\ &\leq \|H_t f - H_t f_n\| + \|H_t f_n - f\| \\ &\leq \|f - f_n\| + \|H_t f_n - f\| \end{aligned}$$

which by strong continuity implies $\lim_{t \downarrow 0} \|H_t f - f\| = 0$ for all $f \in B$.

Now use this to prove the continuity of $H_t f$,

$$\lim_{\nu\downarrow 0} \|H_{t+\nu}f - H_t f\| = \lim_{\nu\downarrow 0} \|H_t(H_\nu f) - H_t f\| \text{ (semi-group property)}$$

$$= \lim_{\nu\downarrow 0} \|H_t(H_\nu f - f)\|, \text{ (linearity of } H_t)$$

$$\lim_{\nu\downarrow 0} \leq \|H_\nu f - f\|, \text{ (contraction property)}$$

$$= 0 \text{ (strong continuity)}.$$

□

Definition D.2 *The generator of a semigroup* $\{H_t, t \geq 0\}$ *on* B *is the linear operator* A *defined by*

$$Af = \lim_{t\to 0} \frac{H_t f - f}{t} \tag{D.3}$$

with domain $D \subset B$ *a subspace of* B *such that*

$$D = \left\{ f \in B : \lim_{t\downarrow 0} \frac{H_t f - f}{t} \overset{sup\text{-}norm}{=} Af \in B \right\}.$$

Theorem D.3 *D is dense in B with sup-norm being the distance measure.*

Proof Let $f \in B$ and define

$$f_t = \frac{1}{t} \int_0^t H_s f \, ds.$$

Since $H_s f$ is continuous in s, use the intermediate value theorem to write $f_t = H_{\tilde{s}} f)$, $\tilde{s} \in [0, t]$. By part 1 of the next theorem $\int_0^t H_s f \, ds \in D$ therefore $1/t \int_0^t H_s f \, ds \in D$ for $t > 0$. From a strong continuity of the semi-group we know that $\lim_{t\downarrow 0} H_t = I$. Now construct a sequence of functions $\{f_{1/n}, n = 1, 2, \ldots\}$. From the strong continuity and the above claim we have defined a countable sequence in D converging to $f \in B$. Therefore D is dense in B. □

Theorem D.4 *Let* $\{H_t, t \geq 0\}$ *be a strongly continuous semigroup on B with the generator A, and if* $f \in B$ *then,*

1. $\int_0^t H_s f \, ds \in D$;
2. *if* $f \in D$, *then* $(d/dt)H_t f = AH_t f$.

Proof

1. To prove that $f_t = \int_0^t H_s f \, ds \in D$, we operate A on it and check if the result makes sense,

$$Af_t = \lim_{r\downarrow 0} \frac{1}{r} \left(H_r \left(\int_0^t H_s f \, ds \right) - \int_0^t H_s f \, ds \right)$$

$$= \lim_{r\downarrow 0} \frac{1}{r} \left(\int_0^t H_{r+s} f \, ds - \int_0^t H_s f \, ds \right)$$

$$= \lim_{r\downarrow 0} \frac{1}{r} \left(\int_r^{t+r} H_s f \, ds - \int_0^t H_s f \, ds \right)$$

$$= \lim_{r\downarrow 0} \frac{1}{r}\left(\int_r^t H_s f\, ds + \int_t^{t+r} H_s f\, ds - \int_0^r H_s f\, ds - \int_r^t H_s f\, ds\right)$$

$$= \lim_{r\downarrow 0} \frac{1}{r}\left(\int_t^{t+r} H_s f\, ds - \int_0^r H_s f\, ds\right)$$

$$= H_t f - f\,.$$

As $H_t f$, f are in B, therefore Af_t is in B implying $f_t = \int_0^t H_s f\, ds$ is in D.

2. For $f \in D$,

$$\frac{d}{dt} H_t f = \lim_{\epsilon\downarrow 0}\left(\frac{H_{t+\epsilon} - H_t}{\epsilon}\right) f$$

$$= \lim_{\epsilon\downarrow 0}\left(\frac{H_\epsilon - I}{\epsilon} H_t\right) f$$

$$= \lim_{\epsilon\downarrow 0}\left(\frac{H_\epsilon - I}{\epsilon}\right) H_t f$$

$$= AH_t f$$

with the last equation holding as long as $H_t f \in D$. Clearly it is as

$$AH_t f = \lim_{\epsilon\downarrow 0} \frac{1}{\epsilon}(H_\epsilon H_t f - H_t f)$$

$$= H_t \lim_{\epsilon\downarrow 0} \frac{1}{\epsilon}(H_\epsilon f - f)$$

$$= H_t A f$$

with the last step following from the fact that $f \in D$ implying that $1/\epsilon(H_\epsilon f - f)$ converges in sup-norm in B. Thus $H_t f \in D$. As $A : D \to B$ and $H_t : B \to B$, $H_t A f \in B$ is well defined and is equal to $AH_t f$, therefore $AH_t f \in B$ and $H_t f \in D$. □

D.1 Resolvent of a semigroup R_λ

Definition D.5 *For $f \in B$ and $\lambda > 0$, we define*

$$(R_\lambda f)(a) = \int_0^\infty e^{-\lambda t} H_t f(a)\, dt \tag{D.4}$$

The family of operators $\{R_\lambda, 0 < \lambda < \infty\}$ is called the **Resolvent** *of the semigroup $\{H_t, 0 \le t < 0\}$.*

The resolvent is useful in the study of the generator A.

Theorem D.6 *$\forall g \in B, \lambda > 0$,*

- $R_\lambda g = \int_0^\infty e^{-\lambda t} H_t g\, dt \in D$

- $f = R_\lambda g$ *is the unique solution to the equation,*

$$\lambda f - Af = g, \quad \textit{where } f \in D \tag{D.5}$$

Proof First, we want to verify that $R_\lambda g \in D$, $\forall g \in \mathbf{B}$.

$$\begin{aligned}\frac{1}{t}(H_t R_\lambda g - R_\lambda g) &= \frac{1}{t}\int_0^\infty e^{-\lambda s}(H_{t+s}g - H_s g)\,ds \\ &= \frac{1}{t}\left(\int_t^\infty e^{\lambda t}e^{-\lambda r}H_r g\,dr - \int_0^\infty e^{-\lambda s}H_s g\,ds\right), \quad \text{where } r = s+t \\ &= \frac{1}{t}(e^{\lambda t}-1)\int_0^\infty e^{-\lambda s}H_s g\,ds - \frac{e^{\lambda t}}{t}\int_0^t e^{-\lambda s}H_s g\,ds. \end{aligned} \tag{D.6}$$

The last equality is obtained by changing the name of the dummy variable from r to s. By L'Hôpital's Rule,

$$\lim_{t\downarrow 0}\frac{1}{t}(e^{\lambda t}-1) = \lambda. \tag{D.7}$$

Hence the first term on the right-hand side of Eqn. D.6 becomes $\lambda R_\lambda g$ as $t \to 0$. Using the Mean Value Theorem,

$$\frac{1}{t}\int_0^t e^{-\lambda s}H_s g\,ds = e^{-\lambda\tilde{t}}H_{\tilde{t}}g, \quad \text{for some } \tilde{t},\ 0 < \tilde{t} < t. \tag{D.8}$$

When t goes to 0, $\tilde{t}$ also goes to 0, hence, the second term on the right-hand side of Equation D.6 becomes $H_0 g = g$. Therefore,

$$AR_\lambda g = \lambda R_\lambda g - g. \tag{D.9}$$

That is, the limit $AR_\lambda g = \lim_{t\downarrow 0} 1/t(H_t R_\lambda g - R_\lambda g)$ exists and belongs to $\mathbf{B}$, which means $R_\lambda g \in D$. Next, we prove the uniqueness of the solution $R_\lambda g$.

Suppose that $f_1 \in D$ and $f_2 \in D$ are two solutions to D.5, then

$$(\lambda I - A)f_1 = g \tag{D.10}$$

$$(\lambda I - A)f_2 = g. \tag{D.11}$$

Let $\varphi = f_1 - f_2$, then $\varphi \in D$ and

$$(\lambda I - A)\varphi = 0. \tag{D.12}$$

Therefore,

$$\begin{aligned}\frac{d}{dt}e^{-\lambda t}H_t\varphi &= -\lambda e^{-\lambda t}H_t\varphi + e^{-\lambda t}AH_t\varphi, \quad \text{by property (f)} \\ &= e^{-\lambda t}H_t(-\lambda I + A)\varphi \\ &= 0, \qquad\qquad \text{by Eqn. D.12.}\end{aligned} \tag{D.13}$$

Thus, $(d/dt)e^{-\lambda t}H_t\varphi = 0$. Then

$$e^{-\lambda t}H_t\varphi - H_0\varphi = 0. \tag{D.14}$$

This is obtained by integrating the preceding differential equation on both of the sides from 0 to t. Hence

$$e^{-\lambda t}H_t\varphi = \varphi. \tag{D.15}$$

Then it follows that

$$0 \le \|\varphi\| = e^{-\lambda t}\|H_t\varphi\| \le e^{-\lambda t}\|\varphi\| \xrightarrow{t\to\infty} 0, \tag{D.16}$$

so $\|\varphi\| = 0$, then $\varphi = 0$ by the definition of norm. Therefore $f = R_\lambda g$ is the unique solution for $(\lambda I - A)f = g$.

□

Reordering Equation D.9 gives us

$$g = (\lambda I - A)R_\lambda g, \tag{D.17}$$

implying that R_λ works as the inverse operator of $(\lambda I - A)$, i.e.

$$R_\lambda = \frac{1}{\lambda - A}. \tag{D.18}$$

Note that $\int_0^\infty e^{-\lambda t} H_t f dt = (1/\lambda - A)f$ is still true, even though in general $H_t \neq e^{tA}$. Rearranging Equation D.9, we get

$$A = \lambda I - R_\lambda^{-1}. \tag{D.19}$$

The next theorem tells us we can construct the semigroup $\{H_t, 0 \le t < \infty\}$ by using R_λ.

Theorem D.7 *Let $\{X(t), 0 \le t < \infty\}$ be a Markov process with stationary transition function $P_t(x,E) = \Pr\{X(t) \in D | X_0 = x\}$ and continuous in probability. Then the transition law $P_t(x,\cdot)$ or equivalently its semigroup $\{H_t, 0 \le t < \infty\}$ is uniquely determined by its generator A.*

Proof As mentioned before, we define the new generator A_λ as follows:

$$A_\lambda = \lambda A R_\lambda = \lambda A(\lambda - A)^{-1}. \tag{D.20}$$

Claim (1). $A_\lambda f \xrightarrow{\lambda\to\infty} Af$, $\forall f \in D$. (2). A_λ is a bounded operator and $\lim_{\lambda\to\infty} e^{tA_\lambda} = H_t$ strongly, i.e., $\forall f \in \mathbf{B}$, $\lim_{\lambda\to\infty} e^{tA_\lambda} f = H_t f$.

First, we are going to prove (1). Using the definition of R_λ, we note that

$$\begin{aligned} \lambda R_\lambda f &= \lambda \int_0^\infty e^{-\lambda t} H_t f\, dt \\ &= \int_0^\infty e^{-r} H_{r/\lambda} f\, dr, \end{aligned} \tag{D.21}$$

where $r = t\lambda$, hence $dr = \lambda\, dt$. So,

$$\begin{aligned} \|\lambda R_\lambda f\| &= \left\| \int_0^\infty e^{-r} H_{r/\lambda} f\, dr \right\| \\ &\le \int_0^\infty e^{-r} \|H_{r/\lambda} f\|\, dr \\ &\le \int_0^\infty e^{-r} \|f\|\, dr = \|f\|, \quad \text{by the contraction property of } H_t, \end{aligned} \tag{D.22}$$

and

$$\begin{aligned} \|\lambda R_\lambda f - f\| &= \left\| \int_0^\infty e^{-t} (H_{t/\lambda} f - f)\, dt \right\| \\ &\le \int_0^\infty e^{-t} \|H_{t/\lambda} f - f\|\, dt. \end{aligned}$$

Let $g_t^{(\lambda)} = H_{t/\lambda}f - f$. Claim $\|g_t^{(\lambda)}\|$ is bounded, since

$$\begin{aligned}\|g_t^{(\lambda)}\| &= \|H_{t/\lambda}f - f\| \\ &\le \|H_{t/\lambda}f\| + \|f\| \le 2\|f\|.\end{aligned} \tag{D.23}$$

Also, $\|g_t^{(\lambda)}\| \overset{\lambda\to\infty}{\longrightarrow} 0$, since $H_{t/\lambda}f \overset{\lambda\to\infty}{\longrightarrow} f$ by property (e). Then,

$$0 \le \lim_{\lambda\to\infty} \|\lambda R_\lambda f - f\| \le \lim_{\lambda\to\infty} \int_0^\infty e^{-t}\|g_t^{(\lambda)}\|\,dt. \tag{D.24}$$

Applying Dominated Convergence Theorem(see Wheeden [479], p. 71) on $\|g_t^{(\lambda)}\|$ in the above equation, we have

$$\lim_{\lambda\to\infty} \int_0^\infty e^{-t}\|g_t^{(\lambda)}\|\,dt = \int_0^\infty e^{-t} \lim_{\lambda\to\infty} \|g_t^{(\lambda)}\|\,dt = 0. \tag{D.25}$$

It follows that

$$\lambda R_\lambda \overset{\lambda\to\infty}{\longrightarrow} I. \tag{D.26}$$

Hence

$$\forall f \in D, \quad A_\lambda = \lambda A R_\lambda \overset{\lambda\to\infty}{\longrightarrow} A. \tag{D.27}$$

Next, we prove that A_λ is a bounded operator. By the definition of the operator A_λ, we have

$$\|A_\lambda f\| = \|\lambda A R_\lambda f\|. \tag{D.28}$$

Substituting Equation D.19 into the above equality gives

$$\begin{aligned}\|A_\lambda f\| &= \lambda\|(\lambda I - R_\lambda^{-1})R_\lambda f\| = \lambda\|\lambda R_\lambda f - f\| \\ &\le \lambda\|\lambda R_\lambda f\| + \lambda\|f\| \le 2\lambda\|f\|,\end{aligned} \tag{D.29}$$

since $\|\lambda R_\lambda f\| \le \|f\|$ as proved in Equation D.22. Before we can prove that $\lim_{\lambda\to\infty} e^{tA_\lambda} = H_t$, the following propositions are introduced so that we can use them later in the proof. □

Proposition D.8

(a) *Resolvent commutes:* $R_\lambda R_\mu f = R_\mu R_\lambda f, \forall f \in \mathbf{B}$.

(b) *Generator commutes:* $A_\lambda A_\mu f = A_\mu A_\lambda f, \forall f \in \mathbf{B}$.

Proof First we will show that $(\lambda - A)(\mu - A) = (\mu - A)(\lambda - A)$.

$$\begin{aligned}(\lambda - A)(\mu - A) &= \lambda\mu - A\mu - \lambda A + AA \\ &= \mu\lambda - \mu A - A\lambda + AA \text{ since } \lambda \text{ and } \mu \text{ are just constants} \\ &= (\mu - A)(\lambda - A).\end{aligned} \tag{D.30}$$

Let us prove (a) first. From Eqn. D.30,

$$\begin{aligned}&(\lambda - A)(\mu - A)(\mu - A)^{-1} = (\mu - A)(\lambda - A)(\mu - A)^{-1} \\ &\implies (\lambda - A) = (\mu - A)(\lambda - A)(\mu - A)^{-1} \\ &\implies I = (\mu - A)(\lambda - A)(\mu - A)^{-1}(\lambda - A)^{-1} \\ &\implies (\mu - A)^{-1} = (\lambda - A)(\mu - A)^{-1}(\lambda - A)^{-1} \\ &\implies (\lambda - A)^{-1}(\mu - A)^{-1} = (\mu - A)^{-1}(\lambda - A)^{-1},\end{aligned} \tag{D.31}$$

which means $R_\lambda R_\mu = R_\mu R_\lambda$ because $R_\mu = (1/\mu - A)$ and $R_\lambda = (1/\lambda - A)$.

Next we want to show that the generator A_λ commutes. Since $A_\lambda = \lambda A R_\lambda$ and $(\lambda - A)R_\lambda = I$, it follows that

$$\lambda R_\lambda - AR_\lambda = I \Longrightarrow \lambda^2 R_\lambda - \lambda A R_\lambda = \lambda I \Longrightarrow A_\lambda = \lambda^2 R_\lambda - \lambda I. \tag{D.32}$$

Then the commutation of A_λ follows from that of R_λ. □

Remark D.1.0 The use of $R_\lambda Af = AR_\lambda f, \forall f \in D$ in the proof the commutation of A_λ restricts f to D space. So in the class we only proved that $A_\lambda A_\mu f = A_\mu A_\lambda f,\ \forall f \in D,$ not in **B**.

Proposition D.9 *Define $H_t^\lambda = e^{A_\lambda t}$, then H_t^λ is a strongly continuous bounded contraction semigroup on* **B**, *and*

$$\|H_t^\lambda f - H_t^\mu f\| = \|e^{A_\lambda t} f - e^{A_\mu t} f\| \le t \|A_\lambda f - A_\mu f\|, \quad \forall f \in \mathbf{B}. \tag{D.33}$$

Proof For all $f \in \mathbf{B}$, $\lim_{t\to o} H_t^\lambda f = \lim_{t \to o} e^{tA_\lambda} f = f$. Hence, H_t^λ is strongly continuous.

For boundedness of H_t^λ, we use the fact that A_λ is bounded.

$$\|H_t^\lambda f\| = \|e^{tA_\lambda} f\| = \left\| \sum_{n=0}^{\infty} \frac{(tA_\lambda)^n}{n!} \right\| \le \sum_{n=0}^{\infty} \frac{(t2\lambda)^n}{n!} \|f\| = e^{2t\lambda}\|f\|, \tag{D.34}$$

since $\|A_\lambda f\| \le 2\lambda \|f\|$, implying $\|A_\lambda^n f\| = \|A_\lambda^{n-1}(A_\lambda f)\| \le \|A_\lambda^{n-1} f\|$.

Using Equation D.32

$$\|e^{tA_\lambda}\| = e^{-t\lambda} \|e^{t\lambda^2 R_\lambda}\| \le e^{-t\lambda} e^{t\lambda^2 \|R_\lambda\|} \le 1. \tag{D.35}$$

The last inequality results from Equation D.22, i.e. $\|R_\lambda f\| \le (1/\lambda)\|f\|$. This means H_t^λ is a contraction mapping.

The semigroup property of H_t^λ is shown as following:

$$H_{t+s}^\lambda = e^{(t+s)A_\lambda} = e^{tA_\lambda} e^{sA_\lambda} = H_t^\lambda H_s^\lambda = H_s^\lambda H_t^\lambda. \tag{D.36}$$

Since $e^{tA_\lambda} = \sum_{n=0}^\infty (tA_\lambda)^n / n!$,

$$\frac{d}{dt} e^{tA_\lambda} = e^{tA_\lambda} A_\lambda = A_\lambda e^{tA_\lambda}. \tag{D.37}$$

This will be used in the following proof.

$$\begin{aligned} e^{tA_\lambda} f - e^{tA_\mu} f &= \int_0^t \frac{d}{ds} (e^{sA_\lambda} e^{(t-s)A_\mu}) f \, ds \\ &= \int_0^t (e^{sA_\lambda} A_\lambda e^{(t-s)A_\mu} - e^{sA_\lambda} e^{(t-s)A_\mu} A_\mu) f \, ds \\ &= \int_0^t (e^{sA_\lambda} e^{(t-s)A_\mu} A_\lambda - e^{sA_\lambda} e^{(t-s)A_\mu} A_\mu) f \, ds \quad \text{commutation of } A_\lambda \end{aligned} \tag{D.38}$$

$$= \int_0^t e^{sA_\lambda} e^{(t-s)A_\mu} (A_\lambda - A_\mu) f \, ds. \tag{D.39}$$

Then

$$\|e^{tA_\lambda}f - e^{tA_\mu}f\| \le \int_0^t \|e^{sA_\lambda}e^{(t-s)A_\mu}(A_\lambda f - A_\mu f)\|\,ds$$
$$\le \int_0^t \|e^{(t-s)A_\mu}(A_\lambda f - A_\mu f)\|\,ds.$$

The last inequality results from the contraction property of H_s^λ, i.e. $\|H_s^\lambda f\| \le \|f\|$. Using the contraction property again on H_{t-s}^μ, the following inequality results:

$$\|e^{tA_\lambda}f - e^{tA_\mu}f\| \le \int_0^t \|A_\lambda f - A_\mu f\|\,ds = t\|A_\lambda f - A_\mu f\| \tag{D.40}$$

□

Since $A_\lambda \overset{\lambda\to\infty}{\longrightarrow} A$, then $\forall f \in D$, $\|A_\lambda f - A_\mu f\| \to 0$ as λ and μ go to ∞, implying H_t^λ is a Cauchy sequence in $\mathbf{B}$, hence convergent $\forall f \in D$. Claim $\forall f \in \mathbf{B}$, $H_t^\lambda f$ is also convergent and denote the limit as $\widetilde{H}_t f$. The proof goes as follows: since D is dense in $\mathbf{B}$, $\forall f \in \mathbf{B}$, there exists a sequence $(f_k) \subseteq D$, so that $(f_k) \overset{k\to\infty}{\longrightarrow} f$. Then,

$$\forall f \in \mathbf{B},\quad \|H_t f - H_t^\lambda f\| = \|H_t f - H_t f_k + H_t f_k - H_t^\lambda f_k + H_t^\lambda f_k - H_t^\lambda f\|$$
$$\le \|H_t(f - f_k)\| + \|(H_t - H_t^\lambda)f_k\| + \|H_t^\lambda(f_k - f)\|, \tag{D.41}$$

which has a limit zero as λ goes to ∞. Denote $\lim_{\lambda\to\infty} H_t^\lambda f = \widetilde{H}_t f$, hence $\widetilde{H}_t$ maps from $\mathbf{B}$ to $\mathbf{B}$. Claim $\widetilde{H}_t$ is also a contraction semigroup with generator A.

We will prove the semigroup property of $\widetilde{H}_t$ and leave the contraction property as the Homework. Since

$$\widetilde{H}_{s+t}f - \widetilde{H}_s\widetilde{H}_t f = (\widetilde{H}_{s+t} - H_{s+t}^\lambda)f + H_s^\lambda(H_t^\lambda - \widetilde{H}_t)f + (H_s^\lambda - \widetilde{H}_s)\widetilde{H}_t f, \tag{D.42}$$

$$\|\widetilde{H}_{s+t}f - \widetilde{H}_s\widetilde{H}_t f\| \le \|(\widetilde{H}_{s+t} - H_{s+t}^\lambda)f\| + \|H_s^\lambda(H_t^\lambda - \widetilde{H}_t)f\| + \|(H_s^\lambda - \widetilde{H}_s)\widetilde{H}_t f\|. \tag{D.43}$$

Let $\lambda \to \infty$, take the limit on both sides of the above equation gives: $\|\widetilde{H}_{s+t}f - \widetilde{H}_s\widetilde{H}_t f\| = 0$, since $\widetilde{H}_t$ is the limit of H_t^λ as $\lambda \to \infty$. Next, the strong continuity property of $\widetilde{H}_t$ is proved.

$$\|\widetilde{H}_s f - f\| = \|(\widetilde{H}_s - H_s^\lambda)f + H_s^\lambda f - f\| \le \|(\widetilde{H}_s - H_s^\lambda)f\| + \|H_s^\lambda f - f\|. \tag{D.44}$$

First, take the limit on both sides of the above equation as $\lambda \to \infty$, we have $\|\widetilde{H}_s f - f\| \le \|H_s^\lambda f - f\|$ by the fact that $\lim_{\lambda\to\infty} H_s^\lambda = \widetilde{H}_s$. Then take the limit again as $s \to 0$, it follows that

$$\lim_{s\to 0} \|\widetilde{H}_s f - f\| = 0 \tag{D.45}$$

by the strong continuity of H_s^λ.

Now, we are ready to show that A is the generator of semigroup $\{\widetilde{H}_t, 0 \le t < \infty\}$, i.e. $\lim_{\epsilon\to 0}(\widetilde{H}_\epsilon - I)/\epsilon = A$, which is equivalent to prove that

$$\frac{d\widetilde{H}_t}{dt} = A\widetilde{H}_t. \tag{D.46}$$

By Equation D.37, $(d/dt)H_t^\lambda = A_\lambda H_t^\lambda = H_t^\lambda A_\lambda$, which is equivalent to

$$H_t^\lambda f - f = \int_0^t H_s^\lambda A_\lambda f\,ds. \tag{D.47}$$

If we can show that

$$\lim_{\lambda\to\infty} H_s^\lambda A_\lambda = \widetilde{H}_s A, \tag{D.48}$$

then Equation D.47 implies that

$$\widetilde{H}_t f - f = \int_0^t \widetilde{H}_s A\, ds, \tag{D.49}$$

which means $(d\widetilde{H}_t/dt) = \widetilde{H}_t A$, i.e. A is the generator of the semigroup $\widetilde{H}_t$.

To prove Equation D.48, $\forall f \in D$.

$$\begin{aligned}\|H_t^\lambda A_\lambda f - \widetilde{H}_t A f\| &= \|H_t^\lambda (A_\lambda f - Af) + (H_t^\lambda - \widetilde{H}_t)Af\| \\ &\le \|H_r^\lambda (A_\lambda f - Af)\| + \|(H_t^\lambda - \widetilde{H}_t)Af\|.\end{aligned}$$

As λ goes to ∞, then the first term goes to zero for all $f \in D$, because $\lim_{\lambda\to\infty} A_\lambda f = Af$, $\forall f \in D$. The second term goes to zero, since $H_t^\lambda \overset{\lambda\to\infty}{\longrightarrow} \widetilde{H}_t$, implying $H_s^\lambda A_\lambda f \overset{\lambda\to\infty}{\longrightarrow} \widetilde{H}_s Af$, $\forall f \in D$.

D.2 Problems

1. Prove that $\widetilde{H}_t : \mathbf{B} \to \mathbf{B}$ is a contraction mapping.

Proof

$$\|\widetilde{H}_t f\| \le \|\widetilde{H}_t f - H_t^\lambda f\| + \|H_t^\lambda f\| \le \|(\widetilde{H}_t - H_t^\lambda)f\| + \|f\| \tag{D.50}$$

The last inequality results from the contraction property of H_t^λ. Let $\lambda \to \infty$ on both sides of the above inequality, we have

$$\|\widetilde{H}_t f\| \le \|f\|. \tag{D.51}$$

□

E ERGODIC PROPERTIES OF JUMP-DIFFUSIONS

E.1 Operator Perturbation Theorem (extension of Ethier and Kurtz)

First define notation for the subsets of measurable functions. Let $B(E)$ be the set of bounded, measurable functions on domain E, and the set of continuous functions $C(E)$ with $\bar{C}(E)$ being both bounded and continuous. Let $C_c^\infty(E)$ be the set of infinitely differentiable functions with compact support, and $\hat{C}(E)$ the set of continuous functions vanishing at infinity. Also define $\mathcal{P}(E)$ the set of probability measures on E, and the characteristic function $1_A(x)$ of set A.

Theorem E.1 *Let $(E_k, r_k), k = 0, 1, \ldots$ be complete, separable metric spaces, with $E = \cup_k E_k$ and metric $r(x, y) = r_k(x, y)$ for $x, y \in E_k$, and 1 otherwise ((E, r) is complete and separable). Suppose the closure of the infinitesimal generator of the semigroup $A_k^d \subset \bar{C}(E_k) \times \bar{C}(E_k)$, generates a strongly continuous contraction semigroup on $\mathcal{L}_k \equiv \bar{\mathcal{D}}(A_k^d)$, that $\mathcal{L}_k$ is separating, and that for each probability measure $\nu \in \mathcal{P}(E_k)$ on the initial condition $X(0)$ there exists a solution of the martingale problem.*

The jump processes with bounded generators are added as follows. Assume $x \to Q(x, \cdot)$ of E into $\mathcal{P}(E)$ is continuous, and let $q(x)$ be nonnegative and $q(x) \in \bar{C}(E)$. Generators

$$A^j f(x) = q(x) \int (f(y) - f(x)) Q(x, dy), \tag{E.1}$$

define bounded linear operators on $B(E)$.

(i) Then the closure of $A^d \subset \bar{C}(E) \times \bar{C}(E)$

$$A^d = \left\{ \Big(\sum_{k=0}^{m} 1_{E_k} f_k \, , \sum_{k=0}^{m} 1_{E_k} A_k^d f_k \Big) \, , \, f_k \in \mathcal{D}(A_k^d), \forall m \geq 0 \right\}, \tag{E.2}$$

generates a strongly continuous contraction semigroup on $\mathcal{L} \equiv \bar{\mathcal{D}}(A^d)$ which is single-valued, and there exists a solution of the martingale problem for A^d.

(ii) Also, the closure of $A = A^d + A^j$ generates a strongly continuous contraction semigroup on $\mathcal{L} \equiv \bar{\mathcal{D}}(A)$, there exists a solution of the martingale problem for $A^d + A^j$ and each initial measure ν.

Proof For part (i) see the first half of problem 28, p. 266 of Ethier and Kurtz [463], with the single-valuedness following from it's dissipativity and Lemma 4.2, p. 21. For part (ii), A^j is bounded and linear on $B(E)$ and since it satisfies the positive maximum principle on all of $B(E)$ it is dissipative (Lemma 2.1, p. 165 [463]). From the same Lemma 4.2 of p. 21 it is single-valued and since the closure of A^d generates a single-valued strongly continuous contraction semigroup on $\mathcal{L}$ with $\mathcal{D}(A^d) \subset \mathcal{D}(A^j)$, then the closure of $A = A^d + A^j$ generates a strongly continuous contraction semigroup on $\mathcal{L}$ (Theorem 7.1, p. 37 [463]). That there exists a solution to the martingale problem for $A = A^d + A^j$ and ν follows from Thereom 10.2, p. 256 of [463], completing the proof. □

E.2 Irreducibility

Let chain $X_i = X(i\Delta)$ with $X(t)$ the Markov process satisfying the first part of Theorem 1. Then X_i is $\nu-$irreducible, i.e. there exists a probability measure ν on $\mathcal{X}$ with for any measurable set $\mathcal{A} \subset \mathcal{X}$ and $\nu(\mathcal{A}) > 0$, then $\forall x \in \mathcal{X}$, $Pr(X_i \in \mathcal{A}$ for some integer $i | X(0) = x) > 0$. In this case, $\nu = \mu$.

Proof To ease notation, define $\tilde{q}(x, \mathcal{A}) = q(x, \mathcal{A} \setminus x)$. The proof of irreducibility involves decomposing the transition probabilities for the jump-diffusion process into the probability the process enters the set $\mathcal{A}$ with exactly n jumps, denoted as $P^{(n)}$ and given by

$$P^{(0)}(t,x,v,\mathcal{A}) = E\{1_{\mathcal{A}}(X(v))e^{-\int_t^v q(X(m))dm} | X(t) = x\}\,, \tag{E.3}$$

$$\begin{aligned} P^{(n+1)}(t,x,v,\mathcal{A}) &= \int_t^v \int_{\mathcal{X}} P^{(n)}(m,z,v,\mathcal{A}) P^{(1)}(t,x,m,dz)dm, \quad n = 1,2,\ldots, \\ &= \int_t^v \int_{\mathcal{X}} P^{(n)}(m,z,v,\mathcal{A}) \int_{\mathcal{X}} P^{(0)}(t,x,m,dr)\tilde{q}(r,dz)\, dm. \end{aligned} \tag{E.4}$$

The diffusion within any given subspace implies irreducibility within each of the separate spaces $\mathcal{X}(k)$ according to the following lemma.

Lemma E.2 *For each $\mathcal{X}(k)$ and all measurable $\mathcal{A}(k) \subset \mathcal{X}(k)$ with $\mu(\mathcal{A}(k)) > 0$, then* $\forall v > t, \forall x \in \mathcal{X}(k), P(t,x,v,\mathcal{A}(k)) > 0$.

Proof of Lemma 2 above on irreducibility of the diffusions in each subspace: For each $\mathcal{X}(k)$ and all measurable $\mathcal{A}(k) \subset \mathcal{X}(k)$ with $\mu(\mathcal{A}(k)) > 0$, then $\forall v > t, \forall x \in \mathcal{X}(k)$, $P(t,x,v,\mathcal{A}(k)) > 0$.

For this associate with any pair $(x(k), \mathcal{A}(k))$ the closed and bounded set $\mathcal{B}(k) \subset \mathcal{X}(k)$ with the property that $x(k) \cup \mathcal{A}(k) \subset \mathcal{B}(k)$. Then define the new killed diffusion process with state space $\mathcal{B}(k) \cup \Delta$, with Δ the state the process enters if there is either a jump, or the process diffuses outside of $\mathcal{B}$. Specifically, let $X(v), v \geq t$ be the diffusion within subspace $\mathcal{X}(k)$ with initial condition $X(t) = x(k)$, and define the lifetime

$$\rho = \inf\left\{ s \geq t \;:\; \int_t^s q(X(m))\, dm \geq Y \text{ or } X(s) \notin \mathcal{B} \right\}, \tag{E.5}$$

with Y an independent exponentially distributed random variable with mean 1. The killed process $\hat{X}$ becomes

$$\hat{X}(v) = \begin{cases} X(v) & t \leq v < \rho \\ \Delta & v \geq \rho, \end{cases} \tag{E.6}$$

with the semigroup of $\hat{X}$ given by $\hat{T}_{t,v}f(x) = E\{f(\hat{X}(v)) | \hat{X}(t) = x\}$. Then the semigroup satisfies the backward equation for all $f \in C_c^\infty(\mathrm{Re}^{n(k)})$ (Karatzas and Shrieve [214], p. 368)

$$-\frac{\partial \hat{T}_{t,v}f(x(k))}{\partial t} = \sum_{i=1}^{n(k)} a_i(x(k)) \frac{\partial \hat{T}_{t,v}f(x(k))}{\partial x_i} + \frac{1}{2}\sum_{i=1}^{n(k)} b_{i,i} \frac{\partial^2 \hat{T}_{t,v}f(x(k))}{\partial x_i^2} - q(x(k))\hat{T}_{t,v}f(x(k)) \tag{E.7}$$

with $\lim_{t \uparrow v} \hat{T}_{t,v}f = f$. Then $\hat{T}_{t,v}f(x(k)) = \int_{\mathcal{B}(k)} \hat{g}(t,x;v,\xi) f(\xi) m(d\xi)$ for $\hat{g}(t,x;v,\xi)$ the fundamental solution of the Cauchy problem having continuous first and second partial derivatives with respect to both arguments x, ξ, and continuous first derivative with respect to t, v (see Karatzas and Shrieve [214, p. 368], or Friedman [480, p. 82] and Chapter 1).

Now complete the irreducibility argument as follows. Construct a positive sequence $\{f^r\} \subset C_c^\infty$ such that $f^r(\xi) \overset{L^1(m)}{\longrightarrow} 1_{\mathcal{A}}(\xi)$ (C_c^∞ is dense in $L^p(m)$), implying

$$|\hat{T}_{t,v} f^r(x) - \hat{P}(t,x,v,\mathcal{A})| \le \int_{\mathcal{B}} |f^r(\xi) - 1_{\mathcal{A}}(\xi)| \hat{g}(t,x;v,\xi) m(d\xi) \tag{E.8}$$

$$\le \sup_{\xi \in \mathcal{B}} \hat{g}(t,x;v,\xi) \int_{\mathcal{B}} |f^r(\xi) - 1_{\mathcal{A}}(\xi)| m(d\xi) \overset{L^1(m)}{\longrightarrow} 0. \tag{E.9}$$

Note, in the final step we used the fact that $\hat{g}(t,x;v,\xi)$ is continuous in ξ over the compact set $\mathcal{B}$, implying it has a supremum. This gives

$$\lim_{r\to\infty} \hat{T}_{t,v} f^r(x) = \hat{P}(t,x,v,\mathcal{A}) = \int_{\mathcal{A}} \hat{g}(t,x;v,\xi) m(d\xi) \tag{E.10}$$

which implies

$$P(t,x(k),v,\mathcal{A}(k)) \ge P(t,x(k),v,\mathcal{A}(k)) \ge \hat{P}^{(0)}(t,x(k),v,A(k)) \tag{E.11}$$

$$\ge \inf_{\xi \in \mathcal{A}(k)} \hat{g}(t,x(k);v,\xi) m(\mathcal{A}(k)) > 0, \tag{E.12}$$

where the last inequality follows from the the fact that the fundamental solution $\hat{g}(t,x;v,\xi) > 0$ for $v > t$ (see Friedman [480, p. 44]), the compactness of $\mathcal{A}$ and the continuity of the fundamental solution $\hat{g}$, and the fact that μ is absolutely continuous with respect to m implying $\mu(\mathcal{A}(k)) > 0$ gives $m(\mathcal{A}(k)) > 0$.

Proof of Lemma QED:

Now back to the main Theorem For any $\mathcal{A}$, $\mu(\mathcal{A}) > 0$ implies for some k_0 there exists a subset $\mathcal{A}(0) \subset \mathcal{X}$, $\mathcal{A}(0) \subset \mathcal{X}(k_0)$ with $\mu(\mathcal{A}(0)) > 0$. This follows from the countability of $\mathcal{N}$. Then, since $P(t,x,v,\mathcal{A}) \ge P(t,x,v,\mathcal{A}(0))$ we work only with the subset $\mathcal{A}(0) \subset \mathcal{A}$. Now we decompose the configuration space $\mathcal{X}$ according to the subsets of configurations which require $n = 0,1,2,\ldots$ jumps to enter $\mathcal{A}(0)$. Define

$$\mathcal{X}(n) = \{x'(k') : k_0 \in \mathcal{T}^n(k')\}, \quad n = 0,1,2,\ldots, \tag{E.13}$$

with $\mathcal{X} = \cup_{n=0}^{\infty} \mathcal{X}(n)$. That $\mathcal{X}$ can be decomposed according to the last statement follows from the connectedness assumption. Now the irreducibility is proven using an induction argument in which we actually prove a somewhat stronger lower bound for the transition probabilities. That is $\forall \mathcal{B}(n) \subset \mathcal{X}(n)$ compact, $\inf_{x \in \mathcal{B}(n)} P^{(n)}(t,x,v,\mathcal{A}(0)) > 0$.

First for $\mathcal{X}(0)$. From the previous Lemma, clearly

$$\forall x \in \mathcal{X}(0), \quad \Pr(X_i \in \mathcal{A} \text{ for some integer } i | X(0) = x) > \hat{P}(0,x,\Delta,\mathcal{A}(0)) > 0\,, \tag{E.14}$$

which implies $\inf_{x \in \mathcal{B}(0)} P^{(0)}(t,x,v,\mathcal{A}(0)) > 0$ for all $\mathcal{B}(0) \subset \mathcal{X}(0)$ compact since $\hat{P}$ is continuous in x.

Now $\forall x \in \mathcal{X}(1)$ and $\forall v > t$, we prove $P^{(1)}(t,x,v,\mathcal{A}(0)) > 0$ via the following set of inequalities:

$$\begin{aligned}
P^{(1)}(t,x,v,\mathcal{A}(0)) &= \int_t^v \int_{\mathcal{X}} P^{(0)}(m,z,v,\mathcal{A}(0)) \int_{\mathcal{X}(1)} P^{(0)}(t,x,m,dr)\tilde{q}(r,dz)\,dm \\
&> \int_t^v \int_{\mathcal{X}(0)} P^{(0)}(m,z,v,\mathcal{A}(0)) \int_{\mathcal{A}(1)} P^{(0)}(t,x,m,dr)\tilde{q}(r,dz)\,dm \\
&> \int_t^v \int_{\mathcal{B}(0)} P^{(0)}(m,z,v,\mathcal{A}(0)) \inf_{r\in\mathcal{A}(1)} \tilde{q}(r,dz)P^{(0)}(t,x,m,\mathcal{A}(1))\,dm \\
&> \inf_{r\in\mathcal{A}(1)} \tilde{q}(r,\mathcal{B}(0)) \int_{t+\delta}^{v-\delta} \inf_{z\in\mathcal{B}(0)} P^{(0)}(m,z,v,\mathcal{A}(0))P^{(0)}(t,x,m,\mathcal{A}(1))\,dm \\
&> (v-t+2\delta) \inf_{r\in\mathcal{A}(1)} \tilde{q}(r,\mathcal{B}(0)) \\
&\quad \times \inf_{z\in\mathcal{B}(0),m\in[t+\delta,v-\delta]} P^{(0)}(m,z,v,\mathcal{A}(0)) \inf_{m\in[t+\delta,v-\delta]} P^{(0)}(t,x,m,\mathcal{A}(1)).
\end{aligned} \tag{E.15}$$

The first inequality is for all measurable sets $\mathcal{A}(1) \subset \mathcal{X}(1)$, the second inequality for all measurable sets $\mathcal{B}(0) \subset \mathcal{X}(0)$, and the third for all $\delta \in (0,(v-t)/2)$. That the lower bound is greater than 0 follows from the fact that we can choose $\mathcal{A}(1)$ and $\mathcal{B}(0)$ both compact with the property that $\inf_{r\in\mathcal{A}(1)} \tilde{q}(r,\mathcal{B}(0)) > 0$, and from the previous Lemma $P^{(0)}(m,z,v,\mathcal{A}(0)) \geq \hat{P}^{(0)}(m,z,v,\mathcal{A}(0)) > 0$ with the inf positive because of continuity of $\hat{P}$ with respect to both m,z. The identical argument for the final inf term implies $P^{(1)}(t,x,v,\mathcal{A}(0)) > 0, \forall x \in \mathcal{X}(1)$, and by the same continuity $\inf_{x\in\mathcal{B}(1)} P^{(1)}(t,x,v,\mathcal{A}(0)) > 0$ for all compact $\mathcal{B}(1) \subset \mathcal{X}(1)$.

Now assume that for all compact $\mathcal{B}(n) \subset \mathcal{X}(n), \forall v > t, \inf_{x\in\mathcal{B}(n)} P^{(n)}(t,x,v,\mathcal{A}(0)) > 0$. Then we have $\forall x \in \mathcal{X}(n+1)$,

$$\begin{aligned}
P^{(n+1)}(t,x,v,\mathcal{A}(0)) &= \int_t^v \int_{\mathcal{X}} P^{(n)}(m,z,v,\mathcal{A}(0)) \int_{\mathcal{X}(n+1)} P^{(0)}(t,x,m,dr)\tilde{q}(r,dz)\,dm \\
&> \int_t^v \int_{\mathcal{X}(n)} P^{(n)}(m,z,v,\mathcal{A}(0)) \int_{\mathcal{A}(n+1)} P^{(0)}(t,x,m,dr)\tilde{q}(r,dz)\,dm \\
&> \int_t^v \int_{\mathcal{B}(n)} P^{(n)}(m,z,v,\mathcal{A}(0)) \\
&\quad \times \inf_{r\in\mathcal{A}(n+1)} \tilde{q}(r,dz)P^{(0)}(t,x,m,\mathcal{A}(n+1))\,dm \\
&> \inf_{r\in\mathcal{A}(n+1)} \tilde{q}(r,\mathcal{B}(n)) \\
&\quad \times \int_{t+\delta}^{v-\delta} \inf_{z\in\mathcal{B}(n)} P^{(n)}(m,z,v,\mathcal{A}(0))P^{(0)}(t,x,m,\mathcal{A}(n+1))\,dm \\
&> (v-t+2\delta) \inf_{r\in\mathcal{A}(n+1)} \tilde{q}(r,\mathcal{B}(n)) \\
&\quad \times \inf_{z\in\mathcal{B}(n),m\in[t+\delta,v-\delta]} P^{(n)}(m,z,v,\mathcal{A}(0)) \\
&\quad \times \inf_{m\in[t+\delta,v-\delta]} P^{(0)}(t,x,m,\mathcal{A}(n+1)) \;,
\end{aligned}$$

which from the induction gives the two inf terms as positive, and again by the continuity and lower bounding of $\hat{P}$ gives for all compact $\mathcal{B}(n+1) \subset \mathcal{X}(n+1)$, $\forall v > t$, $\inf_{x \in \mathcal{B}(n+1)} P^{(n+1)}(t, x, v, \mathcal{A}(0)) > 0$.

Since $\forall x \in \cup_{n=0}^{\infty} \mathcal{X}(n)$, $Pr(X(n) \in \mathcal{A}$ for some integer $n | X(0) = x) > 0$, the μ-irreducibility Lemma is proved. $\square$

F HOMEWORK AND SELECTED SOLUTIONS

F.1 Bayes Posterior Chapter 2

Problem F.1.0 Consider Shannon's *source/channel* model. Let $\mathcal{X} = \{1, \ldots, m\}$ be an enumeration of M distinct states, and $X \in \mathcal{X}$ with prior distribution $\pi(i), i = 1, \ldots, m$. Associated with each state, i, is a unique vector of data, $s_i = (s_{i,1}, \ldots, s_{i,n})$. The output Y is observed through a *channel* which produces Gaussian additive noise $N = (N_1, \ldots, N_n)$, with $E(N) = 0, E(NN^T) = \sigma^2 I$. Thus, if s_i is the transmitted data, the observer sees $s_i + N$. Determine the densities $p(y|X = i)$, and show that the minimum risk, equal cost decision, $i^\star$, for observation y is given by

$$i^\star = \underset{i=1,\ldots,m}{\arg\min} \ \|y - s_i\|^2 - 2\sigma^2 \log \pi(i). \tag{F.1}$$

Proof For each i we have

$$p(y|X = i) = p(s_i + N) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{1}{2\sigma^2}\|y - s_i\|^2\right).$$

By Theorem 2.15, the optimum decision for observation y is given by

$$\begin{aligned} i^\star = \underset{i}{\arg\max} \ \pi(i) p(y|X = i) &= \underset{i}{\arg\max} \ \pi(i) \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{1}{2\sigma^2}\|y - s_i\|^2\right) \\ &= \underset{i}{\arg\max} \ \exp\left(\log \pi(i) - \frac{1}{2\sigma^2}\|y - s_i\|^2\right) \\ &= \underset{i}{\arg\min} \ \frac{1}{2\sigma^2}\|y - s_i\|^2 - \log \pi \\ &= \underset{i}{\arg\min} \ \|y - s_i\|^2 - 2\sigma^2 \log \pi. \end{aligned}$$

□

Problem F.1.1 Let $\mathcal{X} = \{1, 2, 3\}$ and $X \in \mathcal{X}$ with priors $\pi(1) = 1/6$, $\pi(2) = 2/6$ and $\pi(3) = 3/6$. Let $p(y|X = x) = (x + 1)\exp(-(x + 1)y)u(y)$ where $u(y)$ is the unit step function, that is $u(y) = 1$ for $y \geq 0$ and $u(y) = 0$ otherwise. Find the optimum partitioning of the observation space for equal costs of misclassification.

Proof Comparing likelihood ratios we see that for each $k = 1, 2, 3$

$$\frac{p(y|X = 1)}{p(y|X = k)} > \frac{\pi(k)}{\pi(1)},$$

when $y > \log 3$. Thus by Theorem 2.15, $D_1^\star = \{y \in \mathcal{Y} : y > \log 3\}$. Similarly, we find $D_2^\star = \{y \in \mathcal{Y} : \log 2 < y \leq \log 3\}$ and $D_3^\star = \{y \in \mathcal{Y} : 0 \leq y \leq \log 2\}$. □

Problem F.1.2 Let X be a zero-mean Gaussian random variable with unknown variance θ. We observe $Y = X + N$, where N is a zero-mean Gaussian noise with variance σ^2. Assuming X and N are independent,

(i) find the MLE of θ given the observation $Y = y$ and

(ii) apply the EM algorithm to this problem. Show analytically that the stationary point $\theta^\star$ coincides with the MLE.

Problem F.1.3 (Adding detector efficiency) Since CCD detectors are imperfect with varying pixel quantum efficiencies, it has been modeled extensively as in [30, 481–483] assuming each photon is independently deleted with probability $d_j, j \in Y$ and detected with probability $s_j = 1 - d_j$.

If the deletions are conditionally independent given the emissions with rate $d(\cdot)$ then the Poisson log-likelihood of the measurements becomes

$$\log g(Y|\lambda) = -\sum_j s_j \sum_i p(j|i)\lambda_i + \sum_j Y_j \log \left(s_j \sum_i p(j|i)\lambda_i \right) . \tag{F.2}$$

As shown in [30], the EM algorithm can be used for maximizing the likelihood of F.2. Calculating the sequence of iterates corresponding to the maximization gives

$$\lambda^{\text{new}} = \arg\max_\lambda - \sum_i \lambda_i + \sum_i \lambda_i^{\text{old}} \sum_y Y_j \left((1-s_j)p(j|i) + \frac{p(j|i)}{\sum_i p(j|i)\lambda_i^{\text{old}}} \right) \log \lambda_i .$$

Demonstrate that stable points of the iteration satisfy the necessary maximizer conditions for the maximizer of the likelihood of Eqn. F.2.

F.2 Probability Models on Directed Acyclic Graphs Chapter 3

Problem F.2.4 Given a two-state Markov chain corresponding to coin flipping, $X_i, i = 1, 2, \ldots \in \{H, T\}$. Let the two state transition matrix be $\begin{pmatrix} p & 1-p \\ 1-q & q \end{pmatrix}$. Assume $P(X_1 = H) = 1/(1-p+q), P(X_1 = T) = (-p+q)/(1-p+q)$, then what is $P(X_n = H)$?

Problem F.2.5 For the trees in Section 3.4, calculate the neighborhood structures of the random field.

Problem F.2.6 Given generating function $f(s)$ for the 1-type, prove

$$(f_n(s))^j = \sum_{k \in I^+} \Pr\{Z_n = k | Z_0 = 1\} s^k . \tag{F.3}$$

For the $V-$type, prove

$$(f_n^{(1)}(s))^{j_1} (f_n^{(2)}(s))^{j_2}, \ldots, (f_n^{(V)}(s))^{j_V} = \sum_{k \in I^{+V}} \Pr\{Z_n = k | Z_0 = j\} s_{\mathbf{1}}^{k_1}, \ldots, s_V^{k_V} . \tag{F.4}$$

Problem F.2.7 For the single-type processes, prove the expected value of the mean and variance of Z_n given $Z_0 = 1$ is given by

$$E\{Z_n | Z_0 = 1\} = m^n \tag{F.5}$$

$$E\{Z_n^2 | Z_0 = 1\} - (E\{Z_n | Z_0 = 1\})^2 = \begin{matrix} (\sigma^2 m^n (m^n - 1))/(m^2 - m), & m \neq 1; \\ n\sigma^2, & m = 1 . \end{matrix} \tag{F.6}$$

Problem F.2.8 Prove the geometric properties of the expectation of the branching matrix, Eqn. 3.37.

Problem F.2.9 Calculate the largest eigenvalue for the mean matrix M in example 3.19. As proved in Harris, Theorem 7.1 p. 42, that the extinction probability vector $\gamma = \lim_{n\to\infty} f_n(\tilde{\gamma})$, for any $\tilde{\gamma} \in (0,1)^d$. Write a computer program to calculate the extinction probability vector $\gamma = [\gamma^{(1)}, \gamma^{(2)}, \gamma^{(3)}]$.

$$0.1262 \le \gamma^{(1)} = \gamma^{(3)} \ge 0.125, \tag{F.7}$$

$$0.0291 \ge \gamma^{(2)} \ge 0.0290. \tag{F.8}$$

F.3 Entropies and Combinatorics of Patterns Chapter 3

Problem F.3.10 (Maximum Entropy Distribution) Mark proves that the maximum entropy distribution combining random branching processes with bigram/trigram relationships has the neighborhood structure of the hierarchical DAG model for languages described in Section 3.11.

Assume we are given a prior distribution on word strings $\pi(W_{1,n})$ which corresponds to a random branching process or stochastic context-free language. The moment constraints come in the form of bigram relative frequencies:

$$h_{\sigma,\sigma'}(W_{1,n}) = \sum_{i=1}^{n-1} 1_{\sigma,\sigma'}(w_i, w_{i+1}), \tag{F.9}$$

where $\sigma, \sigma' \in \{\text{Words}\}$.

Prove the following theorem from Mark stating the maximum entropy distribution.

Theorem F.1 (Mark) *Let $c = W_{1,n}$ and $f(c) = \pi(W_{1,n})$. The distribution maximizing the generalized entropy $-\sum p(c)\log(p(c))/(f(c))$ subject to the constraints $\{E[h_{\sigma,\sigma'}(W_{1,n})] = H_{\sigma,\sigma'}, \sigma, \sigma'$ over all word pairs becomes*

$$\begin{aligned}\Pr(W_{1,n}) &= Z^{-1}\exp\left(\sum_{\sigma\in V_T}\sum_{\sigma'\in V_T}\alpha_{\sigma,\sigma'}h_{\sigma,\sigma'}(W_{1,n})\right)\pi(W_{1,n})\\ &= Z^{-1}\sum_{T\in\text{Parses}(W_{1,n})}\exp\left(\sum_{k=1}^{n-1}\alpha_{w_k,w_{k+1}} + \sum_{i=1}^{n_T}\log P_{r_i^T}\right),\end{aligned} \tag{F.10}$$

where Z is the normalizing constant and n_T is the number of rules in tree T and r_i^T is the ith rule in tree T.

Problem F.3.11 (continued) Demonstrate the that neighborhood structure of the maximum entropy distribution is identical to the layered hierarchical DAG model depicted in Figure 3.12.

F.4 Markov Random Fields on Directed Graphs Chapter 4

Problem F.4.12 Show that the 2D Ising model of Section 4.10 is not a Markov process!

Problem F.4.13 Simulate some MA Cauchy field in 2D with M being an averaging operator over a small neighborhood. Simulate some ARMA Cauchy field with L being an elasticity operator.

F.5 Gaussian Random Fields on Undirected Graphs 5

Problem F.5.14 (Periodic Gaussian process) Examine the well-known example of an n-dimensional Gaussian vector. Given are sites $D = \{1, 2, \ldots, n\}$; assume $x = [x_1, \ldots, x_n]$ forms an n-dimensional Gaussian vector with density

$$p(x) = \frac{1}{Z} e^{-xK^{-1}x} . \tag{F.11}$$

Construct the $n \times n$ covariance K so that x is an n-dimensional Gaussian random field with cyclic neighborhoods $N_i = \{i-p, \ldots, i-1, i+1, \ldots, i+p\}$, for $p = 1$, and then more generally for $p = 2, \ldots$.

Problem F.5.15 (Matrix Operators and Karhunen-Loeve Expansion of Gaussian Fields) Let us construct the covariance K_X of our Gaussian random fields $\{X_i(\omega), i \in D \subset \mathbb{Z}^d\}$ corresponding to the solution of the matrix equations

$$L\, X_i = W_i,$$

where $L : \mathbb{R}^n \to \mathbb{R}^n$ is an $n \times n$ symmetric ($L = L^*$), real invertible matrix operator, and $W \in \mathbb{R}^n$ is an n-dimensional multivariate Gaussian vector, a collection of i.i.d. unit variance Gaussians. (Think of L as arising from a local (bounded support) self-adjoint difference operator with appropriate boundary conditions as in section 5.1.)

Use the n-eigenelements $\{\phi(k), \lambda_k, k = 1, \ldots, n\}$, $L\phi_k = \lambda_k \phi_k$, and their orthogonality (symmetric operator) to establish the following.

1. Show

$$Z_k(\omega) = \sum_{i=1}^{n} X_i(\omega)\phi_i(k) , \quad k = 1, \ldots, n$$

are independent, Gaussian random variables variance $1/\lambda_k^2$.

2. Show the covariance is given by

$$K_X(i,j) = \sum_{k=1}^{n} \frac{1}{\lambda_k^2} \phi_k(i)\phi_k(j) \tag{F.12}$$

$$= (LL)^{-1}, \tag{F.13}$$

and calculate the determinant of K_X^{-1} in terms of the eigenvalues $\{\lambda_k\}$.
Define the eigenelements $\{\phi_k, \lambda_k\}$ satisfying the operator equation

$$L\phi_i(k) = \lambda_k \phi_i(k), \quad i = 1, \ldots, n . \tag{F.14}$$

Then

$$EZ_k(\omega)Z_j(\omega) = \frac{\langle W, \phi_k \rangle}{\lambda_k} \frac{\langle W, \phi_j \rangle}{\lambda_j} \tag{F.15}$$

$$= \frac{\phi_k^* K_W \phi_j}{\lambda_k \lambda_j} = \frac{1}{\lambda_k^2} \delta(k-j) . \tag{F.16}$$

Last equality follows from orthogonality of eigenfunctions associated with symmetric matrix L.

Problem F.5.16 Using Matlab demonstrate that the asymptotic eigenvalues have the limit shown in Example 5.20:

$$\lim_{n\to\infty} \frac{2\sum_{i=1}^{n} (a+2-2\cos(2\pi i/n))}{n} = \frac{1}{2\pi}\int_0^{2\pi} \log\left(a+2-2\cos\omega\right)\, d\omega\,. \quad \text{(F.17)}$$

Problem F.5.17 A stationary stochastic process has covariances $K_{s-t} = \rho^{|s-t|}$, $0 \le \rho \le 1$. Derive its spectral density $S(\omega)$. Let the covariance matrix

$$R^{(n)} = \left(\rho^{|s-y|}; s,t = 1,2,\ldots,n\right)$$

have the eigenvalues $\{\lambda_1, \lambda_2, \ldots \lambda_n\}$. For $n = 20, 40, 60, \ldots$ compute the eigenvalues numerically, plot the histogram of their distribution as well as the histogram of the set of values $S(\pi/n), S(2\pi/n), \ldots, S(\pi)$. Discuss what you see.

Problem F.5.18 Consider a polynomial $P(\cdot)$:

$$P(z) = \sum_{k=0}^{m} a_k z^k$$

with no zeros on the boundary $|z| = 1$ on the unit circle in the complex plane. Show that there exists a stationary solution to the stochastic difference equation

$$\sum_{k=0}^{m} a_k X_{i+k} = W_i; \quad -\infty < i < \infty,$$

where $W_i; -\infty < i < \infty$ is an orthogonal sequence of random variables with

$$EW_i \equiv 0; EW_iW_j = \delta[j-i]\,.$$

Problem F.5.19 (continuation) What is the spectral density $S(\lambda)$ of the stochastic process, X_i, expressed as a Fourier integral?

Problem F.5.20 (continuation) Let the $n \times n$ matrix,

$$R^{(n)} = (r_{t-s}; -1 \le t,s \le n) = (EX_tX_s; \le t,s \le n),$$

be the covariance matrix of the X-process and denote the eigenvalues of $R^{(n)}$ by $\{\lambda_\nu^{(n)}; 1 \le \nu \le n\}$. Derive an anlytical expression in terms of the spectral density for the asymptotics

$$lim_{n\to\infty} \sum_{\nu=1}^{n} g(g(\lambda)),$$

where $g(\cdot)$ is a continuous real valued function.

Problem F.5.21 (continuation) Choose some values for the coefficients a_k and write a Matlab program that simulates the stochastic process $X_i; 1 \le i \le T$ conditioned by the initial condition $X_t = 0; t = 1,2,\ldots,m$ and plot the sample functions. Study their behavior, especially when the polynomial has a zero close to the boundary of the unit circle in the complex plane. You may choose m as a small positive integer.

F.6 The Canonical Representations of General Pattern Theory 6

Problem F.6.22 (Ising model) Regular lattice examples corresponding to magnetic domains are a natural. Consider the *Ising model* with $\mathcal{G} = \{1, -1\}$, $\Sigma = \text{TORUS}(L, L)$ meaning a 2D discrete torus with vertices $(x, y) \in \{0, 1, \dots, L-1\}^2$ and where site $x = L$ coincides with site $x = 0$ and similarly for the y coordinate. Introduce an acceptor matrix

$$A = \begin{pmatrix} 1 & b \\ b & c \end{pmatrix}$$

with positive constants b, c. What is the intuitive meaning of the constants in the acceptor matrix?

Problem F.6.23 With $Q(dg)$ a probability defined by the density $q(g) = A(x, y)$ with the distribution restricted to $\mathcal{C}_{\mathcal{R}}$. Prove this is the same probability distribution given by the standard Markov chain factorization.

Problem F.6.24 Let $\mathcal{G} = \mathcal{B} = \mathbb{R}$, arity 2, with bond values $\beta_j(g) = g$, $\sigma = \text{LINEAR}$, and potential $U(\beta_1, \beta_2) = (\beta_1 - \beta_2)^2$ so that

$$A_0(\beta_1, \beta_2) = e^{-(\beta_1 - \beta_2)^2}$$

with the weight function

$$Q(\beta) = e^{-\beta^2/v}.$$

Then P will be a Gaussian probability matrix with some covariance matrix; calculate it.

Problem F.6.25 Consider the Ising model on the $l \times l$ lattice

$$\mathcal{L} = \{(x, y) : x, y = 1, 2, \dots, l\}$$

with the acceptor function (matrix)

$$A = \begin{pmatrix} 1 & b \\ b & c \end{pmatrix}.$$

Use Dirichlet conditions, which means that we assume that $I(0, y) = I(l+1, y) = I(x, 0) = I(x, l+1) = 0$. Write a Matlab program using stochastic relaxation that synthesizes and displays the resulting image $I(x, y)$. Experiment with values of b, c, in particular for $c = 1$ and when b takes various positive values. When are the resulting images cohesive so that the 1 and 0 sets in I hang together?

Problem F.6.26 (continuation) Use an output I from the previous problem and introduce a stochastic deformation mechanism acting on each separate pixel of I independently

$$\mathcal{D} : I^{\mathcal{D}}(x, y) = I(x, y) \textit{ with probability } 1 - \epsilon,$$

$$\mathcal{D} : I^{\mathcal{D}}(x, y) = 1 - I(x, y) \textit{ with probability } \epsilon.$$

Write a program that simulates the deformed image, $I^{\mathcal{D}}$, when the pure image, I, (obtained in problem 4.7.16) is given and ϵ is small, perhaps $\epsilon = .1$. Write an expression for the conditional probability of I given $I^{\mathcal{D}}$

$$p[I(x,y)|I^{\mathcal{D}}(\cdot,\cdot)]$$

and derive an inference algorithm, also employing stochastic relaxation, for restoring the pure image I when the deformed $I^{\mathcal{D}}$ is known, called the restored image $I_t^*(x,y)$ after t iterations and display the iterations. During the iterations of the stochastic relaxation $t = 1,2,3,\ldots$ compute the Hamming distance

$$d(I,I^*) = \sum_{x,y=1}^{l} \mathbf{1}_{\{I(x,y)=I^*(x,y)\}}(x,y),$$

where $\mathbf{1}_{\text{relation}}$ is the indicator function of the *relation*.

Problem F.6.27 Prove that the law of composition, definition 7.2, makes the product $\mathcal{S}_1 \cdots \times \mathcal{S}_n$ a group.

Solution: The law of composition is well defined. Suppose s and s' are defined as in the above definition. Then $s \circ s' = (s_1 \circ_1 s_1', \ldots, s_n \circ_n s_n') \in \mathcal{S}$ since $\forall i \in \{1,\ldots,n\} s_i \circ s_i') \in \mathcal{S}_i$ so $s \in \mathcal{S}$. It is also associative, since $\forall i \in \{1,\ldots,n\}$ $\mathcal{S}_i$ is associative. Suppose $t = (t_1,\ldots,t_n) \in \mathcal{S}$. Then, $(s \circ s') \circ t = ((s_1 \circ_1 s_1') \circ_1 t_1, \ldots, (s_n \circ_n s_n\prime) \circ_n t_n) = (s_1 \circ_1 (s_1' \circ_1 t_1), \ldots, s_n \circ_n (s_n' \circ_n t_n)) = s \circ (s' \circ t)$. $(\mathcal{S}, \circ)$ has identity element $e = (e_1, \ldots, e_n)$ where e_i denotes the identity element of $(\mathcal{S}_i, \circ_i)$; s^{-1} is given by $(s_1^{-1}, \ldots, s_n^{-1})$.

Problem F.6.28 Prove that the special linear group

$$\mathbf{SL}(n) = \{A \in \mathbf{GL}(n) : \det A = 1\} \tag{F.18}$$

is a group.

Problem F.6.29 Prove that the matrices

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \tag{F.19}$$

are elements of the coset

$$\left[\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right]_{\mathbf{SO}(3)} . \tag{F.20}$$

Solution: Let the three matrices in the problem be denoted a,b,c for convenience. Then $a = pc$ and $b = qc$, where $p, q \in \mathbf{SO}(3)$ are given, respectively, by

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \tag{F.21}$$

Problem F.6.30 Prove left and right cosets agree for normal subgroups, $aH = Ha$.

Solution: By definition of normal subgroup, $aHa^{-1} = H$, so $aHa^{-1}a = aH = Ha$.

Problem F.6.31 What is the quotient group for

$$\begin{pmatrix} u_1 & u_2 \\ -u_2 & u_1 \end{pmatrix} / \begin{pmatrix} \rho & 0 \\ 0 & \rho \end{pmatrix} . \tag{F.22}$$

Notice the severe abuse of notation, where we have represented the groups by elements.

Solution: Write the scale-rotation matrix as the product of a rotation matrix and a scale matrix. The rotation angle is given by $\arctan(u_2/u_1)$; the scale is given by $\det(U)$. Then, it is clear that each rotation matrix (i.e. each element of $\mathbf{SO}(2)$ represents a coset. Thus, $\mathbf{SO}(2)$ is the quotient group.

Problem F.6.32 Prove that $\phi\colon \mathbf{A}(2) \to \mathbf{A}(2)$ according to $\phi(A, a) = A$ is a homomorphism.

Solution: That $\phi : \mathbf{A}(2) \to \mathbf{A}(2)$ according to $\phi(A, a) = A$ is a homomorphism follows from the fact that

$$\phi((A, a) \circ (B, b)) = \phi((AB, Ab + a)) \tag{F.23}$$

$$= AB = \phi(A, a) \circ \phi(B, b) \ . \tag{F.24}$$

The kernel $K = \{(I, a), a \in \mathbb{R}^2\}$.

Problem F.6.33 For symmetry around the *sagittal plane*, choose the convention that the sagittal plane through the center of the longitudinal cerebral fissure is the set of all points

$$\text{PLANE} - \text{SAGITTAL} = \{x = (x_1, x_2, x_3) \ : \ x_1 \ = 0\} \ .$$

Of particular interest are the lines in $\mathbb{R}^2$ and planes in $\mathbb{R}^3$ of symmetry, points that are invariant to the Euclidean motion:

$$x = Ox + a. \tag{F.25}$$

In the plane $\mathbb{R}^2$, write the equation for the symmetry line using the fact that one of the eigenvalues of $I - O$ is zero.

Solution: implying O has eigenvector, ψ, with $O\psi = \psi$ with eigenvalue 1. The *symmetry line* is the subset $\tilde{X} \subset \mathbb{R}^2$ defined according to

$$\tilde{X} = \{x \in \mathbb{R}^2 : x = \tilde{x}_0 + \alpha\psi, \alpha \in \mathbb{R}, (I - O)\tilde{x}_0 = a\} \ . \tag{F.26}$$

Problem F.6.34 In the subsection 7.29, assume as the template generators a singleton $\mathcal{G}^0 = \{\begin{pmatrix} 1 \\ 1 \end{pmatrix}\}$. What is the determinant of the Jacobian of the similarity transformation.

Problem F.6.35 As pointed out in the section 7.29, the piecewise linear simple contours in the plane are specified via the n-vertices according to $v_m = \sum_{k=1}^{m} g_k \in \mathbb{R}^2$, , $m \geq 0$. These vertices are themselves of interest. Given the density $p(c)m(dc)$ on the configuration space $\mathbb{R}^n$ of Eqn. 7.46, calculate the probability distribution on the vertices $\tilde{p}(v_1, v_2, \ldots, v_n)m(dv)$.

Problem F.6.36 Prove the projection formula Eqns. 7.57 and 7.58 under the linear manifold condition, Eqn. 7.56.

Problem F.6.37 Choose the general linear group as the similarity group, $\mathcal{S}_0 = \mathbf{GL}(3)$. Then, for any two generators g_1 and g_2 in the subsection 7.33, none of which is on a great circle, show that the equation $g_1 = sg_2$ has a unique solution in $s \in \mathcal{S}_0$.

Problem F.6.38 (*UNSOLVED PROBLEM) For the above subsection 7.33, find the $15/2n - 6$ independent constraints on the group elements.

Problem F.6.39 Define set(c) = *inside of the curve* and shows that the equivalence relation, R, defined by

$$c \sim c' \Leftrightarrow \text{set}(c) = \text{set}(c') \tag{F.27}$$

satisfies the conditions to be an identification rule.

F.7 Second Order and Gaussian Fields 9

Problem F.7.40 Write the general expression for the sequence $X(\omega, n)$ in the above example 9.4, and show that the $X(n)$ are $\mathcal{A}$-measurable.

$$\{\omega : X(\omega, n) = 1\} = \left(\bigcup_{i=0}^{2^n} \left[\frac{2i+1}{2^{n+1}}, \frac{2i+2}{2^{n+1}} \right) \right), \tag{F.28}$$

$$\{\omega : X(\omega, n) = 0\} = \left(\bigcup_{i=0}^{2^n} \left[\frac{2i}{2^{n+1}}, \frac{2i+1}{2^{n+1}} \right) \right). \tag{F.29}$$

which are both $\mathcal{A}$ measurable as finite unions of measurable sets are measurable.

Problem F.7.41 Show that q.m. continuity of $\{X(t), t \in T\}$ implies separability of the process.

Problem F.7.42 Given $dx(t) = ax(t)dt + dw(t)$, $w(t)$ a Wiener process.

1. Calculate the adjoint of $A = (d/dt) - a$. Is it self-adjoint?
2. Calculate the covariance $K[t, s]$ on $[0, 1]$, and eigenfunctions and eigenvalues for the Karhunen-Loeve representation.

Problem F.7.43 (Lebesgue Integral and Joint-Measurability) Define the Lebesgue integral

$$\int_{T \times \Omega} X(\omega, t) d(m \times P), \tag{F.30}$$

and establish that it involves measurability of sets of the form 9.15.

Proof goes something like the following. Define the product measure $\mu = m \times P$, and then define simple functions

$$f_n(\omega, t) = \sum_{k=0}^{2^{2n}} \tilde{f}_k 1_{A_k}(\omega, t), \tag{F.31}$$

where

$$A_k = \{(\omega, t) \in T \times \Omega : k2^{-n} \leq f(\omega, t) < (k+1)2^{-n}\}, \quad k = 0, 1, \ldots, 2^{2n}. \tag{F.32}$$

But this is the intersection of two measurable sets in the product topology of $\mathcal{L} \times \mathcal{A}$; hence $\mu(A_k)$ is well defined, and the integral of the simple function becomes $\sum_{k=0}^{2^{2n}} \tilde{f}_k \mu(A_k)$.

Problem F.7.44 Let $\{X(\omega, n), n \in \mathbb{Z}\}$ be a second order process on $(\Omega, \mathcal{A}, P)$ with

$$\Omega = l^2 \subset \mathbb{R}^\infty \text{ the set of square} - \text{summable sequences},$$
$$\mathcal{A} = \mathcal{B}(\mathbb{R}^\infty) \text{ the Borel } \sigma\text{−algebra on } \mathbb{R}^\infty,$$

and P the infinite dimensional Gaussian product measure with n-dimensional densities on cylinders given by

$$p_n(X_1, \ldots, X_n) = \prod_{k=1}^{n} p(X_k). \tag{F.33}$$

Define the stochastic process

$$X(\omega, t) \overset{\text{q.m.}}{=} \lim_{n\to\infty} \sum_{k=0}^{n} X_k(\omega)(e^{2\pi ikt} + e^{-2\pi ikt}), \quad t \in [0, 1]. \tag{F.34}$$

1. Calculate conditions for $\{X(\omega, t), t \in [0, 1]\}$ to be a SOP on $(\Omega, \mathcal{A}, P)$.
2. Calculate conditions for $X(\omega, t)$ to be a q.m. continuous Gaussian process.

Problem F.7.45 (Example 9.10) Consider the Hilbert space of random variables H_X generated by the random process

$$X(t) = X_0(\omega)\cos 2\pi t - X_1(\omega)\sin 2\pi t, \quad t \in [0, 1]. \tag{F.35}$$

Prove that H_X is 2D spanned by $\{X_0(\omega), X_1(\omega)\}$.

Proof Let $Y \in H_X$. By definition there exists a sequence $Y_N = \sum_{n=1}^{N} \alpha_n X(\omega, t_n)$ such that $Y_N \to Y$ in q.m. Define $a_N = \sum_{n=1}^{N} \alpha_n \cos 2\pi t_n$ and $b_N = \sum_{n=1}^{N} \alpha_n \sin 2\pi t_n$. Since Y_N converges in q.m., it is also Cauchy in q.m. Furthermore,

$$E|Y_N - Y_m|^2 = E|X_0(\omega)(a_N - a_m) - X_1(\omega)(b_N - b_M)|^2 \tag{F.36}$$

$$= |a_N - a_m^2| + |b_N - b_M|^2, \tag{F.37}$$

where the last equality follows from $E|X_0|^2 = E|X_1|^2 = 1$ and $EX_0X_1^* = EX_0^*X_1 = 0$. Thus $\{a_N\}$ and $\{b_N\}$ are Cauchy in $\mathbb{C}$, hency they converge to complex numbers, say a and b, respectively. We then have

$$E|Y_N - (aX_0(\omega) + bX_1(\omega))|^2 = E|X_0(\omega)(a_N - a) + X_1(\omega)(b_N - b)|^2 \tag{F.38}$$

$$= |a_N - a|^2 + |b_N - b|^2, \tag{F.39}$$

which converges to zero as $N \to \infty$. □

Problem F.7.46 Proof of vector case of Mercer's Theorem 9.24. Hint: introduce the m-vector norm and Frobenius norm for matrix approximation:

$$\langle f(\cdot), g(\cdot)\rangle = \int f^* g\ ,\ \langle A(\cdot), B(\cdot)\rangle = \int \mathrm{tr} A(t)^* B(t)\, dt\ . \tag{F.40}$$

Proof Let $L_m^2(T)$ denote the space of finite norm m-vector valued functions and $L_{m\times m}^2(T)$ denote the space of finite norm matrix valued functions, with inner products

$$\langle f(\cdot), g(\cdot)\rangle = \int \langle f(s), g(s)\rangle_{\mathbb{R}^m}\, ds = \int f^*(s) g(s)\, ds\ .$$

Linear algebra gives

$$\|A\|^2 = \int \mathrm{tr}(A^*(t)A(t))\, dt = \int \sum_{j=1}^{m} \|A_j(t)\|_{\mathbb{R}^m}^2\, dt = \sum_{j=1}^{m} \int \|A_j(t)\|_{\mathbb{R}^m}^2\, dt = \sum_{j=1}^{m} ||A_j||, \tag{F.41}$$

where $A_j(\cdot)$ is the m-vector valued function corresponding to the jth row of $A(\cdot)$. Then K is self-adjoint by the same argument above, where every occurence of $*$ is replaced by $*$. Similarly, for each $j = 1, \ldots, n$ we apply the Fourier series argument to $A_j(\cdot, t)$, obtaining convergence in $L^2_m(T)$ for each column, and hence convergence in $L^2_{m\times m}(T)$ by (F.41). Again, the same argument as in the scalar case extends this to convergence in $L^2_{m\times m}(T \times T)$. The remaining arguments remain the same, replacing $*$ with $*$ with the appropriate use of the norms. □

Problem F.7.47 Proof of vector case of Karhunen-Loeve Theorem 9.25.

Proof Let us show that the error in the finite-term approximation of the random field using the vector expansion equation tends to zero when the number of terms N tends to infinity. Denote the finite sum $X_N(x) = \sum_{n=1}^N X_n \phi_n(x)$; the jth component of $X_N(x)$ is given by $X_{jN}(x) = \sum_{n=1}^N X_n \phi_{jn}(x)$. The error in representing the process with this finite-term approximation is

$$\|X(x) - X_N(x)\|^2 = EX^*(x)X(x) - EX^*(x)X_N(x) - EX_N^*(x)X(x) + EX_N^* X_N(x), \tag{F.42}$$

$$\text{with } EX^*(x)X(x) = E(X_1^*(x)\ X_2^*(x)\ \ldots X_m^*(x)) \begin{pmatrix} X_1(x) \\ X_2(x) \\ \vdots \\ X_m(x) \end{pmatrix}$$

$$= \sum_{i=1}^m EX_i^*(x)X_i(x) = \mathrm{tr}K(x,x), \tag{F.43}$$

$$EX^*(x)X_N(x) = \sum_{i=1}^m EX_i^*(x)X_{iN}(x) = \sum_{i=1}^m E(X_i^*(x) \sum_{n=1}^N X_n\phi_{in}(x))$$

$$= \sum_{i=1}^m \sum_{n=1}^N \lambda_n \phi_{in}^*(x)\phi_{in}(x), \tag{F.44}$$

$$EX_N^* X_N(x) = \sum_{i=1}^m \sum_{n=1}^N \sum_{n'=1}^N E\left(X_n^*(\omega)X_{n'}(\omega)\right) \phi_{in}^*(x)\phi_{in'}(x)$$

$$= \sum_{i=1}^m \sum_{n=1}^N \lambda_n \phi_{in}^*(x)\phi_{in}(x). \tag{F.45}$$

Thus, equation F.42 simplifies to

$$\|X(x) - X_N(x)\|^2 = \mathrm{tr}K(x,x) - \sum_{i=1}^m \sum_{n=1}^N \lambda_n \phi_{in}^*(x)\phi_{in}(x), \tag{F.46}$$

and following the application of Mercers theorem for matrix valued functions goes to zero as $N \to \infty$. □

Problem F.7.48 Prove the following corollary to the Karhunen-Loeve theorem expanding random processes in an orthogonal vector sequence.

Corollary F.2 *Let $\{X(\omega,x) = (X_1(\omega,x),\ldots,X_m(\omega,x))^*, x \in D \subset \mathbb{R}^d, \mathbb{Z}^d\}$ be an m-dimensional quadratic mean continuous random field. Then $X(\cdot)$ with scalar eigenfunctions $\{\phi_k\}$ has covariance*

$$\int_D K(x,y)\phi_k(y)\,dy = \Lambda_k\phi_k(x)\,, \qquad \Lambda_k = \underbrace{\begin{pmatrix} \Lambda_{11k} & \ldots & \Lambda_{1mk} \\ \vdots & \vdots & \vdots \\ \Lambda_{m1k} & \ldots & \Lambda_{mmk} \end{pmatrix}}_{\Lambda_k, m\times m \text{ matrix}} . \tag{F.47}$$

if and only if $X(\cdot)$ is the quadratic mean limit of random m-vectors $\{X_k = (X_{1k},\ldots,X_{mk})^*\}$,

$$X(x) \stackrel{q.m.}{=} \lim_{n\to\infty}\sum_{k=1}^{n} X_k\phi_k(x)\,, \tag{F.48}$$

with $X_k = \int_D \phi_k(x)^* X(x)\,dx$ orthogonal vectors $EX_kX_{k'}^* = \Lambda_k\delta(k-k')$.

Problem F.7.49 This problem should be compared with example 9.39.

Let

$$\frac{dX(\omega,t)}{dt} = -cX(\omega,t) + W(\omega,t)\,,\; U(0) = 0, \tag{F.49}$$

$W(\omega,\cdot)$ white-noise. Using the sifting property $EW(\omega,t)W(\omega,s) = \delta(t-s)$ to establish the covariance.

Then the impulse response (Green's function) $e^{-c(t-s)}u(t-s)$ implies

$$X(\omega,t) = \int_0^t e^{-c(t-s)}W(\omega,s)\,ds\,. \tag{F.50}$$

Solving for the covariance gives

$$\begin{aligned} K(t,s) = EX(\omega,t)X(\omega,s) &= \int_0^t\int_0^s e^{-c(t-r)}e^{-c(s-r')}EW(\omega,r)W(\omega,r')\,dr\,dr' \\ &= \int_0^t\int_0^s e^{-c(t-r)}e^{-c(s-r')}\delta(r-r')\,dr\,dr' \\ &\stackrel{t\geq s}{=} \frac{e^{-c(t-s)} - e^{-c(t+s)}}{2c} \qquad\qquad (\text{F.51}) \\ &= \int_0^t\int_0^s e^{-c(t-r)}e^{-c(s-r')}\delta(r-r')\,dr\,dr' \\ &\stackrel{t\leq s}{=} \frac{e^{+c(t-s)} - e^{-c(t+s)}}{2c}. \qquad\qquad (\text{F.52}) \end{aligned}$$

Problem F.7.50 Solve for the Green's function from Example 9.48 with arbitrary boundary conditions $G(0,y) = c_0, G(1,y) = c_1$. What non-trivial function is in the null-space of the operator.

Problem F.7.51 Solve for the Green's function from Example 9.48 with boundary conditions $G(0,y) = c_0, G'(0,y) = c_1$.

Problem F.7.52 Prove that the covariance operator K_X defined in 9.29 is positive definite and self-adjoint.

Problem F.7.53 Prove the following theorem.

Theorem F.3 *If X is Gaussian with respect to $L^2([0,1])$ with mean and covariance m_X, K_X of trace class, implies $\{X(t), t \in [0,1]\}$ is a q.m. continuous Gaussian process as defined by definition 9.29 in which finite linear combinations are gaussian.*

Proof The proof requires choosing a sequence $\{f_n, n = 1, 2 \ldots\} \subset L^2$ which sifts $\langle f_n, X\rangle \to X(t_i)$, and converges to a Gaussian random variable with the proper covariance. This provides the mechanism for showing that $X(t_1), \ldots, X(t_k)$, $t_i \in T$ is multivariate Gaussian with covariance K_X.

Let $X(\omega, t)$ be a q.m continuous process $t \in [0,1]$. Let

$$f_n(\cdot) = n 1_{[t_i - \frac{1}{2n}, t_i + \frac{1}{2n}]}(\cdot) , \qquad Y_n(\omega) = \int X(\omega, t) f_n(t)\, dt , \tag{F.53}$$

then $\lim_{n\to\infty} E|Y_n - X(t_i)|^2 = 0$.

Let $A_n = [t_i - (1/2n), t_i + (1/2n)]$, then

$$\begin{aligned} E|Y_n - X(t_i)|^2 &= \int | \int_{A_n} nX(\omega, t) - nX(\omega, t_i)\, dt|^2\, dP(\omega) \\ &= \int \int_{A_n} \int_{A_n} n^2 (X(\omega, t)X(\omega, s) - X(\omega, t_i)X(\omega, s) \\ &\quad - X(\omega, t)X(\omega, t_i) + X(\omega, t_i)X(\omega, t_i))\, ds\, dt\, dP(\omega) \\ &\overset{(a)}{=} \int_{A_n}\int_{A_n} n^2 K_X(t,s) - K_X(t_i, s) - K_X(t, t_i) + K_X(t_i, t_i)\, dt\, ds \\ &\overset{(b)}{=} K_X(\tilde{t}, \tilde{s}) + K_X(t_i, t_i) - K_X(t_i, \tilde{s}) - K_X(\tilde{t}, t_i) , \end{aligned}$$

for some $(\tilde{t}, \tilde{s}) \in A_n \times A_n$. Notice, (a) follows from Fubini's theorem and (b) uses the fact that $X(\omega, t)$ is q.m. continuous implying that $K_X(t,s)$ is continuous in both variables with the mean value theorem. Now as $A_n \to t_i$, with K_X continuous implies that for all $\epsilon > 0$ there exists N such that if $n > N$,

$$|K_X(\tilde{t}, \tilde{s}) - K_X(t_i, \tilde{s})| + |K_X(t_i, t_i) - K_X(\tilde{t}, t_i)| < \epsilon , \tag{F.54}$$

implying $\lim_{n\to\infty} E|Y_n - X(t_i)|^2 = 0$. □

Problem F.7.54 For the 3D cubic domain $D = [0,1]^3$ the boundary conditions become

$$\begin{aligned} u_1(0, x_2, x_3) &= u_1(1, x_2, x_3) = 0 \\ u_2(x_1, 0, x_3) &= u_2(x_1, 1, x_3) = 0 \\ u_3(x_1, x_2, 0) &= u_3(x_1, x_2, 1) = 0 \end{aligned}$$

$$\begin{aligned} \frac{\partial u_1(x_1, 0, x_3)}{\partial x_2} &= \frac{\partial u_1(x_1, 1, x_3)}{\partial x_3} = \frac{\partial u_1(x_1, x_2, 0)}{\partial x_3} = \frac{\partial u_1(x_1, x_2, 1)}{\partial x_3} = 0 \\ \frac{\partial u_2(0, x_2, x_3)}{\partial x_1} &= \frac{\partial u_2(1, x_2, x_3)}{\partial x_1} = \frac{\partial u_2(x_1, x_2, 0)}{\partial x_3} = \frac{\partial u_2(x_1, x_2, 1)}{\partial x_3} = 0 \\ \frac{\partial u_3(0, x_2, x_3)}{\partial x_2} &= \frac{\partial u_3(1, x_2, x_3)}{\partial x_3} = \frac{\partial u_3(x_1, 0, x_3)}{\partial x_3} = 0 . \end{aligned} \tag{F.55}$$

Prove that the set of eigenfunctions $\{\phi_{i,j,k}(1), \phi_{i,j,k}(2), \phi_{i,j,k}(3)\}$ in the 3D setting, with $D = [0,1]^3$, with the mixed boundary conditions becomes

$$\phi^{(1)}_{i,j,k} = \begin{pmatrix} \sqrt{\frac{8}{i^2+j^2+k^2}}\, i \cos \pi i x_1 \sin \pi j x_2 \sin \pi k x_3 \\ \sqrt{\frac{8}{i^2+j^2+k^2}}\, j \sin \pi i x_1 \cos \pi j x_2 \sin \pi k x_3 \\ \sqrt{\frac{8}{i^2+j^2+k^2}}\, k \sin \pi i x_1 \sin \pi j x_2 \cos \pi k x_3 \end{pmatrix},$$

$$\phi^{(2)}_{i,j,k} = \begin{pmatrix} -\sqrt{\frac{8}{i^2+j^2}}\, j \cos \pi i x_1 \sin \pi j x_2 \sin \pi k x_3 \\ \sqrt{\frac{8}{i^2+j^2}}\, i \sin \pi i x_1 \cos \pi j x_2 \sin \pi k x_3 \\ 0 \end{pmatrix},$$

$$\phi^{(3)}_{i,j,k} = \begin{pmatrix} \sqrt{\frac{8}{(i^2+j^2+k^2)(i^2+j^2)}}\, ik \cos \pi i x_1 \sin \pi j x_2 \sin \pi k x_3 \\ \sqrt{\frac{8}{(i^2+j^2+k^2)(i^2+j^2)}}\, jk \sin \pi i x_1 \cos \pi j x_2 \sin \pi k x_3 \\ -\sqrt{\frac{8}{(i^2+j^2+k^2)(i^2+j^2)}}\, (i^2+j^2) \sin \pi i x_1 \sin \pi j x_2 \cos \pi k x_3 \end{pmatrix}, \tag{F.56}$$

with each of the coefficients $\alpha_1, \alpha_2, \alpha_3$ chosen to give the eigenvectors unit energy, and with eigenvalues $\lambda^{(1)}_{i,j,k} = -\pi^2(2a+b)(i^2+j^2+k^2)$ and $\lambda^{(2)}_{i,j,k} = \lambda^{(3)}_{i,j,k} = -\pi^2 a(i^2+j^2+k^2)$.

F.8 Computational Anatomy: Comparison and Estimation of Anatomical Manifolds via Diffeomorphisms Chapter 16

Problem F.8.55 Simulate and display the result of the additive noise deformation with i.i.d. N(0,v) *e*-values applied to the images mentioned in the example F.9.60.

Problem F.8.56 (Background Space Deformations) Define some contrast image I_{temp}, a template, in the unit square $X = (0,1)^2$, for example

$$I(x_1, x_2) = \frac{1}{1 + (x_1 - .5)^2 + (x_2 - .5)^2}$$

or some photograph. Discretize X into the lattice $\mathbb{Z}_{L^2}$ with coordinates $(i,j) \in \mathbb{Z}_{L^2}$ and introduce the image from a background deformation

$$I^D(i,j) = I_{\text{temp}}\left[\frac{i}{L} + u_1(i,j), \frac{j}{L} + u_2(i,j)\right],$$

where $u_1(\cdot,\cdot), u_2(\cdot,\cdot)$ is a 2D vector field.

F.9 Markov Processes and Random Sampling Chapter 18

Problem F.9.57 (Brownian motion with drift.) Define $X(t) = \sigma W(t) + \mu t$, $W(t)$ standard Brownian motion. Then the infinitesimal parameters are $a(x) = \mu, b(x) = \sigma^2$ and the transition law

$$P_t(x, dy) = \frac{1}{\sqrt{(2\pi\sigma^2 t)}} e^{-(y-x-\mu t)^2/2t}\, dy\ .$$

Verify that the transition law satisfies the PDE for a diffusion.

Problem F.9.58 Show for $f \in C^\infty(\mathbb{R})$, that for any $y \in \mathbb{R}, \exists c$ in the interval $\alpha x + (1-\alpha)y$ such that

$$f(y) = f(x) + f'(x)(y-x) + f''(x)\frac{(y-x)^2}{2!} + (f''(c) - f''(x))\frac{(y-x)^2}{2!}.$$

Proof Define a value K implicitly by the equation:

$$f(y) = f(x) + f'(x)(y-x) + K\frac{(y-x)^2}{2}.$$

We wish to show that $K = f''(c)$ for some c in the interval with endpoints x and y. Define a function h on this interval by the rule:

$$h(t) = f(t) - f(x) - f'(x)(t-x) - K\frac{(t-x)^2}{2}.$$

Then $h(x) = 0$ and $h(y) = 0$. Because h is differentiable, from Rolle's Theorem $\exists b$ in the interval such that $h'(b) = 0$:

$$h'(t) = f'(t) - f'(x) - K(t-x),$$

with

$$K = \frac{f'(b) - f'(x)}{b - x}.$$

Since f' is differentiable, apply the Mean Value Theorem which says there is a c in the interval with endpoints x and b such that the right side of the last equation equals $f''(c)$. Thus, $K = f'$. □

Problem F.9.59 For a fixed connector, σ, as a linear graph with n sites prove that the generators g_t in the random configuration $c = \sigma(g_1, \ldots, g_t, \ldots, c_t)$ form a Markov chain, possibly heterogeneous so that the transition probability matrices depend upon t.

Because of the Markov chain, its simulation is straightforward. Starting with g_1, simulate it by its marginal distribution, then continue with g_2 and simulate its conditional distribution given g_1, and so on.

Problem F.9.60 With $\mathcal{G} = \{1, 2, \ldots, r\}$, full information bonds, and

$$A(g_1, g_2) \;=\; 1 - \frac{|g_1 - g_2|}{r},$$

simulate the random configuration.

Problem F.9.61 Consider the *Ising model* with $\mathcal{G} = \{1, 2\}$, $\Sigma = \text{TORUS}(L, L)$ meaning a 2D discrete torus with sites $(x, y); x, y = 0, 1, \ldots L-1$ and where site $x = L$ coincides with site $x = 0$ and similarly for the y coordinate. Introduce an acceptor matrix

$$A = \begin{pmatrix} 1 & b \\ b & c \end{pmatrix},$$

with positive constants b, c. Synthesize these patterns for different values of L, b, c, starting with $L = 32, b = 0.1, c = 1$ and display the results. What is the intuitive meaning of the constants in the acceptor matrix?

Problem F.9.62 Consider the torus $\text{TORUS}(L, L)$ but with another neighborhood definition. Let a site (x, y) have the 8 neighbors:

$$(x+1, y), \quad (x+1, y+1), \quad (x, y+1), \quad (x-1, y+1), \quad (x-1, y), \quad (x-1, y-1),$$
$$(x, y-1), \quad (x+1, y-1).$$

What is the chromatic number. How about a torus in 3D?

Problem F.9.63 Given $X(t) \in \mathbb{R}^n$ and $W(t) = n$-dimensional Wiener process, i.e. $W_1, W_2, \ldots, W_n$ are independent. Prove that $\pi(dx) = e^{-E(x)}/Z$ is invariant for

$$dX(t) = -\tfrac{1}{2}\nabla E(x(t)) + dW(t).$$

Problem F.9.64 Sometimes it is convenient to write A (or A_0) in exponential form. To control the degree of randomness one uses the temperature, T, as in

$$E(\beta_1, \beta_2) = (1/T)E_0(\beta_1, \beta_2),$$

or, if $\mathcal{B}$ is a vector space, the *scale parameter* ϵ as in

$$a(\beta_1, \beta_2) = a_0\left(\frac{\beta_1 - \beta_2}{\epsilon}\right).$$

Let $\mathcal{G} = \mathcal{B} = \mathbb{R}$, arity 2, *full information bonds* meaning that $\beta_j(g) \equiv g$, $\sigma =$ LINEAR, and affinity given by

$$a_0(x) = e^{-x^2},$$

with the weight function

$$Q(x) = e^{-x^2/v}.$$

Then P will be a Gaussian probability matrix with some covariance matrix; calculate it.

Problem F.9.65 With the same assumptions, let the scale parameter $\epsilon \to 0$; how does the probability measure behave? Or, in the less interesting case, $\epsilon \to \infty$?

Problem F.9.66 Simulate and graph the random configuration for various values of the scale parameter ϵ and examine the behavior.

Problem F.9.67 Let $\mathcal{G} = \mathbb{R}$, with $\omega(g) \equiv 4$, full information bonds, $\sigma = \text{LATTICE}(L, L)$, and the acceptor function

$$A_0(g_1, g_2) = \frac{1}{1 + (g_1 - g_2)^2}.$$

For different values of the scale factor ϵ synthesize the random configurations by solving the S.D.E.

To complete the problem formulation we must specify *boundary conditions* at the boundary $\partial\sigma$ of the connector. Reasoning in analogy with partial differential equations we shall speak of *Dirichlet condions* if we define $g = 0$ outside σ. We speak of *Neumann conditions* if we define $g_i = g_{i'}$ when the site i' outside σ is a neighbor of the site i in σ. Use Dirichlet conditions in the exercise.

Problem F.9.68 For the Ising model, $\mathcal{G} = \{1, 2\}$, and with the deformations

$$e(g^{\mathcal{D}} \mid g) \; = \epsilon \quad \textit{if} \quad g^{\mathcal{D}} \neq g, 1 - \epsilon \text{ else}$$

carry out image restoration by stochastic relaxation. Try the value $\epsilon = 0.1$. Display and discuss results.