C8.2: Stochastic analysis and PDEs
Solutions to Problem sheet 2

The questions on this sheet are divided into two sections. Those in the first section are compulsory and should be handed in for marking. Those in the second are extra practice questions and should not *be handed in.*

Section 1 (Compulsory)

1. *Show that a Markov pregenerator has the property that for $f \in \mathcal{D}(A)$, $\lambda \geq 0$ and $f - \lambda Af = g$, then $\|f\| \leq \|g\|$. Deduce that, in particular, g determines f uniquely.*

   By definition of Markov pregenerator, if $f \in \mathcal{D}(A)$, $\lambda \geq 0$ and $f - \lambda Af = g$, then

   $$\min_{\zeta \in E} f(\zeta) \geq \min_{\zeta \in E} g(\zeta).$$

   Apply this to both f and $-f$ we obtain

   $$-\min f \leq -\min g$$

   and

   $$\max f = -\min(-f) \leq -\min(-g) = \max g,$$

   so that $\|f\| \leq \|g\|$.

   For uniqueness, suppose that f_1 and f_2 both solve $f - \lambda Af = g$, then $f_1 - f_2$ solves the same equation with $g = 0$, and by the first part of the question $\|f_1 - f_2\| \leq 0$.

2. *Calculate the infinitesimal generator of the pure jump process X_t modelling the motion of a particle which, if it is currently at location x, it will wait an exponentially distributed amount of time with parameter $\alpha(x)$ before jumping to a new location determined by the probability measure $\mu(x, dy)$. You may assume that $\alpha(x)$ is uniformly bounded.*

   The result is easy, but the idea is to illustrate a method. We are going to condition on what happens to the process in a small time increment of length h. With probability $1 - e^{-\alpha(x)h} = \alpha(x)h + \mathcal{O}(h^2)$ it jumps to a new position. The probability that it jumps twice is $\mathcal{O}(h^2)$. The new position is determined by $\mu(x, dy)$. Thus

   $$\mathbb{E}_x[f(X_h)] = \alpha(x)h \int f(y)\mu(x, dy) + (1 - \alpha(x)h)f(x) + \mathcal{O}(h^2).$$

   Now subtract $f(x)$ from both sides, divide by h and let $h \to 0$, to obtain

   $$Af(x) = \alpha(x) \int (f(y) - f(x))\, \mu(x, dy),$$

   where we have used that $\int \mu(x, dy) = 1$.

3. *Check that each of the following is a Markov generator:*

   (a) $A = G - I$ *where* G *is a positive operator defined on all of* $C(E)$ *such that* $G1 = 1$.

   (b) $E = [0, 1]$ *and* $Af(\eta) = \frac{1}{2}f''(\eta)$ *with*

$$\mathcal{D}(A) = \{f \in C(E) : f'' \in C(E), f'(0) = 0 = f'(1)\}.$$

   (c) $E = [0, 1]$ *and* $Af(\eta) = \frac{1}{2}f''(\eta)$ *with*

$$\mathcal{D}(A) = \{f \in C(E) : f'' \in C(E), f''(0) = 0 = f''(1)\}.$$

(To check that the operator is closed in the second two cases, use that differentiation is a closed (unbounded) operator, so f'' is the composition of closed operators and therefore closed.) Or write $f_n(x) = f_n(0) + xf_n'(0) + \int_0^x \int_0^y f_n''(z)dzdy$ where $(f_n Af_n) \to (f, g)$ and deduce that $Af = g$.

That we have a Markov *pre*-generator in each case is easy. The only condition to check is that if $f \in \mathcal{D}(A)$, $\lambda \geq 0$ and $f - \lambda Af = g$ then $\min f(\zeta) \geq \min g(\zeta)$. For the second two cases, this is just elementary calculus. For the first, it is enough to check that if $f \in \mathcal{D}(A)$ and $f(\eta) = \min f(\zeta)$, then $Af(\eta) \geq 0$. Now $f - \min f \geq 0$ and G is a positive operator with $G1 = 1$, so $G(f - \min f) = Gf - \min f \geq 0$ and so $Gf \geq \min f$ which proves that $(G - I)f(\eta) \geq 0$ as required.

The more challenging task is to show that the range $\mathcal{R}(I - \lambda A) = C(E)$ for sufficiently small λ. For the first case this is straightforward: since G is bounded, choose λ such that $\lambda\|G\| < 1$, say. Then $(I - \lambda(G - I))f = g$ becomes $((1 + \lambda)I - \lambda G)f = g$ which can be solved as $f = \sum_{n=0}^{\infty} \lambda^n G^n g/(1 + \lambda)^{n+1}$.

For the other two cases, we must solve a differential equation of the form

$$f - \frac{\lambda}{2}f'' = g \tag{1}$$

on $[0, 1]$ with appropriate boundary conditions. This is standard: Set $\alpha^2 = 2/\lambda$, then the homogeneous equation has independent solutions $u(x) = e^{\alpha x}$, $v(x) = e^{-\alpha x}$ and the corresponding Wronskian $uv' - u'v = -2\alpha$. We seek a solution of the form $f = \phi u + \psi v$ and make the ansatz that $\phi' u + \psi' v = 0$. Substituting into the original equation, we find ourselves with a pair of simultaneous equations for ϕ' and ψ':

$$\phi' u + \psi' v = 0$$

(our ansatz) and

$$\phi' u' + \psi' v' = -\alpha^2 g.$$

Solving

$$\phi' = \frac{-\alpha^2 gv}{uv' - u'v} = -\frac{\alpha gv}{2}; \qquad \psi' = \frac{\alpha^2 gu}{u'v - uv'} = \frac{\alpha gu}{2}.$$

The general solution to (1) is then of the form

$$f(x) = e^{\alpha x} \int_x^1 \frac{\alpha}{2} g(y) e^{-\alpha y} dy + e^{-\alpha x} \int_0^x \frac{\alpha}{2} g(y) e^{\alpha y} dy + Ae^{\alpha x} + Be^{-\alpha x},$$

and evidently we can choose A and B in such a way that the boundary conditions are satisfied.

4. *Let $E = [0,1]$ and consider the operator $\mathcal{L}$ defined by $\mathcal{L}f(x) = f'(0)$ with*

$$\mathcal{D}(\mathcal{L}) = \{f \in C([0,1]) : f'(0) \text{ exists}\}.$$

Show that the closure of the graph of $\mathcal{L}$ does not correspond to the graph of a linear operator.

The point is that we obtain a 'multivalued' operator in the limit. For example, if we take the functions $f_{n,k}$ with $f_{n,k}(x) = 0$ for $x \in [1/n, 1]$, $f_{n,k}(0) = k/n$ and $f_{n,k}(x)$ linear on $[0, 1/n]$, then each of the sequences $f_{n,k}$ (with k fixed) converges to the (same) zero function as $n \to 0$, but the operator returns the value $-k$ for the kth sequence.

5. *(Discrete time martingale problem, Ethier & Kurtz, Chapter 4, Exercise 16)*

   (a) *Let E be a compact (or locally compact) space and $B(E)$ the bounded Borel measurable functions on E. Let $\mu(x, \Gamma)$ be a transition function on $E \times \mathcal{B}(E)$ and let $\{X(n)\}_{n\in\mathbb{N}}$ be a sequence of E-valued random variables. Define $A : B(E) \to B(E)$ by*

   $$Af(x) = \int_E f(y)\mu(x, dy) - f(x),$$

   and suppose that, for each $f \in B(E)$,

   $$f(X(n)) - \sum_{k=0}^{n-1} Af(X(k))$$

   is a martingale with respect to the natural filtration generated by X. Show that X is a Markov chain with transition function $\mu(x, \Gamma)$.
   Let $\mathcal{F}_n$ denote the natural filtration. Then, by the martingale property,

   $$\mathbb{E}\left[f(X_{n+1}) - \sum_{k=0}^{n} Af(X_k) \middle| \mathcal{F}_n \right] = f(X_n) - \sum_{k=0}^{n-1} Af(X_k),$$

   and so, rearranging,

   $$\mathbb{E}\left[f(X_{n+1}) \middle| \mathcal{F}_n \right] = f(X_n) + Af(X_n) = \int f(y)\mu(X_n, dy).$$

   Since this holds for all f, the result follows.

   (b) *Let $X(n)$, $n = 0.1. \ldots,$ be a sequence of $\mathbb{Z}$-valued random variables such that for each $n \geq 0$, $|X(n+1) - X(n)| = 1$. Let $g : \mathbb{Z} \to [-1, 1]$ and suppose that*

   $$X(n) - \sum_{k=0}^{n-1} g(X(k))$$

   is a martingale with respect to the natural filtration generated by X. Show that X is a Markov chain and calculate its transition probabilities in terms of g.
   Using the martingale property,

   $$\mathbb{E}\left[X(n+1) \middle| \mathcal{F}_n \right] = X(n) + g(X(n)),$$

   and since we know that $|X(n+1) - X(n)| = 1$, this says that

   $$\mathbb{P}\left[X(n+1) - X(n) = 1 \middle| \mathcal{F}_n \right] - \mathbb{P}\left[X(n+1) - X(n) = -1 \middle| \mathcal{F}_n \right] = g(X(n)).$$

Thus

$$\mathbb{P}\left[X(n+1)-X(n)=1 \mid \mathcal{F}_n\right]=\frac{g(X(n))+1}{2},$$

and the result follows.

6. *Recall that the Wright-Fisher diffusion, which takes values in* $[0,1]$ *has generator*

$$Af(x)=\frac{1}{2}x(1-x)f''(x),$$

when restricted to an appropriate subset of the twice continuously differentiable functions on $[0,1]$. *By considering the martingale problem with suitable functions x:*

(a) *Show that* $X_\infty=\lim_{t\to\infty}X_t$ *exists and find its expectation;*
Since $f(x)=x$ has $Af=0$, X_t is a bounded martingale and X_∞ exists by martingale convergence and $\mathbb{E}[X_\infty]=x$.

(b) *Show that* $\mathbb{P}[X_\infty\in\{0,1\}]=1$ *and (using your previous calculation) find* $\mathbb{P}[X_\infty=1]$*;*
Note that setting $f(x)=x(1-x)$, $Af(x)=-x(1-x)$ and so

$$X_t(1-X_t)+\int_0^t X_s(1-X_s)ds$$

is a positive martingale. As such it converges to a bounded limit, but this is only possible if $\mathbb{P}[X_\infty\in\{0,1\}]=1$.

(c) *Find* $\mathbb{E}[\int_0^\infty X_s(1-X_s)ds]$. As a corollary of the previous part, $\mathbb{E}_x[\int_0^\infty X_s(1-X_s)ds]=x(1-x)$.

Now take $f(x)=2x\log x+2(1-x)\log(1-x)$. *Although f isn't in the domain of the generator, we can find a twice continuously differentiable function which equals f on* $[\epsilon,1-\epsilon]$ *and is in the domain. Taking this on trust, find an expression for the expected hitting time of* $\{\epsilon,1-\epsilon\}$ *and hence of* $\{0,1\}$.

For this choice of function, $Af(x)\equiv 1$ and so

$$f(X_t)-\int_0^t 1ds$$

is a martingale. Setting τ_ϵ to be the first hitting time of $\{\epsilon,1-\epsilon\}$,

$$\mathbb{E}_x[f(X_{\tau_\epsilon})-\tau_\epsilon]=f(x),$$

or

$$\begin{aligned}\mathbb{E}_x[\tau_\epsilon]&=-2x\log x-2(1-x)\log(1-x)+2\epsilon\log\epsilon+2(1-\epsilon)\log(1-\epsilon)\\&\to -2x\log x-2(1-x)\log(1-x)\ \text{as}\ \epsilon\to 0.\end{aligned}$$

Section 2 (Extra practice questions, not for hand-in)

A. *Show that if a Markov pregenerator is everywhere defined and is a bounded operator, then it is automatically a Markov generator. [Hint: A bounded operator is automatically closed. To check that $\mathcal{R}(I-\lambda A) = C(E)$ for sufficiently small λ, it suffices to solve $f - \lambda A f = g$, for which you can try a 'geometric series' $f = \sum_{n=0}^{\infty} \lambda^n A^n g$, just as we did on the previous problem sheet.]*

   The hint says it all really. Since A is bounded, choose λ so that $\|\lambda A\| < 1$ and then $(I-\lambda A)f = g$ is solved by

$$f = \sum_{n=0}^{\infty} (\lambda A)^n g$$

   (which is well defined).

B. *(Brownian Motion with sticky boundary.) Show that $Af = \frac{1}{2}f''$ on*

$$\mathcal{D}(A) = \{f \in C([0,\infty)) : f', f'' \in C([0,\infty)), f'(0) = cf''(0)\}$$

   for a fixed $c > 0$ defines a Markov pregenerator.

   This is just a question of fitting another set of boundary conditions for equation (1).

   The corresponding stochastic process is called sticky *Brownian motion. It interpolates between absorbing and reflecting Brownian motion on the half line. In particular, unlike reflecting Brownian motion, the Lebesgue measure of $\{t : X_t = 0\}$ is positive. Indeed one can check that*

$$\mathbb{E}_0\left[\int_0^{\infty} \alpha e^{-\alpha t} \mathbf{1}_{X_t > 0} dt\right] = \frac{1}{1 + c\sqrt{2\alpha}}. \tag{2}$$

   If you want to try to show this, use the fact that since the semigroups are known explicitly in the absorbing and reflecting cases, in an obvious notation we can solve the equations

$$f_a - \lambda A_a f_a = g, \qquad f_r - \lambda A_r f_r = g$$

   explicitly. Since the form of the generators is the same (only the domains differ), one can solve $f - \lambda A f = g$ by taking f to be a constant multiple of $f_r''(0) f_a(x) + c f_a'(0) f_r(x)$. That provides an expression for

$$\mathbb{E}\left[\int_0^{\infty} \alpha e^{-\alpha t} g(X_t) dt\right].$$

C. *Let X be a strong Markov process with Markov generator A on a compact set E. Let $P_t^{\lambda} f(x) = \mathbb{E}^x \exp(-\lambda t) f(X_t)$.*

   (a) *Show by the strong Markov property that for a stopping time τ we have*

$$P_{\tau}^{\lambda} P_t^{\lambda} = P_{\tau+t}^{\lambda},$$

   and hence that we have Dynkin's formula for the resolvent $R_{\lambda} = \int_0^{\infty} e^{-\lambda t} P_t dt$; for $g \in C(E), \lambda > 0, x \in E$ that

$$R_{\lambda} g(x) = \mathbb{E}^x \int_0^{\tau} e^{-\lambda t} g(X_t) dt + P_{\tau}^{\lambda} R_{\lambda} g(x).$$

(b) Apply this to $g = (\lambda - A)f$, for $f \in \mathcal{D}(A)$, to obtain

$$\mathbb{E}^x e^{-\lambda\tau} f(X_\tau) - f(x) = \mathbb{E}^x \int_0^\tau e^{-\lambda s}(A-\lambda)f(X_s)ds. \tag{3}$$

Now let $\lambda \to 0$ to obtain for x such that $E^x\tau < \infty$,

$$\mathbb{E}^x f(X_\tau) - f(x) = \mathbb{E}^x \int_0^\tau Af(X_s)ds.$$

(c) Let X be a Brownian motion and define $T_a = \inf\{t : X_t = a\}$ to be the first hitting time of the point $a > 0$. Working over $C_0(\mathbb{R})$ and applying formula (3) to $f(x) = \exp(\theta x)I_{x\leq a}$ for a suitably chosen θ show that

$$\mathbb{E}^x e^{-\lambda T_a} = e^{-a\sqrt{2\lambda}}, \ \ \forall\lambda \geq 0.$$

(a) This is a simple application of the strong Markov property

$$\begin{aligned} P^\lambda_{\tau+t}f(x) &= \mathbb{E}^x e^{-\lambda(\tau+t)}f(X_{\tau+t}) \\ &= \mathbb{E}^x e^{-\lambda\tau}\mathbb{E}(e^{-\lambda t}f(X_{\tau+t})|\mathcal{F}_\tau) \\ &= \mathbb{E}^x e^{-\lambda\tau}\mathbb{E}^{X_\tau}(e^{-\lambda t}f(X_t)) \\ &= \mathbb{E}^x e^{-\lambda\tau}P^\lambda_t f(X_\tau \\ &= P^\lambda_\tau P^\lambda_t f(x) \end{aligned}$$

Using this we have

$$\begin{aligned} R_\lambda g(x) &= \int_0^\infty P^\lambda_t f(x)dt \\ &= \mathbb{E}^x\int_0^\infty e^{-\lambda t}g(X_t)dt \\ &= \mathbb{E}^x\int_0^\tau e^{-\lambda t}g(X_t)dt + \int_\tau^\infty e^{-\lambda t}g(X_t)dt \\ &= \mathbb{E}^x\int_0^\tau e^{-\lambda t}g(X_t)dt + \int_0^\infty e^{-\lambda(s+\tau)}g(X_{s+\tau})dt \\ &= \mathbb{E}^x\int_0^\tau e^{-\lambda t}g(X_t)dt + \int_\tau^\infty P^\lambda_{\tau+s}g(x)dt \\ &= \mathbb{E}^x\int_0^\tau e^{-\lambda t}g(X_t)dt + \int_\tau^\infty P^\lambda_\tau P^\lambda_s g(x)dt \\ &= \mathbb{E}^x\int_0^\tau e^{-\lambda t}g(X_t)dt + P^\lambda_\tau R_\lambda g(x). \end{aligned}$$

(b) This is straightforward as $R_\lambda(\lambda - A)f(x) = f(x)$, so

$$f(x) = \mathbb{E}^x\int_0^\tau e^{-\lambda t}(\lambda - A)f(X_t)dt + \mathbb{E}^x e^{-\lambda\tau}f(X_\tau).$$

Rearranging gives the result.
Now as $\mathbb{E}^x|\int_0^\tau e^{-\lambda t}(\lambda - A)f(X_t)dt| \leq \mathbb{E}^x\tau\|f\| < \infty$ we can apply DOM to take $\lambda \to 0$ to get the result.

(c) Applying this in the case where $\tau = T_a$ and X is Brownian motion with generator $Af = \frac{1}{2}f''$ we get setting $\theta = \sqrt{2\lambda}$ that

$$\mathbb{E}^0 e^{-\lambda T_a + \sqrt{2\lambda} X_{T_a}} - 1 = 0,$$

and rearranging gives the result.

D. *Show that almost sure convergence implies convergence in distribution.*

Recall that $X_n \to X$ in distribution if for all bounded continuous functions $g : E \to \mathbb{R}$,

$$\lim_{n\to\infty} \mathbb{E}[g(X_n)] = \mathbb{E}[g(X)].$$

Suppose that $X_n \to X$ almost surely, and that g is bounded and continuous, then $g(X_n) \to g(X)$ almost surely and since g is bounded, the Dominated Convergence Theorem tells us that $\mathbb{E}[g(X_n)] \to \mathbb{E}[g(X)]$ as required.

E. *Prove the Portmanteau Theorem:*

Theorem 0.1 (Portmanteau Theorem)**.** *Let $(X_n)_{n\geq 1}$ be a sequence of random variables taking values in S. The following are equivalent.*

(i) *$X_n \to X$ in distribution.*

(ii) *For any closed set $K \subseteq S$, $\limsup_{n\to\infty} \mathbb{P}[X_n \in K] \leq \mathbb{P}[X \in K]$.*

(iii) *For any open set $O \subseteq S$, $\liminf_{n\to\infty} \mathbb{P}[X_n \in O] \geq \mathbb{P}[X \in O]$.*

(iv) *For all Borel sets $A \subseteq S$ such that $\mathbb{P}[X \in \partial A] = 0$, $\lim_{n\to\infty} \mathbb{P}[X_n \in A] = \mathbb{P}[X \in A]$.*

(v) *For any bounded function f, denote by D_f the set of discontinuities of f. Then for any f such that $\mathbb{P}[X \in D_f] = 0$, $\mathbb{E}[f(X_n)] \to \mathbb{E}[f(X)]$ as $n \to \infty$.*

(i) $\Longrightarrow$ (ii) Let $g_n(x) = 1 - (nd(x,K) \wedge 1)$, which is continuous and bounded, is 1 on K and converges pointwise to $\mathbf{1}_K$. Then, for every n,

$$\limsup_{k\to\infty} \mathbb{P}[X_k \in K] \leq \limsup_{k\to\infty} \mathbb{E}[g_n(X_k)] = \mathbb{E}[g_n(X)].$$

Now let $n \to \infty$ and use that the $g_n(X) \leq 1$ and converges pointwise to $\mathbf{1}_K(X)$ plus the Dominated Convergence Theorem.

(ii) $\Longrightarrow$ (iii) Set $K = G^c$, which is closed, and use complements.

(iii) $\Longrightarrow$ (iv) Let G be the interior of A and K the closure. Then, by assumption, $\mathbb{P}[X \in G] = \mathbb{P}[X \in K] = \mathbb{P}[X \in A]$ and so we may use (iii) (and (ii) which follows immediately from (iii)) to obtain

$$\limsup_{n\to\infty} \mathbb{P}[X_n \in A] \leq \limsup_{n\to\infty} \mathbb{P}[X_n \in K] \leq \mathbb{P}[X_n \in K] = \mathbb{P}[X \in A],$$

and

$$\liminf_{n\to\infty} \mathbb{P}[X_n \in A] \geq \liminf_{n\to\infty} \mathbb{P}[X_n \in G] \geq \mathbb{P}[X_n \in G] = \mathbb{P}[X \in A],$$

and the result follows.

(iv) $\Longrightarrow$ (v) From (iv) we have convergence for g of the form $g(x) = \sum_{n=1}^N a_n \mathbf{1}_{A_n}$ where A_n satisfies $\mathbb{P}[X \in \partial A_n] = 0$. We call such functions elementary. Given g as in (v), observe that for every $a < b$, with possibly a countable number of exceptions,

$$\mathbb{P}\big[X \in \partial\{x : g(x) \in (a,b]\}\big] = 0.$$

Indeed, if $X \in \partial\{x : g(x) \in (a,b]\}$ then either g is discontinuous in X or $g(X) = a$ or $g(X) = b$. The first event has probability zero and so have the last two except for possibly at a countable set of values of a, b. By decomposing the real axis in suitable small intervals, we thus obtain an increasing sequence g_k and a decreasing sequence h_k of elementary functions, both converging pointwise to g. Now, for all k,

$$\limsup_{n\to\infty} \mathbb{E}[g(X_n)] \leq \limsup_{n\to\infty} \mathbb{E}[h_k(X_n)] = \mathbb{E}[h_k(X)],$$

and

$$\liminf_{n\to\infty} \mathbb{E}[g(X_n)] \geq \liminf_{n\to\infty} \mathbb{E}[g_k(X_n)] = \mathbb{E}[g_k(X)],$$

and the right hand sides of both converge as $k \to \infty$ (by bounded convergence) to $\mathbb{E}[g(X)]$.

(v) $\Longrightarrow$ **(i)** This is trivial.