# 2) Feller processes

. In this chapter we will consider more general state spaces $S$, but also add one continuity constraint (the Feller property) that will ensure the one-to-one correspondence between the analogue of the Q-matrix (the generator) and the Markov process itself. Before going in the definition of the processes and generator, let us give some preliminaries.

## 2.1) Preliminaries

. In this chapter we consider:

. State space $S$: a compact or locally compact, separable metric space.

[Recall: $S$ is compact metric space: every sequence in $S$ has a converging subsequence in $S$.

. $S$ is locally compact if $\forall x \in S$, $\exists \varepsilon > 0$ s.t. the closure of $\{y \in S \mid |x-y| < \varepsilon\}$ is compact.

. $S$ is separable if it contains a countable dense set.)

Example: $S = \mathbb{R}^n$ with euclidian distance is a locally compact separable metric space.

. Space $C(S)$: . For $S$ compact, it is the space of continuous real-valued functions on $S$

. For $S$ locally compact, it is the space of continuous real-valued functions vanishing at infinity.

[Recall: . A fct $f$ is vanishing at $\infty$ if $\forall \varepsilon > 0$, $\exists$ compact set $K \subset S$ s.t. $|f(x)| < \varepsilon$, $\forall x \in S \setminus K$].

Rem.: • Every continuous function that vanishes at infinity is uniformly continuous (this property is often needed).
• $C(S)$ is separable.

• Uniform norm $\|\cdot\|$: In both cases we use the uniform norm defined by

$$\|f\| := \sup_{x \in S} |f(x)|$$

• With this norm, $C(S)$ is a Banach space.

## 2.2) Feller processes

**Def. 2.1)** Let us set:

• $\Omega = D([0,\infty)) \equiv \{$set of right-continuous functions $\omega : [0,\infty) \to S$ with left limits$\}$ (aka càdlàg functions).

• $\mathcal{F}$ = the smallest $\sigma$-algebra on $\Omega$ such that the mapping $\omega \to \omega(t)$ is measurable $\forall t \geq 0$.

• Denote $X(t,\omega) := \omega(t)$ and $(\theta_s \omega)(t) = \omega(t+s)$.

• Given a collection of probability measures $\{\mathbb{P}^x, x \in S\}$ on $(\Omega, \mathcal{F})$ and a right-continuous filtration $\{\mathcal{F}_t, t \geq 0\}$ on $(\Omega, \mathcal{F})$.

• Then, a stochastic process $X$ with state space $S$ is a Feller process if:

(a) $X$ is $\{\mathcal{F}_t, t \geq 0\}$ adapted,

(b) $\mathbb{P}^x(X(0) = x) = 1$,

(c) $\mathbb{E}^x(Y \circ \theta_s \mid \mathcal{F}_s) = \mathbb{E}^{X(s)}(Y)$ $\mathbb{P}^x$-a.s., $\forall x \in S$, $\forall$ bounded measurable $Y$ on $\Omega$.

(d) the mapping

$$x \to \mathbb{E}^x[f(X(t))]$$

is in $C(S)$, $\forall f \in C(S)$ and $t \geq 0$.

Remarks: . © is the Markov property.

. (d) is called the Feller property.

Here is what the continuity in $x$ means:

. If $x_n \xrightarrow[n\to\infty]{} x$, then the distribution of $X(t)$ for the process starting at $x_n$ converges weakly to that for the process starting at $x$.
[i.e., $\forall$ bounded continuous real-valued functions $f$ on $C(S)$, $\lim_{n\to\infty} \mathbb{E}^{x_n}(f(X_t)) = \mathbb{E}^{x}(f(X(t)))$.]

. Every Feller process has the strong Markov property:

Thm. 2.2) If $Y_s(\omega)$ is bounded and jointly measurable on $[0,\infty)\times\Omega$ and $\tau$ a stopping time, then for every $x \in S$,

$$\mathbb{E}^x(Y_\tau \circ \theta_\tau \mid \mathcal{F}_\tau) = \mathbb{E}^{X(\tau)}(Y_\tau) \quad \mathbb{P}^x\text{-a.s. on } \{\tau<\infty\}.$$

. The proof uses the same ingredients as in the Brownian motion case.

2.3) The semigroup.

. The second element is the analogue of the transition function. For Markov Chains, given a transition function $p_t(x,y)$, we can define $T(t)f(x) = \sum_y p_t(x,y) f(y)$ for $f \in C(S)$.
Then, for example the Chapman-Kolmogorov gives

$$T(t)\,T(s)f = T(s+t)f.$$

. For general $S$, one defines the probability semigroup as follows.

**Def. 23)** A probability semigroup is a family of continuous linear operators $\{T(t), t \geq 0\}$ on $C(S)$ satisfying:

(a) $T(0)f = f$ for all $f \in C(S)$,

(b) $\forall f \in C(S)$, $\lim_{t \downarrow 0} T(t)f = f$,

(c) $T(t+s)f = T(t)T(s)f$, $\forall f \in C(S)$,

(d) $T(t)f \geq 0$ for every nonnegative $f \in C(S)$,

(e). For $S$ compact: $T(t)1 = 1$, $\forall t > 0$.

. For $S$ non compact: $\exists f_n \in C(S)$ with $\sup_n \|f_n\| < \infty$ so that $T(t)f_n \longrightarrow 1$ pointwise $\forall t \geq 0$.

. (e): Since $T(0)f_n = f_n$, we have implicitely that $f_n \to 1$ pointwise too.

**. Remarks:** . (b) means that $\lim_{t \downarrow 0} \|T(t)f - f\| = 0$.

. (c) is the analogue of the Chapman-Kolmogorov equations $\Rightarrow$ $T(t)$ and $T(s)$ commute.

(1) . (d) + (e) implies: $\|T(t)f\| \leq \|f\|$, $\forall f \in C(S)$,
i.e., $T(t)$ is a contraction operator.

. (b) is also known as strong continuity;

(2) . (b) + (c) + contraction $\Rightarrow$ $t \longrightarrow T(t)f$ from $[0, \infty)$ to $C(S)$ is continuous.

**Indeed:**

. (1) For $S$ compact: by def of the sup norm, $f \leq \|f\| \cdot 1$. Therefore ($T = T(t)$)
$T(\|f\|1 - f) \overset{(d)}{\geq} 0$ and by linearity: $\|f\| T1 \geq Tf$.
by (e), $T1 = 1$ and taking the $\|\cdot\|$ on both sides $\Rightarrow \|f\| \geq \|Tf\|$.

. (2) $\|T(t+\varepsilon)f - T(t)f\| \overset{(c)}{=} \|T(t)\underbrace{(T(\varepsilon) - \mathbb{1})f}_{\in C(S)}\| \overset{\text{contr.}}{\leq} \|(T(\varepsilon) - \mathbb{1})f\| \xrightarrow[\varepsilon \to 0]{(b)} 0$.

**Example:** Consider $S = \mathbb{R}$ and $X(t)$ be a standard Brownian motion.

Then, $T(t)$ defined by

$$T(t)f(x) := \mathbb{E}^x(f(X(t)))$$

is a probability semigroup (see exercises).

It is called the **Brownian semigroup.**

- Tightly related with the semigroup we have the so-called resolvent of the semigroup.

**Def. 2.4)** Let $T$ be a probability semigroup. Then its Laplace transform

$$U(\lambda)f := \int_0^\infty e^{-\lambda t} T(t)f \, dt, \quad \lambda > 0,$$

is called the **resolvent of $T$.**

**Rem.:** Since $t \to e^{-\lambda t} T(t)f$ is continuous and

$$\|e^{-\lambda t} T(t)f\| \leq e^{-\lambda t} \|f\|$$

$\Rightarrow U(\lambda)f$ is well-defined. (because $\| \int_0^\infty e^{-\lambda t} T(t)f \, dt \| \leq$

$$\leq \int_0^\infty \| e^{-\lambda t} T(t)f \, dt \| \leq \frac{\|f\|}{\lambda}).$$

- Here are some properties:

**Lemma 2.5)**

(a) $U(\lambda)$ is a linear operator on $C(S)$

(b) $\|U(\lambda)f\| \leq \frac{\|f\|}{\lambda}$

(c) $\lim_{\lambda \to \infty} \lambda U(\lambda)f = f$

(d) $U(\lambda) - U(\mu) = (\mu - \lambda) U(\lambda) U(\mu)$. (Resolvent equation)

$\Rightarrow$ $U(\lambda)$ and $U(\mu)$ commute.

**Proof:** (a) follows by the definition.

(b) $\|U(\lambda)f\| = \|\int_0^\infty e^{-\lambda t} T(t) f \, dt\|$

$$\leq \int_0^\infty e^{-\lambda t} \|T(t)f\| dt$$

$$\underset{\text{contr.+ lin.}}{\leq} \|f\| \int_0^\infty e^{-\lambda t} dt = \frac{\|f\|}{\lambda}.$$

(c) $\|\lambda U(\lambda) f - f\| = \|\lambda \int_0^\infty dt\, e^{-\lambda t} T(t) f - f\|$

$$\overset{t\lambda = u}{=} \|\int_0^\infty du\, e^{-u} (T(\tfrac{u}{\lambda}) - \mathbb{1}) f\|$$

$$\leq \int_0^\infty du\, e^{-u} \underbrace{\|(T(\tfrac{u}{\lambda}) - \mathbb{1}) f\|}_{\xrightarrow[\lambda\to\infty]{} 0} \xrightarrow[\lambda\to\infty]{} 0.$$

(c) For $\lambda = \mu$ : $0 = 0$. ✓

Let $\lambda \neq \mu$. Then,

$$U(\lambda)U(\mu) f = \int_0^\infty dt\, e^{-\lambda t} T(t)\, U(\mu) f$$

$$= \int_0^\infty dt\, e^{-\lambda t} T(t) \int_0^\infty ds\, e^{-\mu s} T(s) f$$

$T(t)f(x)$ is jointly measurable in $(x,t)$

$$= \int_0^\infty dt\, e^{-\lambda t} \int_0^\infty ds\, e^{-\mu s} \underbrace{T(t)T(s) f}_{= T(t+s) f}$$

$$\overset{t+s=u}{=} \int_0^\infty du\, T(u) f \underbrace{\int_0^u dt\, e^{-\lambda t} \cdot e^{-\mu(u-t)}}$$

$$= \frac{e^{-\mu u} \cdot e^{-t(\lambda-\mu)}}{\mu - \lambda}\Big|_0^u$$

$$= \frac{e^{-\lambda u} - e^{-\mu u}}{\mu - \lambda}$$

$$= \frac{1}{\mu - \lambda} [U(\lambda) f - U(\mu) f]. \quad \#$$

**Example:** The resolvent of the Brownian motion is given by

$$(U(\lambda) f)(x) = \frac{1}{\sqrt{2\lambda}} \int_{\mathbb{R}} dy\, e^{-\sqrt{2\lambda}|x-y|} f(y).$$

This comes from: $(U(\lambda)f)(x) = \int_0^\infty dt\, e^{-\lambda t} \int_{\mathbb{R}} dy \frac{e^{-\frac{(x-y)^2}{2t}}}{\sqrt{2\pi t}} f(y)$

$$= \int_{\mathbb{R}} dy\, f(y) \int_0^\infty dt \frac{e^{-\lambda t} \cdot e^{-\frac{(x-y)^2}{2t}}}{\sqrt{2\pi t}} \underset{\lambda t = 1}{=} \int_{\mathbb{R}} dy \frac{f(y)}{\sqrt{2\lambda}} \left( \int_0^\infty \frac{ds}{\sqrt{\pi}} \frac{e^{-\frac{1}{s}}}{s^{3/2}} \cdot e^{-\mu s} \right) \quad ✓.$$

with $\mu = \frac{(x-y)^2\lambda}{2}$. Finally: $\int_0^\infty ds \frac{e^{-1/s} e^{-\mu s}}{\sqrt{\pi}\, s^{3/2}} = e^{-2\sqrt{\mu}}$ (for $\mu > 0$).

5.11.2013

## 2.4) The generator.

- The final ingredient is the generator, that describes the infinitesimal dynamics of the process. It is the analogue of the Q matrices for continuous time Markov chains.
- Consider for a moment still the framework of Markov chains and compute $\frac{d}{dt}\mathbb{E}(f(X_t))$ where $X_t$ is the Markov chain at time $t$ with initial distribution $\pi_0$. We denote by $\mu_t = \pi_0 P_t$ the law at time $t$.

$$\Rightarrow \mathbb{E}(f(X_t)) = \sum_y \mu_t(y) f(y) \equiv \langle \mu_t, f \rangle.$$

$$\frac{d}{dt}\mathbb{E}(f(X_t)) = \sum_x \pi_0(x) \sum_y \frac{d}{dt} P_t(x,y) f(y) \quad (*)$$

By the Chapman-Kolmogorov equation,

$$P_{t+s}(x,y) = \sum_z P_s(x,z) P_t(z,y)$$

we have derived the backwards equation

$$\frac{d}{dt} P_t(x,y) = \sum_z q(x,z) P_t(z,y) \qquad (KBE)$$

by differentiating w.r.t. $s$ at $s=0$. Similary, by differentiating w.r.t. $t$ at $t=0$ one gets

$$\frac{d}{dt} P_t(x,y) = \sum_z P_t(x,z) q(z,y) \qquad (KFE)$$

which is called the <u>forward equation</u>.

- Replacing (KFE) into (*) we get

$$\frac{d}{dt}\langle \mu_t, f \rangle = \sum_x \pi_0(x) \sum_z P_t(x,z) \sum_y q(z,y) f(y)$$

Now, since $\sum_y q(z,y) = 0 \Rightarrow$ we can write

$$\cdot \frac{d}{dt}\langle \mu_t, f\rangle = \sum_x \pi_0(x) \sum_z P_t(x,z) \sum_y q(z,y)\big(f(y)-f(z)\big)$$

Now, define the operator $\mathcal{L}$ acting on functions:

$$(\mathcal{L}f)(x) = \sum_y q(x,y)\,\big(f(y)-f(x)\big) \quad (\text{i.e., } \mathcal{L} = Q - \mathbb{I}).$$

Then, $\underline{\frac{d}{dt}\langle \mu_t, f\rangle = \sum_z \mu_t(z)\,(\mathcal{L}f)(z) = \langle \mu_t, \mathcal{L}f\rangle}$ (A)

(Formally: $\langle \mu_t, f\rangle = \langle \mu_0, e^{t\mathcal{L}} f\rangle$)

Rem.: One can also think of $\mathcal{L}^*$ being the adjoint of $\mathcal{L}$ acting on measures s.t. $\underline{\frac{d}{dt}\langle \mu_t, f\rangle = \langle \mathcal{L}^* \mu_t, f\rangle}$. (B)

(Formally, $\langle \mu_t, f\rangle = \langle e^{t\mathcal{L}^*}\mu_0, f\rangle$).

- In (A) one think at the variation of $\langle \mu_t, f\rangle$ being given by a change in the "observable" $f$ by $\mathcal{L}f$, while in (B) is the change of measure that modifies $\langle \mu_t, f\rangle$.

- Now we define the generator $\mathcal{L}$ for the general case of Feller processes.

**Def. 2.6)** Denote by $\underline{D}$ the domain and by $\underline{R}$ the range of an operator.

A probability generator is a linear operator $\mathcal{L}$ on $C(S)$ satisfying:

(a) $D(\mathcal{L})$ is dense in $C(S)$,

(b) If $f \in D(\mathcal{L})$, $\lambda > 0$ and $f - \lambda\mathcal{L}f = g$, then $\inf_{x\in S} f(x) \geq \inf_{x\in S} g(x)$,

(c) $R(\mathbb{I} - \lambda\mathcal{L}) = C(S)$ for all sufficiently small $\lambda > 0$,

(d) For $S$ compact: $\mathcal{L}1 = 0$

For $S$ non compact: for small $\lambda > 0$, $\exists\, f_n \in D(\mathcal{L})$ (which can depend on $\lambda$) s.t. $g_n := f_n - \lambda\mathcal{L}f_n$ satisfies: $\sup_n \|g_n\| < \infty$, $f_n, g_n \to 1$ pointwise.

**Remarks:** (d) is natural, since for the function $f = 1$ in (A) we would get $0 = \frac{d}{dt}\langle \mu_t, 1\rangle = \langle \mu_t, \mathcal{L}1\rangle$.

- There will be a relation with a semigroup. For instance, if $T(t)f := \langle \mu_t, f\rangle$, then formally

$$\frac{d}{dt}\langle \mu_t, f\rangle\Big|_{t=0} = \lim_{t\downarrow 0} \frac{T(t)f - T(0)f}{t} = \lim_{t\downarrow 0} \frac{T(t)f - f}{t}.$$

$$\Rightarrow \langle \mu_t, \mathcal{L}f\rangle\Big|_{t=0} = T(0)\mathcal{L}f = \mathcal{L}f.$$

**Lemma 2.7)** Let $f \in D(\mathcal{L})$, $\lambda \geq 0$ and $f - \lambda\mathcal{L}f = g$. Then $\|f\| \leq \|g\|$.

Further, for $\lambda$ small enough, $(1-\lambda\mathcal{L})^{-1}$ is a contraction that maps non-negative functions to non-negative functions.

**Proof:** - By (b) we have $\inf_{x\in S} g(x) \leq \inf_{x\in S} f(x)$.

But also, replacing $(f,g)$ by $(-f,-g)$,

$$\sup_{x\in S} f(x) \leq \sup_{x\in S} g(x)$$

Thus $\|f\| \leq \|g\|$.

- Further, by (c) for $\lambda$ small enough

$(1-\lambda\mathcal{L})^{-1} g \in D(\mathcal{L})$ for all $g \in C(S)$

and if we take $g \geq 0$,

$$0 \leq \inf_{x\in S} g(x) \leq f(y) \leq \sup_{x\in S} g(x) \equiv \|g\|, \quad \forall y \in S,$$

thus, $\|f\| = \|(1-\lambda\mathcal{L})^{-1} g\| \leq \|g\|$. #

Let us now see an example:

**Example:** . Let $S = \mathbb{R}$, $(\mathcal{L}f)(x) := \frac{d}{dx} f(x)$ with domain

$$\mathcal{D}(\mathcal{L}) := \{ f \in C(\mathbb{R}) \mid f' \in C(\mathbb{R}) \}.$$

. Then, $\mathcal{L}$ is a probability generator.

. Let us verify the (a)-(d) of the definition.

(a) Since $C_0^\infty(\mathbb{R}) \subset \mathcal{D}(\mathcal{L})$ and $C_0^\infty(\mathbb{R})$ is dense in $C(\mathbb{R}) \Rightarrow$ also $\mathcal{D}(\mathcal{L})$ is dense in $C(\mathbb{R})$. ✓

(b) $f$ solves the ODE: $f = \lambda f' + g$.

Its solution is given by

$$f(x) = \int_x^\infty dy \, \frac{1}{\lambda} e^{-\frac{(y-x)}{\lambda}} g(y),$$

where we used the fact that $f(\infty) = 0$.

Then,

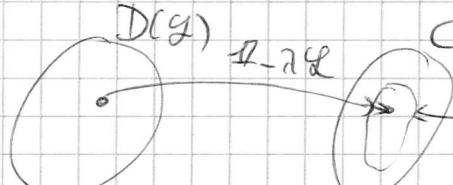
$$f(x) = \int_x^\infty dy \, \frac{1}{\lambda} e^{-\frac{(y-x)}{\lambda}} g(y)$$

$$\overset{z := \frac{y-x}{\lambda}}{=} \int_0^\infty dz \, e^{-z} g(x + \lambda z)$$

$$\geq \int_0^\infty dz \, e^{-z} \inf_{x \in \mathbb{R}} g(x + \lambda z)$$

$$= \inf_{x \in \mathbb{R}} g(x). \Rightarrow \inf_{x \in \mathbb{R}} f(x) \geq \inf_{x \in \mathbb{R}} g(x). \checkmark$$

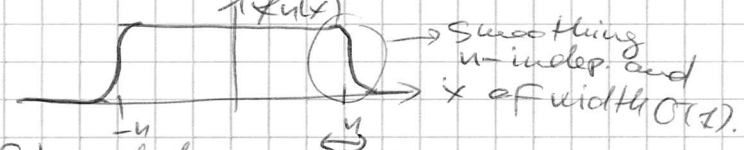
(c) Let $f \in \mathcal{D}(\mathcal{L}) \Rightarrow f - \lambda f' \in C(\mathbb{R}) \Rightarrow R(\mathbb{1} - \lambda\mathcal{L}) \subseteq C(\mathbb{R})$.




But for any $g \in C(\mathbb{R})$, $f(x) = -\int_x^\infty dy \, \frac{1}{\lambda} e^{-\frac{(y-x)}{\lambda}} g(y)$

belongs to $\mathcal{D}(\mathcal{L}) \Rightarrow C(\mathbb{R}) \subseteq R(\mathbb{1} - \lambda\mathcal{L})$. ✓

(d) It is enough to take




$\Rightarrow f_n \to 1$ pointwise, $g_n$ is unif. bounded and

$g_n \to 1$ pointwise since $f_n' \to 0$ for $x$ away $O(1)$ from $\pm n$. ✓

## 2.5) Lévy processes.

. Lévy processes are the most general Feller processes on $\mathbb{R}$ with stationary independent increments.

. Let us describe them.

. let us set

$$\psi(u) := i\beta u - \frac{\sigma^2 u^2}{2} + \int_{\mathbb{R}} \nu(dx) \left( e^{iux} - 1 - iux \mathbb{1}_{|x|\leq 1} \right)$$

with parameters: $\beta \in \mathbb{R}$, $\sigma \geq 0$,

and $\nu$ a measure on $\mathbb{R}^*$ satisfying

$$\int_{\mathbb{R}} \nu(dx) \min\{|x|^2, 1\} < \infty$$

. Lemma 2.8) $\exp(\psi(u))$ is a characteristic function.

. Proof: . let us start with $\nu = 0$. Then

$$\exp(\psi(u)) = \exp\left(i\beta u - \frac{\sigma^2 u^2}{2}\right) \quad \left(= \mathbb{E}(e^{iu\eta})\right).$$

is characteristic function of a random variable

$\eta \sim \mathcal{N}(\beta, \sigma^2)$.

. Next consider a compound Poisson random variable $Z$, i.e.,

$$Z = \sum_{k=1}^{N(\lambda)} Y_k$$

with $Y_1, Y_2, \ldots$ iid random variables with distribution $\mu$, and $N(\lambda) \sim \text{Poi}(\lambda)$ and independent of the $Y_i$'s.

Then, $\mathbb{E}(e^{iuZ}) = \sum_{\ell \geq 0} \mathbb{E}(e^{iuZ} | N(\lambda) = \ell) \cdot \frac{\lambda^\ell e^{-\lambda}}{\ell!}$

$$= \sum_{e \geq 0} \frac{\lambda^e \cdot e^{-\lambda}}{e!} \cdot \left( \int_{\mathbb{R}} \mu(dx)\, e^{iux} \right)^e$$

$$= \exp\left( \lambda \cdot \int_{\mathbb{R}} \mu(dx) \left( e^{iux} - 1 \right) \right)$$

$$= \exp(\psi(u)) \quad \text{for} \quad \nu = \lambda \cdot \mu,\ \gamma = 0,\ \beta = \lambda \int_{[-1,1]} \mu(dx)\, x$$

$\Rightarrow$ If $\nu$ is finite, $e^{\psi(u)}$ is the characteristic function of $a\eta + N(0, \sigma^2)$ and $Z$, where $\eta$ and $Z$ are taken to be independent.

- $\nu$ could be not finite, for instance if $\nu(dx) \sim \frac{1}{x^2}$ around $x = 0$. This case can be obtained by a limiting procedure. #

**Remarks:** The associated semigroup of a Lévy process is given by

$$T(t) f(x) = E\left( f(x + \xi_t) \right)$$

where $\xi_t$ is a random variable with characteristic function $\exp(t\psi(u))$.

7.11.13

## 2.6) From Feller process to the infinitesimal description.

**Thm 2.9)** Given a Feller process, define $T$ by

$$T(t)f(x) := \mathbb{E}^x[f(X(t))]$$

for $f \in C(S)$. Then, $T(t)$ is a probability semigroup on $C(S)$.

**Proof:** (a) $T(0)f_1(x) = f(x)$, $\forall x \in S \Rightarrow T(0)f = f$.

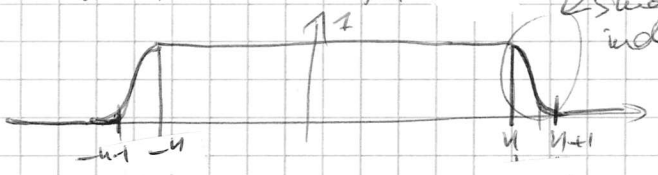
(c) It follows by the Markov property:

$$\begin{aligned} T(s+t)f(x) &= \mathbb{E}^x[f(X(t+s))] \\ &= \mathbb{E}^x\left[\mathbb{E}^x[f(X(t+s))\,|\,\mathcal{F}_s]\right] \\ &\overset{MP}{=} \mathbb{E}^x\left[\mathbb{E}^{X(s)}[f(X(t))]\right] \\ &= \mathbb{E}^x[T(t)f(X(s))] \\ &= T(s)\,T(t)f(x). \end{aligned}$$

(d) Obvious.

(e) For $S$ compact: $T(t)1(x) = \mathbb{E}^x(1) = 1(x)$ ✓.

For $S$ non-compact, take $f_n \in C(S)$ as, ← smoothing indep. of $n$.




$\Rightarrow$ • $\sup_n \|f_n\| \leq 1 < \infty$ ✓.

• $f_n \to 1$ pointwise

• $T(t)f_n(x) = \mathbb{E}^x[f_n(X(t))] \longrightarrow 1$ pointwise

since $\mathbb{P}^x$ is a probability measure (no mass is lost at $\infty$) $\forall t \geq 0$ and $\begin{cases} f_n = 1 \text{ on } [-n,n] \\ f_n = 0 \text{ on } [-n-1,n+1]^c. \end{cases}$

(b) To see this property we need to work a bit more.

. For fixed $x$, $T(t)f(x) - f(x) = \mathbb{E}^x\left(f(X(t)) - f(x)\right) \xrightarrow[t \downarrow 0]{} 0$
because $X$ is right continuous and $f$ is continuous.

We however need to show that

$$\lim_{t \downarrow 0} \sup_{x \in S} |T(t)f(x) - f(x)| = 0.$$

. let us first show that $U(\alpha)f(x) - U(\beta)f(x) = (\beta - \alpha)U(\alpha)U(\beta)f(x)$.

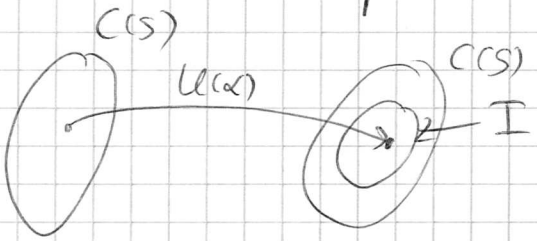
Remind $U(\alpha)f(x) := \int_0^\infty dt\, e^{-\alpha t} T(t)f(x)$

$$\Rightarrow U(\alpha)U(\beta)f(x) = \int_0^\infty dt\, e^{-\alpha t} T(t) \int_0^\infty ds\, e^{-\beta s} \underbrace{T(s)f(x)}_{\overset{\text{def}}{=} \mathbb{E}^x(f(X(s)))}$$

. It is right-cont. in $s$, $\forall x \in S$
. It is cont. in $x$, $\forall s$ (from Feller property)
↳ it is jointly measurable in $(x, s)$
. The same for $T(t)$ applied to $e^{-\beta s} T(s)f(x)$
$\Rightarrow$ can exchange $T(t)$ and $\int_0^\infty ds$.

$= \dots =$ (as in proof of Lemma 2.5) $\dfrac{U(\alpha)f(x) - U(\beta)f(x)}{\beta - \alpha}$.

. As $U(\alpha): C(S) \to C(S)$, let $I = \mathcal{R}(U(\alpha))$.




By the resolvent eq. $\Rightarrow U(\beta) = U(\alpha)\left(1 - (\beta - \alpha)U(\beta)\right)$
we get that $\mathcal{R}(U(\alpha))$ is independent of $\alpha$.

. Now, consider $f = U(\alpha) g \in I$, $g \in C(S)$.

Then, $T(t) f = \int_0^\infty ds\, e^{-\alpha s} T(s+t) g$

$$\overset{s+t=u}{=} \int_t^\infty du\, e^{-\alpha(u-t)} T(u) g$$

$$= e^{\alpha t} \Big[ \underbrace{\int_0^\infty du\, e^{-\alpha u} T(u) g}_{= U(\alpha) g} - \underbrace{\int_0^t du\, e^{-\alpha u} T(u) g}_{\substack{t \downarrow 0 \to 0 \\ \text{uniformly}}} \Big]$$

$$\Rightarrow \| T(t) f - f \| \xrightarrow{t \downarrow 0} 0 .$$

. We would like to have this $\forall f \in C(S)$.

. Let $(f_n)_{n \in \mathbb{N}}$ a sequence in $I$ and denote by

$f = \lim_{n \to \infty} f_n$ (strong limit), i.e., $f \in \overline{I}_s$ (strong closure, i.e., $\| f - f_n \| \xrightarrow[n \to \infty]{} 0$).

$$\Rightarrow \| T(t) f - f \| \le \| T(t) f - T(t) f_n \| + \| T(t) f_n - f_n \| + \| f_n - f \|$$

$$\le \underbrace{2 \cdot \| f_n - f \|}_{\to 0 \text{ as } n \to \infty} + \underbrace{\| T(t) f_n - f_n \|}_{\to 0 \text{ as } t \downarrow 0} \xrightarrow[n \to \infty]{t \downarrow 0} 0 .$$

($T(t)$ is a contraction (for this we did not need property (b)))

. Further, $\lim_{\alpha \to \infty} \alpha U(\alpha) f(x) = f(x)$ pointwise

$\Rightarrow$ the weak closure of $I$, $\overline{I}_w$, is equal $C(S)$

(the weak topology is the weakest that makes all elements of $S^*$ (the dual of $S$ = space of bounded linear fct.) continuous).

. So, if $\overline{I}_w = \overline{I}_s \Rightarrow$ we are done.

This is Corollary A.7 of [Liggett]: $\overline{I}_s = \overline{I}_w$ if $I$ is a linear subset of a Banach space. ∎

. The next step is the connection between a semigroup and a generator.

**Thm. 2.10)** Let $T(t)$ be a probability semigroup and define $\mathcal{L}$ by

$$\mathcal{L}f = \lim_{t \downarrow 0} \frac{T(t)f - f}{t} \qquad (*)$$

on $\mathcal{D}(\mathcal{L}) := \{f \in C(S) \mid \text{the strong limit in } (*) \text{ exists}\}$.

Then, $\mathcal{L}$ is a probability generator.

. It has the following properties:

(a) $\forall g \in C(S)$, $\alpha > 0$,

$f = \alpha U(\alpha) g \iff f \in \mathcal{D}(\mathcal{L})$ and satisfies $f - \alpha^{-1}\mathcal{L}f = g$

(b) If $f \in \mathcal{D}(\mathcal{L}) \Rightarrow T(t)f \in \mathcal{D}(\mathcal{L})$ for all $t \geq 0$, is a continuously differentiable function of $t$ with

$$\frac{d}{dt} T(t)f = T(t)\mathcal{L}f = \mathcal{L}T(t)f.$$

(c) For $f \in C(S)$, $t > 0$:

$$\lim_{n \to \infty} \left(1 - \frac{t}{n}\mathcal{L}\right)^{-n} f = T(t)f.$$

**Rem:** Formally $T(t) = \exp(t\mathcal{L})$. For $\mathcal{L}$ bounded, this can be defined as:

$$\sum_{n \geq 0} \frac{t^n}{n!}\mathcal{L}^n \quad \text{or} \quad \lim_{n \to \infty} \left(1 + \frac{t}{n}\mathcal{L}\right)^n \quad \text{or} \quad \lim_{n \to \infty} \left(1 - \frac{t}{n}\mathcal{L}\right)^{-n}.$$

For $\mathcal{L}$ unbounded, however, only the latter makes sense! (For instance, $\left\| \sum_{n \geq 0} \frac{t^n}{n!}\mathcal{L}^n \right\| \leq \sum_{n \geq 0} \frac{t^n}{n!}\|\mathcal{L}\|^n \Rightarrow \|e^{t\mathcal{L}}\| \leq e^{t\|\mathcal{L}\|}$).

. **Reason:** As $T(t) = e^{t\mathcal{L}}$ is a contraction, and "$\mathcal{L}1 = 0$", the eigenvalues have real part $\leq 0 \Rightarrow$ $\frac{1}{\quad}$ behaves well,

**Remind:**

**Def. 2.6:**

ⓐ $\mathcal{D}(\mathcal{G})$ dense in $C(S)$

ⓑ $f \in \mathcal{D}(\mathcal{G})$, $\lambda > 0$, $f - \lambda \mathcal{G} f = g$
$\Rightarrow \inf_{x \in S} f(x) \geq \inf_{x \in S} g(x)$

ⓒ $R(I - \lambda \mathcal{G}) = C(S)$, $\lambda < 1$

ⓓ For $\lambda < 1$, $\exists f_n \in \mathcal{D}(\mathcal{G})$ s.t. $g_n = f_n - \lambda \mathcal{G} f_n$ satisfies $\sup_n \|g_n\| < \infty$, $f_n, g_n \to 1$ p.t.wise.

**Proof:** Consider $f = \alpha U(\alpha) g$ for some $\alpha > 0$ and $g \in C(S)$, i.e.,

$$f = \alpha \int_0^\infty ds\, e^{-\alpha s} T(s) g.$$

$$\Rightarrow \frac{T(t) f - f}{t} = \alpha \frac{e^{\alpha t} - 1}{t} \int_0^\infty ds\, e^{-\alpha s} T(s) g$$

(as in Proof of the previous Thm (page 39))

$$- \frac{\alpha}{t} \int_0^t ds\, e^{-\alpha s} T(s) g$$

$$\left(T(s) g \xrightarrow{s \to 0} g\right) \xrightarrow[t \downarrow 0]{} \alpha^2 U(\alpha) g - \alpha g = \alpha (f - g).$$

$\Rightarrow$ Property ⓑ $\Rightarrow$ ✓.

Property ⓒ of Def. 2.6 ✓.

By the fact that $\alpha U(\alpha) g \in \mathcal{D}(\mathcal{G})$ and $\alpha U(\alpha) g \xrightarrow[\alpha \to \infty]{} g$, $\mathcal{D}(\mathcal{G})$ is dense in $C(S)$

$\Rightarrow$ Property ⓐ of Def. 2.6. ✓.

Next let us show Prop. ⓓ of Def. 2.6:

For $t > 0$, $f \in \mathcal{D}(\mathcal{G})$, set

$$g_t := \left(1 + \frac{\lambda}{t}\right) f - \frac{\lambda}{t} T(t) f = f - \lambda \frac{T(t) f - f}{t}.$$

Then, $\lim_{t \downarrow 0} g_t = f - \lambda \mathcal{G} f$ and

$$\left(1 + \frac{\lambda}{t}\right) \inf_{x \in S} f(x) \geq \inf_{x \in S} g_t(x) + \inf_{x \in S} \frac{\lambda}{t} T(t) f(x)$$

(Apply $T(t)$ to the positive function $f(x) - \inf_y f(y)$ and use property ⓐ of def. 2-3)

$$\geq \inf_{x \in S} g_t(x) + \inf_{x \in S} \frac{\lambda}{t} f(x)$$

We have obtained $\inf_{x \in S} f(x) \geq \inf_{x \in S} g_t(x) \xrightarrow[t \downarrow 0]{} \inf_{x \in S} (f(x) - \lambda \mathcal{G} f(x))$ by

$\Rightarrow$ Property ⓑ of Def. 2.6 ✓.

. For some $f \in D(\mathcal{G})$, $\alpha > 0$, let $g = f - \alpha^{-1}\mathcal{G}f$.

. By property @⇒, $h = \alpha U(\alpha) g$ satisfies $h - \alpha^{-1}\mathcal{G}h = g$

$\Rightarrow f = h$, by the fact that then $(f-h) - \alpha^{-1}\mathcal{G}(f-h) = 0$

and by Lemma 2.7, $\|f-h\| \le \|0\| = 0$.

⇒ Property @ is completed. ✓

. To show ⓓ of Def. 2.6; let $S$ be non-compact (for compact it is easier as usual).

. Assume $g_n \in C(S)$ satisfying $\sup_n \|g_n\| < \infty$, $g_n$ and $T(t)g_n \to 1$ pointwise for each $t \ge 0$ (compare with the def. 2.3ⓔ).

Then, set

$$f_n \in D(\mathcal{G}) \quad \text{by} \quad g_n = f_n - \alpha \mathcal{G} f_n .$$

By property @, $f_n = \alpha U(\alpha) g_n$.

As $f_n(x) = \alpha \int_0^\infty ds\, e^{-\alpha s} T(s) g_n(x)$ and $T(s) g_n(x) \to 1$ pt-wise.

$\Rightarrow f_n(x) \to 1$ pt-wise (use bounded convergence to take the limit inside the integral).

⇒ Property ⓓ of Def. 2.6 : ✓.

. Now we show Property ⓑ :

$$\frac{d}{dt} T(t) f = \lim_{s \downarrow 0} \frac{T(t+s)f - T(t)f}{s}$$

$$= \lim_{s \downarrow 0} T(t) \left( \frac{T(s)f - f}{s} \right)$$

$$= \lim_{s \downarrow 0} \frac{T(s)\left(T(t)f\right) - T(t)f}{s},$$

If any of these limits exists, then they are the same and all other exists too.

. Since $f \in \mathcal{D}(\mathcal{G})$ and $T(t)$ is a contraction,

$$\left\| T(t)\left(\frac{T(s)f - f}{s}\right)\right\| \leq \left\| \frac{T(s)f - f}{s}\right\|$$

So that the middle expression exists.

$\Rightarrow$ The 3rd limit exists implies $T(t)f \in \mathcal{D}(\mathcal{G})$.

. Finally the second limit is continuous in $t \Rightarrow T(t)f$ is $C^1$ in $t$.

$\Rightarrow$ Property (b) ✓.

. To verify (c): $\underline{(\mathbb{1} - \alpha^{-1}\mathcal{G})^{-n} f = (\mathbb{1} - \alpha^{-1}\mathcal{G})^{-(n-1)} \underbrace{(\mathbb{1} - \alpha^{-1}\mathcal{G})^{-1} f}_{\text{by (a)} = R(\alpha) f}}$

$$= \alpha^n R(\alpha)^n f$$

$$\overset{\text{def. of } R(\alpha)}{=} \alpha^n \int_0^\infty ds_1\, e^{-\alpha s_1} \cdots \int_0^\infty ds_n\, e^{-\alpha s_n} T(s_1 + \cdots + s_n) f$$

$$= \int_0^\infty ds\, \frac{\alpha^n s^{n-1}}{(n-1)!} \cdot e^{-\alpha s} \cdot T(s) f$$

. Now, let $\xi_1, \xi_2, \ldots$ iid $\exp(1)$ random variables.

$$\Rightarrow (\mathbb{1} - \tfrac{t}{n}\mathcal{G})^{-n} f = \mathbb{E}\left( T\left( \frac{\xi_1 + \cdots + \xi_n}{n/t} \right) f \right).$$

($\alpha = \frac{n}{t}$)

. For $f \in \mathcal{D}(\mathcal{G})$, $\dfrac{\|T(t)f - T(s)f\|}{|t-s|} \leq \|\mathcal{G}f\|$.

Indeed, $\|T(t)f - T(s)f\| \leq \sum_{k=1}^{n} \left\| T\left(s + \tfrac{t-s}{n}\cdot k\right) f - T\left(s + \tfrac{t-s}{n}\cdot(k-1)\right) f \right\|$

(w.l.o.g. $t > s$)

$$\leq \sum_{k=1}^{n} \underbrace{\left\| T\left(\tfrac{t-s}{n}\right) f - f \right\|}_{\text{$T$ is a contraction}} \leq \|\mathcal{G}f\| \cdot \tfrac{t-s}{n}$$

by Property (b)

$$\xrightarrow{n \to \infty} \|\mathcal{G}f\|.$$

Therefore,

$$\|(\mathbb{1}-\tfrac{t}{n}\mathcal{G})^{-n}f - T(t)f\|$$

$$= \|\mathbb{E}\left(T\left(\tfrac{\xi_1+\dots+\xi_n}{n}t\right)f - T(t)f\right)\|$$

$$\leq t\cdot\|\mathcal{G}f\|\,\underbrace{\mathbb{E}\left(\left|\tfrac{\xi_1+\dots+\xi_n}{n}-1\right|\right)}_{\xrightarrow[n\to\infty]{} 0 \text{ by the law of large numbers.}}$$

$\Rightarrow$ Also property (c) holds. $\#$

12.11.2013

**Example:** Consider the uniform motion to the right on $\mathbb{R}$. It is a Feller process with

$$\mathbb{P}^x = \delta_{\omega:\omega(t)=x+t}$$

or, equivalently, $T(t)f(x) = f(t+x)$.

Then, the generator of this process is given by $\mathcal{G}f = f'$ on $\{f\in C(\mathbb{R}) \mid f'\in C(\mathbb{R})\}$.

(See exercises: what has to be proven is that this domain is the same as the one of Theorem 2.10).

**Example:** Take $S=\mathbb{R}$ and the generator defined by $\mathcal{G}f = \frac{1}{2}f''$ on

$$D(\mathcal{G}) = \{f\in C(\mathbb{R}) \mid f', f''\in C(\mathbb{R})\}.$$

Then, this is a probability generator with Brownian semigroup.

(See exercises).

**Remark:** The domain is important. For example, if $D(\mathcal{G}) = \{f\in C(\mathbb{R}) \mid f', f''\in C(\mathbb{R}_+),\ f'(0)=0\}$ is a B.M. reflected

## Example (Cauchy process).

- The Cauchy process is a process on $\mathbb{R}$ with stationary independent increments, having the density:

$$\mathbb{P}(X(t+s)-X(s)\in dx) = \frac{1}{\pi}\frac{t\,dx}{t^2+x^2} \quad (t>0).$$

- Let $\mathcal{D} = \{f\in C(\mathbb{R}) \mid f''\in C(\mathbb{R})\}$.

Then,

$$\frac{T(t)f(x)-f(x)}{t} = \frac{1}{\pi}\int_{\mathbb{R}} dy\ \frac{f(x+y)-f(x)}{t^2+y^2}.$$

- As $t\downarrow 0$, $\frac{1}{t^2+y^2}$ is largest around $y\simeq 0$

But ~~for~~ $f\in\mathcal{D}$, $f(x+y) = f(x)+yf'(x)+\underbrace{o(y^2)}_{\text{uniformly in } x}$

$\Rightarrow \frac{f(x+y)-f(x)}{y^2}$ is not integrable around $y=0$.

- However, since $\frac{1}{\pi}\int_{|y|\le 1} dy\ \frac{y}{t^2+y^2} = 0$, we can rewrite, for any $t>0$,

$$\frac{T(t)f(x)-f(x)}{t} = \frac{1}{\pi}\int_{\mathbb{R}} dy\ \frac{f(x+y)-f(x)-yf'(x)\mathbb{1}_{|y|\le 1}}{t^2+y^2}$$

and this has a nice limit as $t\downarrow 0$.

So, we obtain:

$$\mathcal{L}f(x) = \frac{1}{\pi}\int_{\mathbb{R}\setminus\{0\}} dy\ \frac{f(x+y)-f(x)-yf'(x)\mathbb{1}_{|y|\le 1}}{y^2}$$

$$\equiv \int_{\mathbb{R}\setminus\{0\}} \nu(dy)\left(f(x+y)-f(x)-yf'(x)\mathbb{1}_{|y|\le 1}\right)$$

with $\nu(dy) = \frac{dy}{\pi y^2}$.

Rem: One can replace $y\mathbb{1}_{|y|\le 1}$ with any odd fct. $\phi$ with compact support

Remark: . In the first two examples, the operators were local and have continuous paths, while for the Candlay process it is not local, and one can show that is does not have continuous paths.

. This is not an accident, as the connection between continuity of paths and locality of the generator is quite general.

Remark: . There are different ways of stating Thm 2.10.

. First, one can associate a transition kernel, $T_t$, to a semigroup $T(t)f(x) = \int_S T_{(t)}(x, dy)\, f(y)$

so that the resolvent writes

$$U(\alpha) f(x) = \int_0^\infty dt\, e^{-\alpha t}\, T(t) f(x)$$

$$= \int_S U_\alpha(x, dy)\, f(y)$$

where $U_{(\alpha)}(x, A) \equiv \int_0^\infty dt\, e^{-\alpha t}\, T_{(t)}(x, A)$ : resolvent kernel.

($x \in S$)

. Formally, $U(\alpha) = \int_0^\infty dt\, e^{-\alpha t} e^{\mathcal{L} t} = (\alpha - \mathcal{L})^{-1}$

i.e., $\alpha U(\alpha) = (\mathbb{1} - \alpha^{-1}\mathcal{L})^{-1} \leftrightarrow (\mathbb{1} - \alpha^{-1}\mathcal{L})^{-1}\mathcal{L} = \alpha(\alpha U(\alpha) - \mathbb{1})$

Power series which as $\alpha \to \infty$ one recovers $\mathcal{L}$.

## 2.7) From the infinitesimal description to the Feller process.

- What we want to do is, given a probability generator $\mathcal{L}$ (defined on a domain $\mathcal{D}(\mathcal{L})$), first construct the associated semigroup and then from this the Feller process.

- The expression $e^{t\mathcal{L}} = \sum_{n\geq 0} \frac{t^n}{n!}\mathcal{L}^n$ is well-defined only for bounded operators, and as we have seen in our examples, generically $\mathcal{L}$ is unbounded (see the Brownian motion generator). Still it will be useful.

- Consider the approximation $\mathcal{L}_\varepsilon$ of $\mathcal{L}$ given by

$$\mathcal{L}_\varepsilon := \mathcal{L}(\mathbb{1} - \varepsilon\mathcal{L})^{-1}$$

for $\varepsilon$ small enough.

**Lemma 2.11)** For $\varepsilon$ small enough, $\mathcal{L}_\varepsilon$ is well-defined and it is a bounded operator with

$$\|\mathcal{L}_\varepsilon g\| \leq \frac{2}{\varepsilon}\|g\|.$$

**Proof:** By Def 2.6(c), for $\varepsilon$ small enough, $\mathcal{R}(\mathbb{1} - \varepsilon\mathcal{L}) = C(S)$

$\Rightarrow \mathcal{R}(\mathbb{1} - \varepsilon\mathcal{L})^{-1} = \mathcal{D}(\mathcal{L}) \Rightarrow \mathcal{L}_\varepsilon$ is well-defined.

- Further, for $f \in \mathcal{D}(\mathcal{L})$, $f - \varepsilon\mathcal{L}f = g \Leftrightarrow f = (\mathbb{1} - \varepsilon\mathcal{L})^{-1} g$

and $\|\mathcal{L}_\varepsilon g\| = \|\mathcal{L}(\mathbb{1} - \varepsilon\mathcal{L})^{-1} g\| = \|\mathcal{L}f\| \leq \frac{\|f\| + \|g\|}{\varepsilon}$.

But by Lemma 2.7, $\|f\| \leq \|g\|$

$\Rightarrow \|\mathcal{L}_\varepsilon g\| \leq \frac{2}{\varepsilon}\|g\|$. #

. Now, since $\mathcal{L}_\varepsilon$ is a bounded operator, we can define

$$T_\varepsilon(t) := e^{t\mathcal{L}_\varepsilon} = \sum_{n\geq 0} \frac{t^n \mathcal{L}_\varepsilon^n}{n!}$$

. An useful result is:

Lemma 2.12) If $\tilde{\mathcal{L}}$ is a bounded operator, then it satisfies $\mathcal{R}(\mathbb{1} - \lambda\tilde{\mathcal{L}}) = C(S)$ for $\lambda$ small enough.

Proof: For any given $g \in C(S)$, we have to find a $f$ s.t. $f - \lambda\tilde{\mathcal{L}} f = g$.
It is given by $f = \sum_{n\geq 0} \lambda^n \tilde{\mathcal{L}}^n g$
which converges if $\lambda \|\mathcal{L}\| < 1$. #

Lemma 2.13) @ For every $f \in C(S)$,

$$(\mathbb{1} - \varepsilon\mathcal{L})^{-1} f - \varepsilon\mathcal{L}_\varepsilon f = f.$$

(b) $\mathcal{L}_\varepsilon$ is a probability generator and $T_\varepsilon(t)$ is a probability semigroup with generator $\mathcal{L}_\varepsilon$ (as in Theorem 2-(a)).

Proof: exercises.

With the above ingredients we now want to prove the existence of a semigroup associated with $\mathcal{L}$.

Remark: . $\mathcal{L}$ and $(\mathbb{1} - \varepsilon\mathcal{L})^{-1}$ commutes. This follows from

Lemma 2.13@ : $(\mathbb{1} - \varepsilon\mathcal{L})^{-1} - \mathbb{1} = \varepsilon\mathcal{L}(\mathbb{1} - \varepsilon\mathcal{L})^{-1}$ on $f \in C(S)$

$$(\mathbb{1} - \varepsilon\mathcal{L})^{-1}(\mathbb{1} - \mathbb{1} + \varepsilon\mathcal{L}) = \varepsilon(\mathbb{1} - \varepsilon\mathcal{L})^{-1}\mathcal{L}.$$

. By Lemma 2.2 $(\mathbb{1} - \varepsilon\mathcal{L})^{-1}$ is a contraction.

14.11.2013

**Thm 2.14)** For $f \in C(S)$,

$$T(t)f := \lim_{\varepsilon \downarrow 0} T_\varepsilon(t) f$$

exists uniformly on bounded $t$ intervals.

It defines a probability semigroup with generator $\mathcal{L}$ in the sense of Thm. 2.10.

• **Proof:** • First of all, notice that

$$\mathcal{L}_\varepsilon \text{ and } \mathcal{L}_\delta$$

commute for $\varepsilon, \delta > 0$ (for $\varepsilon, \delta$ small enough).

Indeed, by Lemma 2.13②,

$$\mathcal{L}_\delta \mathcal{L}_\varepsilon f = \frac{(I - \delta\mathcal{L})^{-1}(I - \varepsilon\mathcal{L})^{-1} f - (I - \delta\mathcal{L})^{-1} f - (I - \varepsilon\mathcal{L})^{-1} f + f}{\varepsilon\delta}$$

$$= \mathcal{L}_\varepsilon \mathcal{L}_\delta f$$

provided that $(I - \delta\mathcal{L})^{-1}(I - \varepsilon\mathcal{L})^{-1} = (I - \varepsilon\mathcal{L})^{-1}(I - \delta\mathcal{L})^{-1}$.

This holds since

$$(I - \varepsilon\mathcal{L})^{-1}(I - \delta\mathcal{L})^{-1} f = g \iff f = g - (\varepsilon + \delta)\mathcal{L} g + \varepsilon\delta \mathcal{L}^2 g$$

which is symmetric in $\varepsilon, \delta$. $\Rightarrow$ $T_\varepsilon$ and $T_\delta$ commutes too, and they do with $\mathcal{L}_\varepsilon, \mathcal{L}_\delta$.

• Consider now $T_\varepsilon(t)f - T_\delta(t)f$. By integration by parts:

$$T_\varepsilon(t)f - T_\delta(t)f = \int_0^t ds \, \frac{d}{ds}\left(T_\varepsilon(s)\, T_\delta(t-s)\right) f$$

By Lemma 2.13(b) and Thm 2.10(b) $=$

$$= \int_0^t T_\varepsilon(s)\, T_\delta(t-s)\left(\mathcal{L}_\varepsilon - \mathcal{L}_\delta\right) f$$

$$\Rightarrow \| T_\varepsilon(t)f - T_\delta(t)f \| \le t \cdot \| \mathcal{L}_\varepsilon f - \mathcal{L}_\delta f \|. \quad (*)$$

(Contraction)

. Since $(\mathbb{1}-\varepsilon\mathcal{L})^{-1}f - f = \varepsilon(\mathbb{1}-\varepsilon\mathcal{L})^{-1}\mathcal{L}f$ for $f\in D(\mathcal{L})$.

by Lemma 2.7 (rewritten as $\|(\mathbb{1}-\lambda\mathcal{L})^{-1}g\| \le \|g\|$ for $g\in C(S)$)
↳ $(\mathbb{1}-\lambda\mathcal{L})^{-1}$ is a contraction!

we get $\|(\mathbb{1}-\varepsilon\mathcal{L})^{-1}f - f\| \le \varepsilon\|\mathcal{L}f\|$.

$\Rightarrow \lim_{\varepsilon\downarrow 0}(\mathbb{1}-\varepsilon\mathcal{L})^{-1}f = f$, for $f\in D(\mathcal{L})$. (**)

. As $D(\mathcal{L})$ is, by def., dense in $C(S)$ and $(\mathbb{1}-\varepsilon\mathcal{L})^{-1}$ is a contraction $\Rightarrow \lim_{\varepsilon\downarrow 0}(\mathbb{1}-\varepsilon\mathcal{L})^{-1}f = f$ for all $f\in C(S)$ (same computation as in the proof of Thm. 2.9).

. Because $\mathcal{L}_\varepsilon f = (\mathbb{1}-\varepsilon\mathcal{L})^{-1}\mathcal{L}f$, $f\in D(\mathcal{L})$ with $g=\mathcal{L}f\in C(S)$,

$\Rightarrow \lim_{\varepsilon\downarrow 0}\mathcal{L}_\varepsilon f = \lim_{\varepsilon\downarrow 0}(\mathbb{1}-\varepsilon\mathcal{L})^{-1}g \overset{(**)}{=} g = \mathcal{L}f$.

. Then, by (*), $\lim_{\varepsilon\downarrow 0} T_\varepsilon(t)f$ exists uniformly on bounded time intervals for $f\in D(\mathcal{L})$.

. Finally, again by {$D(\mathcal{L})$ dense in $C(S)$} ∪ {$T_\varepsilon(t)$ contraction}
$\Rightarrow$ the same holds $\forall f\in C(S)$ and $T(t)$ satisfies the semigroup property (Prop. ⓒ of Def. 3.4).

---

. We still have other properties to be checked.

. Fix a $\lambda>0$ small enough s.t. $R(\mathbb{1}-\lambda\mathcal{L}) = C(S)$ and choose $\alpha := \frac{1}{\lambda}$.

For $g\in C(S)$, let $f_\varepsilon$ given by $f_\varepsilon - \lambda\mathcal{L}_\varepsilon f_\varepsilon = g$ and $f_\varepsilon = \alpha U_\varepsilon(\alpha)g \equiv \alpha\int_0^\infty dt\, e^{-\alpha t}T_\varepsilon(t)g$.
(equivalently by)

• Define $h_\varepsilon := (\mathbb{1} - \varepsilon \mathcal{L})^{-1} f_\varepsilon$. Then, by

$$T(t) f = \lim_{\varepsilon \downarrow 0} T_\varepsilon(t) f ,$$

$$f := \lim_{\varepsilon \downarrow 0} f_\varepsilon = \alpha \int_0^\infty dt\, e^{-\alpha t} T(t) g .$$

But by Lemma 2.13 @,

$$\lim_{\varepsilon \downarrow 0} h_\varepsilon = \lim_{\varepsilon \downarrow 0} (\mathbb{1} - \varepsilon \mathcal{L})^{-1} f_\varepsilon$$

$$= \lim_{\varepsilon \downarrow 0} \Big( f_\varepsilon + \varepsilon \underbrace{\mathcal{L}_\varepsilon f_\varepsilon}_{= \frac{g - f_\varepsilon}{\lambda}} \Big) = f \quad \text{too.}$$

• We now choose $h \in D(\mathcal{L})$ s.t.

$$h - \lambda \mathcal{L} h = g .$$

Then, $\mathcal{L} h_\varepsilon = \mathcal{L}_\varepsilon f_\varepsilon \underset{\varepsilon \downarrow 0}{\longrightarrow} \alpha (f - g)$

$$\Rightarrow \lim_{\varepsilon \downarrow 0} \big( (h_\varepsilon - h) - \lambda \mathcal{L} (h_\varepsilon - h) \big) = 0$$

By Lemma 2.7 $\Rightarrow h_\varepsilon - h \underset{\varepsilon \downarrow 0}{\to} 0$, i.e., $f = h \in D(\mathcal{L})$ and

$$\mathcal{L} f = \alpha (f - g) .$$

Thus: $(\mathbb{1} - \lambda \mathcal{L})^{-1} g = \alpha \int_0^\infty dt\, e^{-\alpha t} T(t) g .$

—

• To see the last property of Def. 3.4, Take $g_n$ as in by Def 2.6 @. Then, for each $x \in S$,

$$\lim_{n \to \infty} \underbrace{\alpha \int_0^\infty dt\, e^{-\alpha t} (T(t) g_n)(x)}_{= f_n(x)} = 1 .$$

$g_n$ as in
$\Downarrow$
$g_n = f_n - \lambda \mathcal{L} f_n$
$\Downarrow$
$f_n = \alpha R(\alpha) g_n$

• Each linear operator can be written as:

$$T(t) f(x) = \int_S f(y)\, d\mu_{t,x}(y) \quad \text{for some finite, positive measure } \mu_{t,x}.$$

(by Riesz representation theorem).

$\Rightarrow$ Since $g_n(x) \to 1$ (pointwise by Def. 2.6 @),

$$\alpha \int_0^\infty dt\, e^{-\alpha t} \mu_{t,x}(S) = 1 \quad \text{for all } x \in S .$$

. Now, $\mu_{t,x}$ is continuous in $t,x \Rightarrow \mu_{t,x}(S) = 1$

$\Rightarrow T(t)$ satisfies the last property of Def. 3.4. (i.e., $T(t)f(x) \to 1$ otwise).

. Finally, let us see that this semigroup has generator $\mathcal{L}$.

. By $\frac{d}{dt} T(t)f = T(t)\mathcal{L}f = \mathcal{L}T(t)f$, we have

$$T_\varepsilon(t)f - f = \int_0^t \frac{d}{ds} T_\varepsilon(s) f \, ds = \int_0^t ds \, T_\varepsilon(s) \mathcal{L}_\varepsilon f$$

. For $f \in D(\mathcal{L})$, by $T(t)f = \lim_{\varepsilon \downarrow 0} T_\varepsilon(t)f$, $\lim_{\varepsilon \downarrow 0} \mathcal{L}_\varepsilon f = \mathcal{L}f$ and the fact that $T_\varepsilon(t)$ is a contraction, we get

$$T(t)f - f = \int_0^t ds \, T(s) \mathcal{L} f.$$

$\Rightarrow$ For $f \in D(\mathcal{L})$, $\lim_{t \downarrow 0} \frac{T(t)f - f}{t} = \mathcal{L}f$, i.e., the generator defined in Thm. 2.10 is an extension of $\mathcal{L}$ ( the domain in Thm 2.10 might be larger than the $D(\mathcal{L})$ given here).
By Def. 2.6 (c), however, it is not the case. #

. Next we want to see that given a semigroup we can construct a Feller process, i.e., we need to construct the family of $\{\mathbb{P}^x, x \in S\}$ and also check the properties like the Markov property or the fact that there is a version of the process which is càdlàg.

19.11.'13 ↑

• Assume that there is a càdlàg version and let us the semigroup be given. We want to see that $\mathbb{E}^x(f(X(t))) := T(t)f(x)$ has then the Markov property.

Lemma 2.15). Given a semigroup $\{T(t), t\geq 0\}$ and $\mathbb{P}^x$ s.t. $\mathbb{E}^x(f(X(t))) = T(t)f(x)$, if $X$ is càdlàg, then the Markov property holds.

• Proof (sketch): The proof is very close to the one that one does to show it for the Brownian motion by the transition kernel satisfying the Chapman–Kolmogorov equations. Here is how it works:

We need to show that

$$\mathbb{E}^x(Y\circ\theta_s|\mathcal{F}_s) = \mathbb{E}^{X(s)}(Y), \quad \mathbb{P}^x\text{-a.s.},$$

$\forall x\in S$, $Y$ bounded measurable function on $\Omega$.

Equivalently, $\forall A\in\mathcal{F}_s$,

$$\mathbb{E}^x(Y\circ\theta_s \mathbb{1}_A) = \mathbb{E}^x(\mathbb{E}^{X(s)}(Y)\mathbb{1}_A) \qquad (**)$$

(a) First consider $Y$ of a special form,

$$Y(\omega) = \prod_{m=1}^{n} f_m(w(t_m))$$

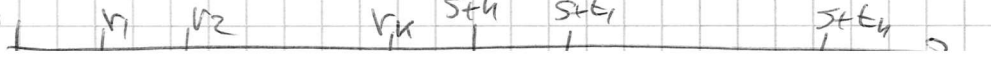
for $0<t_1<t_2<\dots<t_n$, $f_m\in C(S)$.

Also, consider $A\in\mathcal{F}$ finite-dimensional, i.e.,

$$A = \{w : w(r_1)\in A_1, \dots, w(r_k)\in A_k\},$$

for $0<r_1<\dots<r_k$, $A_i\in\mathcal{S}$ open.

• Further, let $h>0$ small (later $h\downarrow 0$) s.t.:



(b) Define $\phi(y,h) := \mathbb{E}^y\left[f_1(X(t_1-h)) \cdots f_k(X(t_k-h))\right]$.

Then, $\mathbb{E}^x(Y \circ \theta_s \mathbb{1}_A) = \mathbb{E}^x\left[\phi(X(s+h), h)\, \mathbb{1}_A\right]$ (*)

and it is jointly continuous for $(y,h) \in S \times [0, t_1)$.

For example, when $k=n=1$, $\phi(y,h) = \mathbb{E}^y(f_1(X(t_1-h)))$

$= T(t_1-h) f_1(y)$

and $\mathbb{E}^x(Y \circ \theta_s \mathbb{1}_A) = \mathbb{E}^x\left[f_1(\omega(s+t_1))\, \mathbb{1}_{\omega(r_1) \in A_1}\right]$

$= \int_{A_1} dy\, T(r_1)(x,y) \int_S dv\, T(s+h-r_1)(y,v) \underbrace{\int_S dz\, T(t_1-h)(v,z) \cdot f_1(z)}_{\equiv \phi(v,h)}$

(Kernel)

$= \mathbb{E}^x(\phi(X(s+h), h)\, \mathbb{1}_A)$.

- At this point one apply the $\pi$-$\lambda$ Theorem (see Thm A.3 in the Book)

with $\mathcal{P}$ = finite-dim. sets in $\mathcal{F}^o_{s+\frac{h}{2}}$ ($\mathcal{F}^o_t$ = smallest $\sigma$-algebra s.t. $\omega \to \omega(s)$ is measurable $\forall s \le t$; $\mathcal{F}_t = \bigcap_{s>t} \mathcal{F}^o_s$ : right-cont.)

and $\mathcal{L} = \{A \in \mathcal{F}^o_{s+\frac{h}{2}} : (*) \text{ holds}\}$.

- Since $\mathcal{F}_s \subset \mathcal{F}^o_{s+\frac{h}{2}} \Rightarrow (*)$ holds $\forall A \in \mathcal{F}_s$ and $Y$ of the special form.

- Take now $h \downarrow 0$ in (*): Since $\phi(y,0) = \mathbb{E}^y(Y)$ (and l.h.s. of (*) is indep. of $h$)

$\Rightarrow (**)$ holds for any $A \in \mathcal{F}_s$, $Y$ of the special form.

- For general $Y$, one uses the monotone class Thm. (See Thm A.4 in the Book) applied to $\mathcal{P}$ = finite-dim sets and

$\mathcal{H} = \{\text{bounded r.v. } Y : (**) \text{ holds } \forall A \in \mathcal{F}_s\}$. #

. Once we will have shown that there is a càdlàg version of the process with given semigroup, by the Markov property, its finite dimensional distributions are determined:

-For example, for $t_1 < t_2$,

$$\mathbb{E}^x\left[f(X(t_1))\, g(X(t_2))\right] = \mathbb{E}^x\left[f(X(t_1))\, \mathbb{E}^{X(t_1)}\left(g(X(t_2 - t_1))\right)\right]$$

$$= \mathbb{E}^x\left[f(X(t_1)) \cdot T(t_2 - t_1)\, g(X(t_1))\right]$$

$$= T(t_1)\, h(x) \quad \text{where } h(x) = f(x)\, T(t_2 - t_1)\, g(x).$$

. **Thm 2.16)** If $T(t)$ is a probability semigroup, then there is a Feller process $X(t)$ satisfying

$$\mathbb{E}^x(f(X(t))) = T(t) f(x)$$

for $x \in S$, $t \geq 0$ and $f \in C(S)$.

. For the proof we will use a supermartingale associated with the resolvent, and apply the following two Lemmas (stated without proof).

**Lemma 2.17)** Suppose that $(M(t), t \in \mathbb{R}_+)$ is a uniformly bounded supermartingale. Then, with probability one,

$$\lim_{s \uparrow t,\, s \in \mathbb{Q}} M(s) \text{ exists } \forall\, t > 0,$$

and

$$\lim_{s \downarrow t,\, s \in \mathbb{Q}} M(s) \text{ exists } \forall\, t \geq 0.$$

**Rem.:** The idea of the proof is to control the upcrossings as one does in the proof of the convergence of submartingale.

**Lemma 2.18)** Let $(M(t), t \in \mathbb{Q}_+)$ be a non-negative supermartingale. Then, $\forall t \in \mathbb{Q}_+$,

$$\mathbb{P}\left(M(t) > 0, \inf_{s \in \mathbb{Q} \cap [0,t]} M(s) = 0\right) = 0.$$

*Proof:* First consider $0 \le s_1 < \dots < s_n \le t$, $s_i \in \mathbb{Q}_+$ and $\varepsilon > 0$.

Set $A_k = \{ M(s_k) < \varepsilon, M(s_i) \ge \varepsilon \text{ for } 1 \le i \le k \}$, $1 \le k \le n$.

Then, $\mathbb{E}(M(t) \mathbb{1}_{A_k}) \le \mathbb{E}(M(s_k) \mathbb{1}_{A_k}) \le \varepsilon \cdot \mathbb{P}(A_k)$.

(supermartingale)

As $A_k$'s are disjoints, $\sum_{k=1}^{n}$ leads to:

$$\mathbb{E}\left(M(t), \min_{1 \le i \le n} M(s_i) < \varepsilon\right) \le \varepsilon \, \mathbb{P}\left(\min_{1 \le i \le n} M(s_i) < \varepsilon\right)$$

. Enumerating the rational in $[0,t]$ and take the limit $n \to \infty$ gives:

$$\mathbb{E}\left(M(t), \inf_{s \in \mathbb{Q} \cap [0,t]} M(s) < \varepsilon\right) \le \varepsilon \cdot \mathbb{P}\left(\inf_{s \in \mathbb{Q} \cap [0,t]} M(s) < \varepsilon\right)$$

$$\xrightarrow{\varepsilon \to 0} 0. \qquad \#$$

---

Now we have the ingredients to prove Thm 2.16.

*Proof of Thm 2.16).* We will use Kolmogorov-Doeblin theorem that ensures the existence of a process $(Y(t), t \in \mathbb{R}_+)$, $\forall x \in S$, with the desired finite-dim. distr. and $Y(0) = x$.

. Take any $f \in C(S)$ non-negative. Then,

$$e^{-\alpha t} T(t) U(\alpha) f = \int_t^\infty ds \, e^{-\alpha s} T(s) f \le U(\alpha) f.$$

So, $\forall t \in \mathbb{R}^+$, $\mathbb{E}\left[e^{-\alpha t} T(t) U(\alpha) f(Y(t))\right] \le U(\alpha) f(Y(0))$

Therefore, $e^{-\alpha t} U(\alpha) f(Y(t))$ is a bounded supermartingale with respect to the filtration $(\mathcal{F}_t, t \in \mathbb{Q})$, where $\mathcal{F}_t = \sigma(Y(s), s \in \mathbb{Q} \cap [0,t])$.

By Lemma 2.17, right / left limits along rationals exist $\forall t$ with probability 1. The set where this fails depends generically on $f$ and $\alpha$.

⇒ We need to work with sets of $f$ and $\alpha$ which are countable, but large enough to recover $C(S)$.

- Take $f$ in a countable dense set $\tilde{S} \subset C(S)$ and $\alpha \in \mathbb{N}$. Then, by $\alpha U(\alpha) f \to f$ as $\alpha \to \infty$ and since $C(S)$ is separable, we use the functions $U(\alpha) f$, $\alpha \in \mathbb{N}$, $f \in \tilde{S}$.

- Take $f \in C(S)$ strictly positive and consider the supermartingale $M(t) = e^{-t} U(1) f(Y(t))$. This is also strictly positive and by Lemma 2.18

$$\Rightarrow \inf_{s \in \mathbb{Q} \cap [0,t]} M(s) > 0 \quad \text{a.s.}, \ \forall t > 0.$$

- From $U(1) f \in C(S) \Rightarrow \{Y(s), s \in \mathbb{Q} \cap [0,t]\}$ is a.s. bounded $\forall t > 0$.

- Thus, with probability one, $Y(s)$ has right and left limits $\forall t \geq 0$. Therefore, define:

$$X(t) := \lim_{s \downarrow t, s \in \mathbb{Q}} Y(s).$$

Then, this is right-cont. and has left limits. But the finite-dim. distr. corresponding to $T(t)$ are weakly continuous ⇒ $X(t)$ has the right finite-dim. distr. ∎

There is one more property:

**Def. 2.19)** The process $X(t)$ is called *quasi-left continuous* if whenever a sequence of stopping times $\tau_n \nearrow \tau$, it follows that

$$X(\tau_n) \to X(\tau) \quad \text{a.s. on } \{\tau < \infty\}.$$

**Prop. 2.20)** The process $X(t)$ of Thm 2.16 is quasi-left continuous.

**Proof:** By considering $\tau_n \wedge m$, we may w.l.o.g. assume that $\tau$ is bounded (as we need to check the property on $\{\tau < \infty\}$).

$\therefore L(t) = \lim_{n\to\infty} X(\tau_n + t)$ exists $\forall t > 0$, since $X(t)$ has left limits.

$\lim_{t \downarrow 0} L(t) = X(\tau)$ since $X(t)$ is right-continuous.

$\Rightarrow \forall f, g \in C(S)$:

$$\mathbb{E}^x\left[f(L(0))\, g(X(\tau))\right] =$$

$$= \lim_{t\downarrow 0} \lim_{n\to\infty} \mathbb{E}^x\left[ f(X(\tau_n)) \cdot g(X(\tau_n + t))\right]$$

$$= \lim_{t\downarrow 0} \lim_{n\to\infty} \mathbb{E}^x\left\{ f(X(\tau_n)) \cdot \underbrace{\mathbb{E}^x\left[ g(X(\tau_n + t)) \mid \mathcal{F}_{\tau_n}\right]}_{\text{strong MP}} \right\}$$

$$= \mathbb{E}^{X(\tau_n)}(g(X(t)))$$

$$= (T(t) g)(X(\tau_n))$$

$$= \lim_{t\downarrow 0} \mathbb{E}^x\left\{ f(L(0)) \cdot (T(t) g)(L(0))\right\}$$

$$= \mathbb{E}^x\left\{ f(L(0))\, g(L(0))\right\}.$$

We can approximate $\mathbb{1}_{(L(\sigma)\in\mathcal{O}_1)}$, $\mathcal{O}_1$ an open set, by an increasing limit of elements in $C(S)$.

Therefore, for $\mathcal{O}_1, \mathcal{O}_2$ open and disjoints,

$$\mathbb{P}^x(L(\sigma)\in\mathcal{O}_1, X(\tau)\in\mathcal{O}_2) = \mathbb{P}^x(L(\sigma)\in\mathcal{O}_1, L(\sigma)\in\mathcal{O}_2) = 0.$$

But $S$ is separable $\Rightarrow \mathbb{P}^x(L(\sigma)\neq X(\tau)) = 0$. #

---

- Although continuity is not always present, there is a class of processes where one.

**Def. 2.21)** A diffusion process is a Feller process with continuous paths.

- To verify continuity one can use for example this statement.

**Thm 2.22)** Suppose $X(\epsilon)$ is an $S$-valued stochastic process satisfying:

$$\mathbb{E}\left[\text{dist}(X(\epsilon), X(s))^\gamma\right] \leq C|t-s|^{\beta+1}, \quad 0\leq s,t\leq T$$

for some $\gamma, \beta, C > 0$.

Then, with probability 1, the paths are uniformly continuous on $D\cap[0,T]$, $D\equiv$ diadic rationals,

Furthermore, if the paths are càdlàg, then they are continuous.

(without proof). See stoch. proc. or stoch. analysis lecture.

21.11.2013