CHAPTER 6: FELLER-DYNKIN PROCESSES

1. Preliminaries

In Chapter 5 on Markov processes with countable state spaces, we have investigated in which sense we may think of transition functions P_t as exponentials $\exp(tQ)$ of matrices Q with certain properties: In finite state spaces, there is a one-to-one correspondence between standard transition functions and conservative Q-matrices, given by

$$P_t = \exp(tQ), \quad Q = \frac{\mathrm{d}}{\mathrm{d}t} P_t|_{t=0}.$$

In countably infinite spaces, the situation was complicated by possibly instataneous states ($c(x) = |q(x,x)| = \infty$), non-conservative matrices and the phenomenon of explosion. It is natural to ask whether analogous relations can be formulated for uncountable state spaces, and what the appropriate substitute for the notion of conservative Q-matrix is.

Brownian motion. To get a feel for what changes in relation with finite or countable state spaces, let us look at the transition function of the Brownian family $\mathbb{X}$

$$P_t(x,A) = \mathbb{P}_x(X_t \in A) = \int_A \frac{1}{\sqrt{2\pi t}} \exp\Big(-\frac{(x-y)^2}{2t}\Big)\mathrm{d}y$$

and the induced family of operators $P_t : \mathrm{b}\mathcal{E} \to \mathrm{b}\mathcal{E}$ given by

$$(P_t f)(x) = \mathbb{E}_x[f(X_t)] = \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi t}} \exp\Big(-\frac{(x-y)^2}{2t}\Big) f(y)\mathrm{d}y.$$

Changing variables as $y = x + \sqrt{t}z$ we see that

$$P_t f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x+\sqrt{t}z) \exp\Big(-\frac{z^2}{2}\Big)\mathrm{d}z = \mathbb{E}[f(x+\sqrt{t}Z)]$$

with $Z \sim \mathcal{N}(0,1)$ a standard normal variable. If $f \in C^2(\mathbb{R})$, then

$$f(x+\sqrt{t}z) = f(x) + f'(x)\sqrt{t}z + \frac{1}{2}f''(x)tz^2 + o(t)$$

as $t \to 0$ at fixed z and x. This suggests (but does not prove!) that

$$P_t f(x) = f(x) + \sqrt{t}f'(x)\mathbb{E}[Z] + \frac{1}{2}f''(x)t\mathbb{E}[Z^2] + o(t^2) = f(x) + t\frac{1}{2}f''(x) + o(t)$$

and

$$\frac{\mathrm{d}}{\mathrm{d}t}(P_t f)(x)\Big|_{t=0} = \frac{1}{2}f''(x).$$

So if we try to generalize the relation between Q-matrices and transition function from countable state spaces, it is reasonable to expect that for Brownian motion, the linear map $f \mapsto \frac{1}{2}f''$ should play a certain role. Notice that $\frac{1}{2}f''$ is *not* well-defined for all $f \in \mathrm{b}\mathcal{E}$, even though $P_t f$ is. So the situation is more complicated than for standard transition functions in countable state spaces, for which we were able to define $q(x,y)$ for *all* x, y. Instead of conservative or weakly conservative Q-matrices, we are going to deal with unbounded linear operators.

Substitute for standardness. Let us go a step back and ask how we might generalize the notion of standardness. The definition in terms of matrix elements makes no sense for uncountable state

---

Date: February 8, 2019.

spaces, but we can try something else. If (P_t) is a standard transition function on a countable space E, then for every bounded function $f : E \to \mathbb{R}$,

$$|P_t f(x) - f(x)| \le \sum_{y \in E} |P_t(x,y) - \delta_{x,y}|\,|f(y)| \le 2||f||_\infty (1 - P_t(x,x)) \to 0$$

with

$$||f|| := \sup_{y \in E} |f(y)|.$$

If E is finite, we even have the stronger property

$$||P_t f - f|| \to 0 \quad \text{as } t \to 0.$$

This leaves us with two natural substitutes for standardness: we ask that $P_t f \to f$ as $t \to 0$, either pointwise on $\mathbb{R}$ or uniformly on $\mathbb{R}$. Pointwise convergence is easier to satisfy but uniform convergence is more convenient from an analytic point of view, so we go for the latter.

Unfortunately, for the transition function of Brownian motion, it is not true that $||P_t f - f||_\infty \to 0$ for all $f \in b\mathcal{E}$ (exercise!). This can be remedied by only considering functions f that are continuous and go to zero as $|x| \to \infty$; the space of such functions is denoted $C_0(\mathbb{R})$ (exercise!). Intuitively, this is not so surprising: we ask for $\mathbb{E}_x[f(X_t)] \to f(x)$ as $t \to 0$. If f is not continuous, there is no reason why this should be true. For the uniformity of the convergence, it is worth observing that for all fixed $y \neq x$, $(2\pi t)^{-1/2} \exp(-(x-y)^2/(2t))$ goes to zero as $t \to \infty$, however the convergence is slower and slower the larger $|y - x|$ is: if y is very far away from $X_0 = x$, it should take the process longer to reach y. So for the convergence $P_t f \to f$ to be uniform, it is natural to ask for some condition on f that guarantees that far away y do not matter that much.

Locally compact spaces with countable base. In general topological state spaces E, we say that a function $f : E \to \mathbb{R}$ *vanishes at infinity* if, for every $\varepsilon > 0$, there exists a compact set $K \subset E$ such that $\sup_{E \setminus K} |f| \le \varepsilon$, and we define

$$\begin{aligned} C_0(E) &:= \{f : E \to \mathbb{R} \mid f \text{ is continuous and vanishes at infinity}\} \\ C_b(E) &:= \{f : E \to \mathbb{R} \mid f \text{ is continuous and bounded}\} \\ C(E) &:= \{f : E \to \mathbb{R} \mid f \text{ is continuous}\}. \end{aligned}$$

If E is compact, then every continuous function is bounded and vanishes at infinity, i.e., $C_0(E) = C_b(E) = C(E)$, but in general we only know

$$C_0(E) \subset C_b(E) \subset C(E).$$

There exist Polish spaces E for which $C_0(E)$ consists only of one element, the function that is everywhere equal to zero. This is clearly not what we want. A sufficient condition that guarantees that $C_0(E)$ is rich enough is that E is *locally compact with countable base*, or *lccb* for short. This means, by definition, that

(i) There exists a countable family $(\mathcal{O}_n)_{n \in \mathbb{N}}$ of open sets such that every open set $\mathcal{O} \subset E$ can be written as a union $\mathcal{O} = \cup_{i \in I} \mathcal{O}_i$ for some $I \subset \mathbb{N}$. This property is automatically satisfied when the space is metric and separable (take the open balls with rational radius and centers in a dense countable set).

(ii) For every $x \in E$, there exists an open set $\mathcal{O} \subset E$ such that $x \in \mathcal{O}$ and the closure $\overline{\mathcal{O}}$ is compact.

For example, $\mathbb{R}$ is lccb but $C[0,1]$ with the supremum norm and topology of uniform convergence is not. Every lccb space is Polish, but the converse is not true.

Some properties of lccb spaces are collected in Appendix A.

2. Main definitions and theorems

2.1. **Feller-Dynkin semi-group and Feller-Dynkin family.**

Definition 1. *Let E be an lccb space. A* Feller-Dynkin semi-group *(abbreviated* FD semi-group*) on E is a family $(P_t)_{t \ge 0}$ such that:*

(i) *Each* $P_t : C_0(E) \to C_0(E)$ *is a linear operator.*
(ii) $P_{t+s} = P_t P_s$ *for all* $s, t \geq 0$ *and* $P_0 = I$.
(iii) *We have* $0 \leq f \leq 1 \ \Rightarrow 0 \leq P_t f \leq 1$, *for all* $f \in C_0(E)$ *and* $t \geq 0$.
(iv) $||P_t f - f|| \to 0$ *as* $t \searrow 0$, *for all* $f \in C_0(E)$.

We call a normal Markov family $\mathbb{X}$ *a* Feller-Dynkin family *(*FD family*) if it is has càdlàg sample paths and the family* $(P_t)_{t\geq 0}$ *given by* $(P_t f)(x) = \mathbb{E}_x[f(X_t)]$ *is a Feller-Dynkin semi-group.*

For a FD family we may assume without loss of generality that the filtration is right-continuous, moreover the family is strong Markov. The proof is completely analogous to our earlier theorems, the only difference is that we use $C_0(E)$ instead of $C_{\mathrm{b}}(E)$.[1]

Remark (Strongly continuous contraction semi-group). Property (iv) is called *strong continuity*. Property (iii) implies that

$$\text{(iii')} \quad \forall f \in C_0(E)\ \forall t \geq 0: \quad ||P_t f|| \leq ||f||,$$

i.e., each P_t is a *contraction*. Families that satisfy (i), (ii), (iii'), and (iv) are called *strongly continuous contraction semi-groups*. They can be defined in general Banach spaces (not necessarily $C_0(E)$).

Let us have a closer look at what we are really asking for a family to be Feller-Dynkin. Let $\mathbb{X}$ be a Markov family with càdlàg sample paths and transition function $P_t(x, A) = \mathbb{P}_x(X_t \in A)$. We have noted earlier that if $f \in \mathrm{b}\mathcal{E}$, then the function $P_t f$ defined by $P_t f(x) = \int_E P_t(x, \mathrm{d}y) f(y) = \mathbb{E}_x[f(X_t)]$ is in $\mathrm{b}\mathcal{E}$ as well. The map P_t is clearly linear. The semi-group property $P_{t+s} = P_t P_s$ is inherited from the Chapman-Kolmogorov equations. The implication $0 \leq f \leq 1 \Rightarrow 0 \leq P_t f \leq 1$ holds true for all $f \in \mathrm{b}\mathcal{E}$ because $\mathcal{E} \ni A \mapsto P_t(x, A)$ is a probability measure for all $t > 0$ and $x \in E$. So we really only need to check two things: First, whether it is true that

$$f \in C_0(E) \ \Rightarrow \ \forall t > 0 : P_t f \in C_0(E). \tag{1}$$

Second, is the semi-group strongly continuous on $C_0(E)$ (property (iv) of Definition 1)? Notice that, because of the right-continuity of sample paths and normality, we know that for all $f \in C_0(E)$, $P_t f(x) = \mathbb{E}_x[f(X_t)] \to \mathbb{E}_x[f(X_0)] = f(x)$ as $t \to 0$. The following lemma tells us that strong continuity then comes for free.

Lemma 2. *Let* E *be an lccb space and* $(P_t)_{t\geq 0}$ *a family of operators that satisfies properties (i), (ii), and (iii) of Definition 1, and in addition*

$$(\mathrm{iv}') \quad \forall f \in C_0(E)\ \forall x \in E: \quad \lim_{t\to 0} (P_t f)(x) = f(x).$$

Then $(P_t)_{t\geq 0}$ *also satisfies (iv) of Definition 1 and it is a FD semi-group.*

The important consequence for us is the following:

> A normal Markov family with càdlàg sample paths is a FD family if and only if its transition function preserves $C_0(E)$, i.e., Eq. (1) holds true.

It is not too difficult to check, with the help of the Riesz-Markov theorem (see Appendix A), that for every FD semi-group there is a uniquely defined sub-Markov transition function $(T_t)_{t\geq 0}$ such that $(P_t f)(x) = \int_E T_t(x, \mathrm{d}y) f(y)$. For $(T_t)_{t\geq 0}$ to be a Markov transition function, it is necessary and sufficient that the semi-group satisfies an additional condition. Notice that $\mathbf{1} \in C_0(E)$ if and only if E is compact.

Definition 3. *Let* E *be an lccb space. We call* $(P_t)_{t\geq 0}$ *a* Markovian FD semi-group *if it is a FD semi-group and in addition*

(v) *If* E *is compact:* $P_t \mathbf{1} = \mathbf{1}$ *for all* $t > 0$. *If* E *is not compact: there is a sequence* $(f_n)_{n\in\mathbb{N}}$ *with* $f_n \to \mathbf{1}$ *pointwise and* $\sup_n ||f_n|| < \infty$ *such that* $P_t f_n \to \mathbf{1}$ *pointwise.*

[1] Remember that one of the key elements was a functional monotone class theorem, applied to the multplicative class $C_{\mathrm{b}}(E)$, which in metric spaces generates the Borel σ-algebra. In lccb spaces, something still works with $C_0(E)$ as well because the Borel σ-algebra is already generated by the smaller class $C_0(E)$.

2.2. **Infinitesimal generator.**

Definition 4. *Let E be an lccb space and $(P_t)_{t\geq 0}$ a FD semi-group on E. The* infinitesimal generator *of $(P_t)_{t\geq 0}$ is the operator $L : \mathscr{D}(L) \to C_0(E)$ with domain*

$$\mathscr{D}(L) = \{f \in C_0(E) \mid \exists g \in C_0(E) : \lim_{t\searrow 0} ||\tfrac{1}{t}(P_t f - f) - g|| = 0\}$$

that maps $f \in \mathscr{D}(L)$ to

$$Lf = \lim_{t\searrow 0} \frac{1}{t}(P_t f - f).$$

Conditions (i)–(iii) in the following theorem provide a substitute for the notion of weakly conservative Q-matrix.

Theorem 5. *Let E be an lccb space. An operator $L : \mathscr{D}(L) \to C_0(E)$ is the infinitesimal generator of a FD semi-group if and only if the following three conditions hold true:*

(i) *$\mathscr{D}(L)$ is a dense subspace of $C_0(E)$.*
(ii) *L satisfies the* positive maximum principle*, i.e., for every $f \in \mathscr{D}(L)$ and every maximizer $x_0 \in E$ of f with $f(x_0) \geq 0$,[2] we have $(Lf)(x_0) \leq 0$.*
(iii) *There exists a $\lambda > 0$ such that $\mathscr{R}(\lambda I - L) = C_0(E)$.*

Condition (i) is the next best thing to ask for if the generator L has a domain smaller than $C_0(E)$. Condition (ii) goes well with the candidate generator $Lf = \frac{1}{2}f''$ (and yet to be determined domain $\mathscr{D}(L)$), it replaces the conditions on the signs of the matrix elements of a weakly conservative Q-matrix and on its row sums. Condition (iii) is usually the hardest to check, it does not work if the domain $\mathscr{D}(L)$ is too small. Conditions (ii') below and (iii) reflect that a weakly conservative Q-matrix in finite state space E has no strictly positive eigenvalue, see the remark below.

Remark (Dissipativity, Hille-Yosida theorem)*.* Condition (ii) implies that L is *dissipative*,[3] i.e.,

$$\text{(ii')} \quad \forall \lambda > 0 \quad \forall f \in \mathscr{D}(L): \quad ||\lambda f - Lf|| \geq \lambda ||f||. \tag{2}$$

See [EK86, Chapter 4.2, Lemma 2.1]. The *Hille-Yosida theorem* says that an operator L is the generator of a strongly continuous contraction semi-group if and only if it satisfies conditions (i),(ii'), and (iii). The Hille-Yosida theorem holds true in general Banach spaces.

Remark (Bijectivity of $\lambda I - L$)*.* Later we will see that if L satisfies (i), (ii), and (iii), then condition (iii) is actually satisfies for *all* $\lambda > 0$ so that the operator $\lambda I - L : \mathscr{D}(L) \to C_0(E)$ is surjective for all $\lambda > 0$. The dissipativity (ii') implies that the operator $\lambda I - L : \mathscr{D}(L) \to C_0(E)$ is injective with bounded inverse, $||(\lambda I - L)^{-1} g|| \leq \frac{1}{\lambda}||g||$. Thus $\lambda I - L$ is in fact bijective with bounded inverse. The counterpart in finite state spaces is the following: If Q is a conservative Q-matrix in a finite state space E, then Q cannot have strictly positive eigenvalues. In fact, one knows a little more: all eigenvalues of Q must lie in the complex half-plane $\{\lambda \in \mathbb{C} \mid \operatorname{Re}\lambda \leq 0\}$ (exercise!).

The additional condition of the following proposition replaces the condition that a conservative Q-matrix has row sums equal to zero.

Proposition 6. *Let E be an lccb space. An operator $L : \mathscr{D}(L) \to C_0(E)$ is the infinitesimal generator of a Markovian FD semi-group if and only if it satisfies the conditions (i) to (iii) from Theorem 5 and in addition:*

- *If E is compact: $\mathbf{1} \in \mathscr{D}(L)$, and $L\mathbf{1} = 0$,*
- *If E is not compact: for all sufficiently small $\lambda > 0$, there exists a sequence $(f_n)_{n\in\mathbb{N}}$ (that may depend on λ) so that $g_n = f_n - \lambda L f_n$ satisfies $\sup_n ||g_n|| < \infty$ and both f_n and g_n converge to $\mathbf{1}$ pointwise.*

[2] CORRECTION!! the additional condition $f(x_0) \geq 0$ was missing in class and in earlier versions of these notes, it is only needed when E is compact and FD semi-group is really sub-Markovian and not Markovian.

[3] The word is explained as follows: Suppose that we work in a Hilbert space with scalar product $\langle \cdot, \cdot \rangle$ and norm $||f|| = \sqrt{\langle f, f\rangle}$ (instead of the supremum norm). Then the condition (2) is equivalent to $\operatorname{Re}\langle f, Lf\rangle \leq 0$ for all f, from which we may in turn deduce that $\frac{\mathrm{d}}{\mathrm{d}t}||\exp(tL)f||^2 \leq 0$. In some PDE applications, the Hilbert space might be $L^2(\mathbb{R}^3)$ and $|u(x,t)|^2 := |(\exp(tL)f)(x)|^2$ may have the interpretation of an energy density. Then $\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^3} |u(x,t)|^2 \mathrm{d}x \leq 0$ says that the total energy can only decrease or be dissipated, but never increase.

2.3. **Canonical process associated with a given FD semi-group.** In order to show that for a given Markovian FD semi-group an associated FD family exists, we construct the canonical version. Let $\Omega = \mathcal{D}_E[0,\infty)$ be the space of càdlàg functions $\omega : [0,\infty) \to E$. For $t \geq 0$, let $X_t(\omega) := \omega(t)$. Further let $\mathcal{F} := \sigma(X_t,\, t \geq 0)$ and $\mathcal{F}_t^0 := \sigma(X_s,\, s \geq t)$. Finally let $(\theta_s\omega)(t) := \omega(s+t)$.

Theorem 7. *Let E be an lccb space and $(P_t)_{t\geq 0}$ a Markovian FD semigroup on E. Then there exists a uniquely defined family $(\mathbb{P}_x)_{x\in E}$ of probability measures on $(\Omega,\mathcal{F})$ such that*

$$\mathbb{X} = (\Omega, \mathcal{F}, (\mathcal{F}_t^0)_{t\geq 0}, (\mathbb{P}_x)_{x\in E}, (X_t)_{t\geq 0}, (\theta_s)_{s\geq 0})$$

is a FD family with $\mathbb{E}_x[f(X_t)] = (P_tf)(x)$ for all $f \in C_0(E), t \geq 0, x \in E$.

Remark (Sub-Markov case). The sub-Markov case is usually dealt with by making the state space a little larger. Let ∂ be some element not in E, called *coffin* or *cemetery*, and $E_\partial := E \cup \{\partial\}$. Equip E_∂ with $\mathcal{E}_\partial := \sigma(\mathcal{B}(E), \{\partial\})$, i.e., the smallest σ-algebra that contains the singleton $\{\partial\}$ and all sets from the Borel σ-algebra E. Extend the sub-Markov transition function $(P_t)_{t\geq 0}$ as follows:

$$\begin{aligned}
P_t^\partial(x,\{\partial\}) &:= 1 - P_t(x,E) \qquad (x \in E),\\
P_t^\partial(x,A) &:= P_t(x,A) \qquad (x \in E, A \in \mathcal{B}(E)),\\
P_t^\partial(\partial,\cdot) &:= \delta_\partial(\cdot).
\end{aligned}$$

Then $(P_t^\partial)_{t\geq 0}$ is a transition function on $(E_\partial, \mathcal{E}_\partial)$. For example, if $E = \{1,2\}$, we may represent P_t and P_t^∂ by 2×2 and 3×3 stochastic matrices, respectively, and

$$P_t^\partial = \begin{pmatrix} P_t(1,1) & P_t(1,2) & 1 - P_t(1,1) - P_t(1,2) \\ P_t(2,1) & P_t(2,2) & 1 - P_t(2,1) - P_t(2,2) \\ 0 & 0 & 1 \end{pmatrix}.$$

Each function $f \in C_0(E)$ is extended to a function $f : E_\partial \to \mathbb{R}$ by defining $f(\partial) := 0$; we have $(P_t^\partial f)(x) = (P_tf)(x)$ for all $x \in E$. So as far as functions $f \in C_0(E)$ are concerned, there is no difference between the extended semi-group and the original semi-group.

The path space is extended as follows: let Ω_∂ be the set of paths $\omega : \mathbb{R}_+ \to E_\partial$ such that:

- either $\omega(t) \in E$ for all $t \geq 0$, and $t \mapsto \omega(t)$ is càdlàg,
- or there exists $\zeta(\omega) > 0$ such that $\omega(t) = \partial$ for all $t \geq \zeta(\omega)$, $\omega(t) \in E$ for all $t < \zeta(\omega)$, and $\omega(\cdot)$ is right-continuous and has left limits in every $t_0 < \zeta(\omega)$,
- or $\omega(t) = \partial$ for all $t \geq 0$.

We *do not* ask for the existence of left limits as $t \nearrow \zeta(\omega)$. Define $X_t^\partial(\omega) := \omega(t)$. Then an analogue of Theorem 7 holds true with Ω replaced with Ω_∂, and the natural choices of σ-algebra, filtration, and shift operators. The definition of $\zeta(\omega)$ is extended to all $\omega \in \Omega$ by

$$\zeta(\omega) := \inf\{t \geq 0 \mid X_t^\partial(\omega) = \partial\}.$$

ζ is the *life-time* of the process. See [RW94, Section III.7].

3. Important examples

3.1. **Finite and countable state spaces.**

Example 1 (Conservative finite Q-matrices). Let E be a finite set with cardinality $n \in \mathbb{N}$, equipped with the discrete topology, and Q a conservative Q-matrix. Then E is lccb and in fact compact, and we can identify $C_0(E) = C(E)$ with $\mathbb{R}^n$. Let L be the operator with domain $\mathscr{D}(L) = C_0(E)$ and

$$Lf(x) = \sum_{y\in E} q(x,y)f(y).$$

Notice that, because of the conservativity of Q,

$$Lf(x) = -c(x)f(x) + \sum_{\substack{y\in E:\\ y\neq x}} q(x,y)f(y) = \sum_{y\in E} q(x,y)\big(f(y) - f(x)\big).$$

Then L satisfies conditions (i)–(iii) from Theorem 5, the associated FD semi-group is given by $P_t = \exp(tQ)$ (exercise!).

More generally, if E be a countable space, equipped with the discrete topology, we may ask two questions:

(1) True or false: A standard transition function $(P_t)_{t\geq 0}$ induces a Feller-Dynkin semi-group if and only if it has a weakly conservative Q-matrix.

(2) Let Q be a conservative Q-matrix. Consider the linear operator L_Q in $C_0(E)$ with domain

$$\mathscr{D}(L_Q) := \Big\{ f \in C_0(E) \mid \text{ the map } x \mapsto \sum_{y\in E} q(x,y)\big(f(y) - f(x)\big) \text{ is in } C_0(E) \Big\}$$

and $L_Q f(x) := \sum_{y\in E} q(x,y)(f(y) - f(x))$. True or false: L_Q is a FD generator if and only if the minimal solution $(P_t)_{t\geq 0}$ is stochastic (i.e., there is no explosion).

It turns out that the answer to both question is: false. We sketch why and leave the details as exercise: Consider for example $E = \mathbb{N}_0$. Then a standard transition function $(P_t)_{t\geq 0}$ in E is FD if and only if, for all $t > 0$ and $y \in \mathbb{N}_0$,

$$\lim_{|x|\to\infty} P_t(x,y) = 0,$$

i.e., for each column of P_t, the matrix elements go to zero as you move further and further down the column. The probabilistic interpretation is that

Feller-Dynkin means that the process cannot come in from infinity in finite time

(think about death chains), while intuitively,

non-explosion means that the process cannot reach infinity in finite time.

(think about birth chains). These are two different things! The difference carries over to Q-matrices and generators: with the inequality

$$P_t(x,y) \geq \mathrm{e}^{-c(x)t} + \int_0^t \mathrm{e}^{-c(x)s} q(x,y) \mathrm{e}^{-c(y)(t-s)} \mathrm{d}s,$$

one can show that any weakly conservative Q-matrix associated with a FD semi-group necessarily satisfies

$$\forall y \in E: \quad \lim_{x\to\infty} q(x,y) = 0,$$

which can be used to build counter-examples for the second question. For another counter-example, look at the death chain from Sheet 12. So in general, Feller-Dynkin property and non-explosion of minimal solutions are not equivalent.

3.2. **Jump processes.**

Example 2. Let $K : E \times \mathcal{B}(E) \to \mathbb{R}_+$ be a kernel (i.e., $x \mapsto K(x,A)$ is measurable for all $A \in \mathcal{B}(E)$ and $A \mapsto K(x,A)$ is a measure, for all $x \in A$) with $K(x,\{x\}) = 0$. Consider the formal operator

$$(Lf)(x) := \int_E K(x,\mathrm{d}y)\big(f(y) - f(x)\big).$$

Under suitable assumptions on K, the operator L with suitable domain $\mathscr{D}(L)$, is a FD generator.

If $K(x,E) < \infty$ for all $x \in E$, then we need not use FD theory and can actually construct the process by techniques quite similar to what we did for countable state spaces. See [EK86, Chapter 4.2] for the bounded case $\sup_x K(x,E) < \infty$ and [Fel71, Chapter X.3] for the general case. The process behaves as follows: when started in x, it waits for an exponentially distributed time $\mathrm{Exp}(c(x))$ with $c(x) = K(x,E)$ and then it jumps to another point according to the kernel $\Pi(x,\mathrm{d}y) := \frac{1}{c(x)} K(x,\mathrm{d}y)$.

Notice that

$$||Lf|| \leq ||f|| \sup_{x\in E} |K(x,E)|,$$

so if $\sup_{x\in E} K(x,E) < \infty$, we may view L as a bounded operator in $\mathrm{b}\mathcal{E}$ and define $P_t = \exp(tL)$ by the exponential series.

An outcome of the Yosida approximation constructed below is that every FD process can be approximated by a jump process with bounded generator.

3.3. **Brownian motion and relatives.**

Example 3 (Brownian motion). The infinitesimal generator of Brownian motion is the operator $Lf = \frac{1}{2}f''$ with domain

$$\mathscr{D}(L) = \{f \in C_0(\mathbb{R}) \cap C^2(\mathbb{R}) \mid f', f'' \in C_0(\mathbb{R})\}.$$

In particular, this operator satisfies conditions (i)—(iii) from Theorem 5.

Example 4 (Uniform motion to the right). Consider the operator L in $C_0(\mathbb{R})$ with domain

$$\mathscr{D}(L) := \{f \in C_0(\mathbb{R}) \cap C^1(\mathbb{R}) \mid f' \in C_0(\mathbb{R})\}$$

given by $(Lf)(x) = f'(x)$. Then L satisfies conditions (i)—(iii) from Theorem 5. The generated FD semi-group is given by $P_t f(x) = f(x+t)$, the associated FD family satisfies $X_t = x+t$, $\mathbb{P}_x$-a.s.

The previous two examples raise the question whether there are processes with a generator L given by a third-order derivative. The answer is no (exercise!). However a whole class of generators can be built by combining second- and first-order derivatives, of the type

$$Lf = \frac{1}{2}a(x)f''(x) + b(x)f'(x)$$

with suitable domain $\mathscr{D}(L)$ and under conditions on a and b. The intuition is that first-order differential operators are associated with deterministic motions while second-order differential operators bring in Brownian motion; the corresponding processes can often be written as solutions to stochastic differential equations. This topic belongs to the chapter of *diffusions*.[4]

The following two examples illustrate that the domain $\mathscr{D}(L)$ really matters and often encodes the behavior of a process at the boundary of an interval.

Example 5 (Brownian motion on $[0,\infty)$ with reflection at 0). Let

$$\mathbb{X} = (\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, (\mathbb{P}_x)_{x\in\mathbb{R}}, (X_t)_{t\geq 0}, (\theta_s)_{s\geq 0})$$

be a Brownian family, think $\mathbb{P}_x(X_t \in A) = \mathbb{P}(x + B_t \in A)$ with $(B_t)_{t\geq 0}$ a standard Brownian motion defined on some probability space $(\tilde{\Omega}, \mathcal{G}, \mathbb{P})$. Define

$$Y_t(\omega) := |X_t(\omega)| \qquad (t \geq 0, \omega \in \Omega).$$

Then $\mathbb{Y} = (\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, (\mathbb{P}_x)_{x\in\mathbb{R}}, (Y_t)_{t\geq 0}, (\theta_s)_{s\geq 0})$ is a FD family with state space $E = [0,\infty)$ (exercise!). The associated generator L^{refl} has domain

$$\mathscr{D}(L^{\mathrm{refl}}) = \{f \in C_0(\mathbb{R}_+) \mid f', f'' \in C_0(\mathbb{R}_+),\ f'(0) = 0\}$$

and is given by $L^{\mathrm{refl}}f = \frac{1}{2}f''$. Indeed, let $f \in C_0([0,\infty))$ and $f_e : \mathbb{R} \to \mathbb{R}$ the even extension, given by $f_e(x) := f(|x|)$. Then $f_e \in C_0(\mathbb{R})$ and for all $x \geq 0$ and $t \geq 0$,

$$\mathbb{E}_x[f(Y_t)] = \mathbb{E}_x[f(|X_t|)] = \mathbb{E}_x[f_e(X_t)].$$

It follows that if $f_e \in \mathscr{D}(L)$ with L be the generator of Brownian motion, then $f \in \mathscr{D}(L^{\mathrm{refl}})$, furthermore

$$L^{\mathrm{refl}}f = Lf_e\Big|_{\mathbb{R}_+} = \frac{1}{2}f''.$$

Conversely, if $f \in \mathscr{D}(L^{\mathrm{refl}})$: let $g := L^{\mathrm{refl}}f$. Then

$$\lim_{t\searrow 0}\sup_{x\geq 0}\Big|\frac{1}{t}\Big(\mathbb{E}_x[f_e(X_t)] - f_e(x)\Big) - g(x)\Big| = 0.$$

[4]A *Feller-Dynkin diffusion* in $\mathbb{R}^n$ is a FD process in $\mathbb{R}^n$ with *continuous* sample paths such that the domain $\mathscr{D}(L)$ of the generator contains the space $C^\infty_\kappa(\mathbb{R}^n)$ of C^∞-functions with compact support. It is a general theorem that the generator of every FD diffusion, restricted to $C^\infty_\kappa(\mathbb{R}^n)$, is a second-order differential operator. See Rogers and Williams, Vol. 1, Chapter III.13.3.

Because of Since the Brownian semi-group $(P_t)_{t\geq 0}$ maps even functions to even functions (exercise! think $\mathbb{E}_x[f_e(X_t)] = \mathbb{E}[f_e(x+B_t)] = \mathbb{E}[f_e(x-B_t)] = \mathbb{E}[f_e(-x+B_t)] = \mathbb{E}_{-x}[f(X_t)]$), we deduce

$$\lim_{t\searrow 0}\sup_{x\in\mathbb{R}}\Big|\frac{1}{t}\Big(\mathbb{E}_x[f_e(X_t)] - f_e(x)\Big) - g_e(x)\Big| = 0$$

with $g_e \in C_0(\mathbb{R})$ the even extension of g. It follows that $f_e \in \mathscr{D}(L)$. Thus $f \in \mathscr{D}(L^{\text{refl}})$ if and only if $f_e \in \mathscr{D}(L)$, which in turn happens if and only if $f \in C_0(\mathbb{R}_+)$ satisfies $f', f'' \in C_0(\mathbb{R}_+)$ and $f'(0) = 0$. This last condition is needed for the continuity of f_e' at $x = 0$.

Example 6 (Brownian motion on $[0,\infty)$ with absorption at 0). Let $\mathbb{X}$ be a Brownian family and $\tau := \inf\{t \geq 0 \mid X_t = 0\}$. Define

$$Y_t(\omega) := \begin{cases} X_t(\omega), & t < \tau(\omega), \\ 0, & t \geq \tau(\omega). \end{cases}$$

Then we obtain again a FD family with state space $\mathbb{R}_+$. The generator is $L^{\text{abs}} f = \frac{1}{2} f''$ with domain

$$\mathscr{D}(L^{\text{abs}}) = \{f \in C_0(\mathbb{R}_+) \mid f', f'' \in C_0(\mathbb{R}_+),\ f(0) = 0,\ f''(0) = 0\}. \tag{3}$$

Let $f_o \in C_b(\mathbb{R})$ be the bounded function given by

$$f_o(x) := \begin{cases} f(x), & x \geq 0, \\ 2f(0) - f(-x), & x < 0. \end{cases}$$

We show that $\mathbb{E}_x[f(Y_t)] = \mathbb{E}_x[f_o(X_t)]$ for all $x \geq 0$ but have to work a little harder than for Brownian motion reflected at 0. Notice $f_o \in C_0(\mathbb{R})$ if and only if $f(0) = 0$, and f_o'' is continuous at $x = 0$ if and only if $f''(0) = 0$.

Using the strong Markov property (in a way similar to the reflection principle), we see that the process $Z_t := X_t \mathbb{1}_{\{t<\tau\}} + (-X_t)\mathbb{1}_{\{t\geq\tau\}}$ leads again to a Brownian family, therefore for $x \geq 0$

$$\mathbb{E}_x\big[f_o(X_t)\mathbb{1}_{\{t\geq\tau\}}\big] = \mathbb{E}_x\big[f_o(Z_t)\mathbb{1}_{\{t\geq\tau\}}\big] = \mathbb{E}_x\big[f_o(-X_t)\mathbb{1}_{\{t\geq\tau\}}\big]$$

hence

$$\mathbb{E}_x\big[f_o(X_t)\mathbb{1}_{\{t\geq\tau\}}\big] = \mathbb{E}_x\big[\frac{1}{2}(f_o(X_t) + f_o(-X_t))\mathbb{1}_{\{t\geq\tau\}}\big] = f(0)\mathbb{P}_x(t \geq \tau)$$

and

$$\begin{aligned} \mathbb{E}_x[f(Y_t)] &= \mathbb{E}_x\big[f(X_t)\mathbb{1}_{\{t<\tau\}}\big] + f(0)\mathbb{P}_x(\tau \leq t) \\ &= \mathbb{E}_x\big[f_o(X_t)\mathbb{1}_{\{t<\tau\}}\big] + \mathbb{E}_x\big[f_o(X_t)\mathbb{1}_{\{t\geq\tau\}}\big] = \mathbb{E}_x\big[f_o(X_t)\big]. \end{aligned}$$

Let us write G for the generator of $(Y_t)_{t\geq 0}$. We wish to show that G is given by L^{abs} with domain (3).

First we show $\mathscr{D}(L^{\text{abs}}) \subset \mathscr{D}(G)$ and $Gf = \frac{1}{2}f'' = L^{\text{abs}} f$ for all $f \in \mathscr{D}(L^{\text{abs}})$: If $f \in C_0(\mathbb{R}_+)$ satisfies $f', f'' \in C_0(\mathbb{R}_+)$ and $f(0) = f'(0) = 0$, then $f_o \in C_0(\mathbb{R})$ with $f_o', f_o'' \in C_0(\mathbb{R})$, thus $f_o \in \mathscr{D}(L)$ with L the generator of Brownian motion on $\mathbb{R}$. It follows that $f \in \mathscr{D}(G)$ with $Gf = Lf|_{\mathbb{R}_+} = \frac{1}{2}f''$.

The proof is complete once we check that not only is $\mathscr{D}(L^{\text{abs}}) \subset \mathscr{D}(G)$ but in fact $\mathscr{D}(G) = \mathscr{D}(L^{\text{abs}})$. The proof is concluded as follows: 1. Show that L^{abs} with domain (3) satisfies conditions (i)—(iii) in Theorem 5. 2. Notice that the generator G satisfies these conditions as well because of Theorem 5. 3. Show that if L_1, L_2 are two operators that satisfy conditions (i)—(iii) in Theorem 5 and $\mathscr{D}(L_1) \subset \mathscr{D}(L_2)$ with $L_2 f = L_1 f$ for all $f \in \mathscr{D}(L_1)$, then $\mathscr{D}(L_1) = \mathscr{D}(L_2)$. (A generator cannot be the extension of another generator.) 4. Conclude that $\mathscr{D}(G) = \mathscr{D}(L^{\text{abs}})$.

3.4. **Cauchy process.**

Example 7 (Cauchy process). The Cauchy process is a process with state space $\mathbb{R}$ and stationary independent increments such that $X_{t+s} - X_s$ has probability density function $\rho_t(x) = \frac{1}{\pi}\frac{t}{t^2+x^2}$. The associated semi-group is

$$(P_t f)(x) := \int_{\mathbb{R}} \rho_t(y-x) f(y)\mathrm{d}y = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{t}{t^2+(y-x)^2} f(y)\mathrm{d}y.$$

Then $(P_t)_{t\geq 0}$ is a FD semi-group. Furthermore if $f \in C_0(\mathbb{R})$ is twice differentiable, then $f \in \mathscr{D}(L)$ and

$$(Lf)(x) = \lim_{\varepsilon\searrow 0}\int_{\mathbb{R}\setminus[-\varepsilon,\varepsilon]} \frac{f(x+u)-f(x)}{u^2}\mathrm{d}u = \frac{1}{\pi}\int_0^\infty \frac{f(x+u)+f(x-u)-2f(x)}{u^2}\,\mathrm{d}u.$$

Notice that $1/u^2$ is not integrable near $u=0$, so in general, the integral $\int_{\mathbb{R}}(f(x+u)-f(x))\frac{\mathrm{d}u}{u^2}$ is not absolutely absolutely convergent. The divergence is cured by working instead with the middle expression, which is the so-called *Cauchy principal value* of $\int_{\mathbb{R}}(f(x+u)-f(x))\frac{\mathrm{d}u}{u^2}$. For the existence of the Cauchy principal value, we use a change of variables for the integral over $(-\infty,-\varepsilon)$, which yields

$$\int_{\mathbb{R}\setminus[-\varepsilon,\varepsilon]} \frac{f(x+u)-f(x)}{u^2}\mathrm{d}u = \int_0^\infty \frac{f(x+u)+f(x-u)-2f(x)}{u^2}\,\mathrm{d}u,$$

Next we notice that $u \mapsto \frac{f(x+u)+f(x-u)-2f(x)}{u^2}$ is integrable near $u=0$ because it converges to $f''(0)$. For large u we use the bound by $4||f||/u^2$ and remember $\int_1^\infty \frac{\mathrm{d}u}{u^2} < \infty$. It follows that integrand in the right-hand side is integrable on $(0,\infty)$ and the limit as $\varepsilon \searrow 0$ exists.

Remark (Cauchy process and uniform angles). Changing variables as $z = (y-x)/t$ and then $\theta = \arctan z$, we get

$$(P_t f)(x) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{f(x+tz)}{1+z^2}\mathrm{d}z = \frac{1}{\pi}\int_{-\pi/2}^{\pi/2} f(x+t\tan\theta)\mathrm{d}\theta.$$

Thus

$$(P_t f)(x) = \mathbb{E}[f(x+t\tan\Theta)]$$

with Θ a random variable uniformly distributed in $(-\frac{\pi}{2},\frac{\pi}{2})$. The Markov property translates into the following surprising fact: if $\Theta_1,\Theta_2 \sim \mathcal{U}((-\frac{\pi}{2},\frac{\pi}{2}))$ are two additional independent variables, then for all $s,t>0$,

$$s\tan\Theta_1 + t\tan\Theta_2 \stackrel{d}{=} (s+t)\tan(\Theta),$$

which in turn can be interpreted as a special case of the *Huygens principle* in geometric optics (see the footnote to Example (e) in [Fel71, Chapter II.4]). Notice that the naive guess $\frac{\mathrm{d}}{\mathrm{d}t}P_t f(x)|_{t=0} = f'(x)\mathbb{E}[\tan\Theta] = f'(x)\frac{1}{\pi}\int_{\mathbb{R}}\frac{z}{1+z^2}\mathrm{d}z$ fails because

$$\mathbb{E}[|\tan\Theta|] = \frac{1}{\pi}\int_{\mathbb{R}} \frac{|z|}{1+z^2}\,\mathrm{d}z = \infty.$$

Remark (Cauchy process as a jump process). Formally (not worrying about domains, convergence of integrals etc.), the generator of the Cauchy process is of the form

$$(Lf)(x) = \int_{\mathbb{R}} K(x,\mathrm{d}y)\big(f(y)-f(x)\big)$$

with kernel $K(x,A) = \int_{\mathbb{R}} \mathbb{1}_A(x+u)\frac{\mathrm{d}u}{u^2}$. This should remind you of the formal generator for jump processes given earlier, even though here $K(x,\mathbb{R}) = \int_{\mathbb{R}}\frac{\mathrm{d}u}{u^2} = \infty$. Despite this difference, it turns out that the sample paths of the Cauchy process are piecewise constant functions. However the number of jump discontinuities in bounded intervals is infinite with positive probability. This is no contradiction with the property of having càdlàg paths because jump heights can be arbitrarily small.

It is not too difficult to check that the Cauchy process has stationary independent increments. Looking back, we notice that by now we know three examples of processes that are Feller-Dynkin and have stationary increments: Brownian motion, the Poisson process,[5] and the Cauchy process. All three are examples of *Lévy processes*, i.e., processes with stationary independent increments and càdlàg paths.

4. Resolvent

In addition to the transition function $(P_t)_{t\geq 0}$ and the generator L another family of operators plays an important role: the *resolvent*. Let $\mathbb{X}$ be a Markov family with transition function $(P_t)_{t\geq 0}$. Assume that $\mathbb{R}_+ \ni t \mapsto P_t(x, A)$ is measurable for all $x \in E$ and $A \in \mathcal{E}$. This is the case, for example, when $\mathbb{X}$ has càdlàg sample paths. Then the *resolvent* of the process is the family $(U_\alpha)_{\alpha>0}$ of maps $U_\alpha : \mathrm{b}\mathcal{E} \to \mathrm{b}\mathcal{E}$ given by

$$(U_\alpha f)(x) := \int_0^\infty \mathrm{e}^{-\alpha t}(P_t f)(x)\mathrm{d}t \qquad (x \in E).$$

The probabilistic interpretation is the following: let $T \sim \mathrm{Exp}(\alpha)$ be an exponential random variable, with parameter α, independent of $(X_t)_{t\geq 0}$. Then

$$(U_\alpha f)(x) = \mathbb{E}_x\big[f(X_T)\big].$$

If $P_t(C_0(E)) \subset C_0(E)$ for all $t \geq 0$, then $U_\alpha(C_0(E)) \subset C_0(E)$ for all $\alpha > 0$ (exercise!), so we may also view the resolvent as a map $U_\alpha : C_0(E) \to C_0(E)$.

The resolvent of a semi-group is defined in a completely analogous way.

Example 8. If $E = \{1, \ldots, n\}$ with $n \in \mathbb{N}$, then we can identify $C_0(E)$ with $\mathbb{R}^n$ (think of functions as column vectors) and linear operators from $C_0(E)$ to itself with matrices. If Q is a conservative Q-matrix and $P_t = \exp(tQ)$, then the resolvent is given by

$$U_\alpha = \int_0^\infty \mathrm{e}^{\alpha t}\mathrm{e}^{tQ}\mathrm{d}t = (\alpha - Q)^{-1} \qquad (\alpha > 0).$$

(remember our earlier remark that Q has no strictly positive eigenvalue). For $\alpha = 0$, we recover the Green's function G, which need not be finite.

Proposition 8. *Let $(P_t)_{t\geq 0}$ be a sub-Markov semi-group in $C_0(E)$, i.e., it satisfies properties (i), (ii), and (iii) of Definition 1. Assume that $t \mapsto (P_t f)(x)$ is measurable for all $x \in E$ and $f \in C_0(E)$. Then $(U_\alpha)_{\alpha>0}$ is a sub-Markov resolvent in $C_0(E)$, meaning that:*

(i) *Each U_α is a linear map from $C_0(E)$ to $C_0(E)$.*
(ii) *We have $0 \leq f \leq 1 \ \Rightarrow 0 \leq \alpha U_\alpha f \leq 1$, for all $\alpha > 0$ and $f \in C_0(E)$.*
(iii) *The* resolvent identity *holds: for all $\alpha, \beta > 0$,*

$$U_\alpha - U_\beta = (\beta - \alpha)U_\alpha U_\beta.$$

Moreover $(P_t)_{t\geq 0}$ is strongly continuous if and only if in addition

(iv) *For all $f \in C_0(E)$: $||\alpha U_\alpha f - f|| \to 0$ as $\alpha \to \infty$.*

The resolvent identity implies that $U_\alpha U_\beta = U_\beta U_\alpha$ for all $\alpha, \beta > 0$.

Remark (Strongly continuous contraction resolvent). Condition (ii) implies the condition

$$\text{(ii')} \qquad \forall \alpha > 0\ \forall f \in C_0(E) : \ ||\alpha U_\alpha f|| \leq ||f||.$$

Families of operators that satisfy (i), (ii'), and (iii) are called *contraction resolvents*. If in addition (iv) holds true then $(U_\alpha)_{\alpha>0}$ is called a *strongly continuous contraction resolvent*.

---

[5]In our definition, the Poisson process $(N_t)_{t\geq 0}$ was an $\mathbb{N}_0$-valued process with $N_0 = 0$ a.s., but because of $\mathbb{N}_0 \subset \mathbb{R}$ we can also view it as a real-valued process, and one can define an associated family with $\mathbb{P}_x(X_t \in A) = \mathbb{P}(x + N_t \in A)$. The proof that this results in a FD family is left as an exercise.

Remark (Range of the resolvent). A close look at the proof of Proposition 8 reveals that strong continuity is equivalent to the density of $\mathcal{R} = \mathscr{R}(U_\alpha)$ in $C_0(E)$. The range does not depend on α because of the resolvent identity. Indeed, if $f \in \mathscr{R}(U_\alpha)$, write $f = U_\alpha g$. Then

$$f = U_\alpha g = U_\beta g + (\beta - \alpha) U_\alpha U_\beta g = U_\beta (g + (\beta - \alpha) U_\alpha g) \in \mathscr{R}(U_\beta)$$

hence $\mathscr{R}(U_\alpha) \subset \mathscr{R}(U_\beta)$. The roles of α and β can be inverted and we deduce $\mathscr{R}(U_\alpha) = \mathscr{R}(U_\beta)$ for all $\alpha, \beta > 0$.

When computing the action of the semi-group on functions in the resolvent, the following observation proves useful: Let $f \in C_0(E)$ and $\alpha > 0$. Then for all $t > 0$ and $x \in E$,

$$(P_t U_\alpha f)(x) = \int_0^\infty \mathrm{e}^{-\alpha s} P_{t+s} f(x) \mathrm{d}s.$$

Indeed, let $(T_t)_{t \geq 0}$ be the transition function associated with the semi-group (whose existence is proven with the Riesz-Markov theorem). Set $\mu_{t,x}(A) := T_t(x, A)$. Then by Fubini's theorem,

$$\begin{aligned}
(P_t U_\alpha f)(x) &= \int_E T_t(x, \mathrm{d}y)(U_\alpha f)(y) \\
&= \int_E \Big(\int_0^\infty \mathrm{e}^{-\alpha s} \big(P_s f\big)(y) \mathrm{d}s \Big) \mu_{t,x}(\mathrm{d}y) \\
&= \int_0^\infty \Big(\int_E \mathrm{e}^{-\alpha s} \big(P_s f\big)(y) \mu_{t,x}(\mathrm{d}y) \Big) \mathrm{d}s \\
&= \int_0^\infty \mathrm{e}^{-\alpha s} (P_{t+s} f)(x) \mathrm{d}s.
\end{aligned}$$

Proof of Proposition 8 . The proof of (i) is left as an exercise. For (ii), we note that if $0 \leq f \leq 1$, then using (ii) of Definition 1, we get

$$0 \leq \alpha \int_0^\infty \mathrm{e}^{-\alpha t} (P_t f)(x) \mathrm{d}t \leq \int_0^\infty \alpha \mathrm{e}^{-\alpha t} \mathrm{d}t = 1.$$

For (iii), we evaluate

$$\begin{aligned}
(U_\alpha U_\beta f)(x) &= \int_0^\infty \mathrm{e}^{-\alpha t} \Big(\int_0^\infty \mathrm{e}^{-\beta s} (P_{s+t} f)(x) \mathrm{d}s \Big) \mathrm{d}t \\
&= \int_0^\infty (P_r f)(x) \Big(\int_0^r \mathrm{e}^{-\alpha t - \beta(r-t)} \mathrm{d}r \Big) \\
&= \int_0^\infty (P_r f)(x) \frac{\exp(-\alpha r) - \exp(-\beta r)}{\beta - a\alpha} \mathrm{d}r \\
&= \frac{1}{\beta - \alpha} \Big((U_\alpha f)(x) - (U_\beta f)(x) \Big).
\end{aligned}$$

Finally if $(P_t)_{t \geq 0}$ is strongly continuous, then for all $t_0 > 0$,

$$|\alpha U_\alpha f(x) - f(x)| \leq \int_0^\infty \alpha \mathrm{e}^{-\alpha t} |P_t f(x) - f(x)| \mathrm{d}t \leq t_0 \sup_{t \in [0, t_0]} ||P_t f - f|| + 2||f|| \mathrm{e}^{-\alpha t_0}.$$

Because of the strong continuity of $(P_t)_{t \geq 0}$, given $\varepsilon > 0$, we can find $t = t_0(\varepsilon) > 0$ such that for all $t \in [0, t_0]$, we have $||P_t f - f|| \leq \varepsilon/2$. Making t_0 smaller if needed, we may always assume that $t_0 \leq 1$. Then we can find $\alpha_0 = \alpha_0(t_0, \varepsilon)$ such that for all $\alpha \geq \alpha_0$, we have $2||f|| \exp(-\alpha t_0) \leq \varepsilon/2$. Altogether $||\alpha U_\alpha f - f|| \leq \varepsilon$, for all $\alpha \geq \alpha_0$. It follows that $||\alpha U_\alpha f - f|| \to 0$.

Conversely, if $||\alpha U_\alpha f - f|| \to 0$ for all $f \in C_0(E)$: We prove first that $||P_t f - f|| \to 0$ whenever f lies in the range of one of the αU_α's. Thus let $f = \alpha U_\alpha g$ with $\alpha > 0$, $g \in C_0(E)$. Then

$$\begin{aligned} P_t f(x) - f(x) &= \int_0^\infty \alpha \mathrm{e}^{-\alpha s}\Big(P_{t+s} g(x) - P_s g(x)\Big)\mathrm{d}s \\ &= \int_t^\infty \alpha \mathrm{e}^{\alpha(t-u)} P_u g(x)\mathrm{d}u - \int_0^\infty \alpha \mathrm{e}^{-\alpha s} P_s g(x)\mathrm{d}s \\ &= -\int_0^t \alpha \mathrm{e}^{-\alpha s} P_s g(x)\mathrm{d}s + (\mathrm{e}^{\alpha t} - 1)\int_t^\infty \alpha \mathrm{e}^{-\alpha s} P_s g(x)\mathrm{d}s \end{aligned}$$

hence

$$||P_t f - f|| \le t||g|| + (\mathrm{e}^{\alpha t} - 1)||g||$$

which goes to zero as $t \to 0$. This proves the claim if f lies in the range of αU_α. For general f: let $\varepsilon > 0$ and $\alpha > 0$ small enough so that $||\alpha U_\alpha f - f|| \le \varepsilon/3$. Set $F := \alpha U_\alpha f$. Then

$$||P_t f - f|| \le ||P_t f - P_t F|| + ||P_t F - F|| + ||F - f|| \le 2\tfrac{\varepsilon}{3} + ||P_t F - F||.$$

We have already checked that $||P_t F - F|| \to 0$ so there exists $t_0 > 0$ such that for all $t \le t_0$, we have $||P_t F - F|| \le \varepsilon/3$. Then also $||P_t f - f|| \le \varepsilon/3$ for all $t \le t_0$. $\square$

Proof of Lemma 2. Let $(P_t)_{t\ge 0}$ be a weakly continuous sub-Markov semi-group, i.e., it satisfies conditions (i)–(iii) from Definition 1 and condition (iv') from the statement of the lemma. By the semi-group property, the map $\mathbb{R}_+ \ni t \to P_t f(x)$ is right-continuous, hence measurable, for every $f \in C_0(E)$ and $x \in E$. Therefore the resolvent $U_\alpha f$ is well-defined and we are in the situation of Proposition 8. We prove that $\mathscr{R}(U_\alpha)$ is dense in $C_0(E)$. An argument similar to the proof of the implication "strongly continuous $\Rightarrow$ (iv)" in Proposition 8 shows that the weak continuity of $(P_t)_{t\ge 0}$ implies

$$\forall f \in C_0(E)\ \forall x \in E: \quad \lim_{\alpha\to\infty} \alpha U_\alpha f(x) = f(x).$$

Note in addition $||\alpha U_\alpha f|| \le ||f||$. As a consequence, the α-independent range $\mathcal{R} = \mathscr{R}(U_\alpha)$ of the resolvent is dense in $C_0(E)$ with respect to pointwise convergence of uniformly bounded sequences. A functional analytic argument[6] allows us to deduce that $\mathcal{R}$ must be also dense in $C_0(E)$ with respect to uniform convergence. The strong continuity of $(P_t)_{t\ge 0}$ is then deduced by an argument similar to (iv) in Proposition 8. $\square$

5. From semigroup to generator

Here we prove the implications "$\Rightarrow$" in Theorem 5 and Proposition 6. First we explain how the resolvent is related to the infinitesimal generator.

Proposition 9. *Let $(P_t)_{t\ge 0}$ be a FD semi-group, $(U_\alpha)_{\alpha>0}$ its resolvent, and L its infinitesimal generator. Then L satisfies the positive maximum principle. Moreover for all $\alpha > 0$, we have $\mathscr{D}(L) = \mathscr{R}(U_\alpha)$, and for all $f \in \mathscr{D}(L)$ and $g \in C_0(E)$, we have*

$$g = (\alpha - L) f \iff f = U_\alpha g. \tag{4}$$

So the resolvent U_α is the inverse of the operator $\alpha I - L : \mathscr{D}(L) \to C_0(E)$, we write $U_\alpha = (\alpha - L)^{-1}$.

Proof. Step 1: L satisfies the positive maximum principle. Let $f \in \mathscr{D}(L)$ and $x_0 \in E$ a maximizer of f. Assume first $f(x_0) > 0$. Then $0 \le \frac{f_+}{f(x_0)} \le 1$ so by the sub-Markov property, $0 \le P_t \frac{f_+}{f(x_0)} \le 1$ i.e. $0 \le P_t f_+ \le f(x_0)$. If $f(x_0) = 0$, then $f \le 0$, $f_+ = 0$, and $0 \le P_t f_+ \le f(x_0)$ holds true as well. Therefore

$$P_t f = P_t(f_+ - f_-) \le P_t f_+ \le f(x_0)$$

[6]Suppose by contradiction that $\mathcal{R}$ is not dense with respect to uniform convergence, i.e., convergence with respect to $||\cdot||$. Then, by the Hahn-Banach theorem, there exists a bounded linear map $\varphi : C_0(E) \to \mathbb{R}$ that vanishes on $\mathcal{R}$ but not on all of $C_0(E)$; let $f_0 \in C_0(E)$ with $\varphi(f_0) \ne 0$. By the Riesz-Markov theorem, there exists a finite signed measure μ such that $\varphi(f) = \int_E f \mathrm{d}\mu$ for all $f \in C_0(E)$. Let $(g_n)_{n\in\mathbb{N}}$ be a uniformly bounded sequence in $\mathcal{R} \subset C_0(E)$ with $g_n \to f_0$ pointwise, then $0 = \varphi(g_n) = \int_E g_n \mathrm{d}\mu \to \int_E f \mathrm{d}\mu$ by the bounded convergence theorem for finite signed measure, hence $\varphi(f) = 0$, contradiction.

for all $t \geq 0$ and

$$(Lf)(x_0) = \lim_{t \searrow 0} \frac{(P_t f)(x_0) - f(x_0)}{t} \leq 0.$$

Step 2: $\mathscr{R}(U_\alpha) \subset \mathscr{D}(L)$. Let $f \in \mathscr{R}(U_\alpha)$ and $g \in C_0(E)$ with $f = U_\alpha g$. Then

$$P_t f(x) = \int_0^\infty \mathrm{e}^{-\alpha s} P_{s+t} g(x) \mathrm{d}s = \mathrm{e}^{\alpha t} \int_t^\infty \mathrm{e}^{-\alpha u} P_u g(x) \mathrm{d}u,$$

so we expect the derivative at time t to be

$$\varphi_t = \alpha \mathrm{e}^{\alpha t} \int_t^\infty \mathrm{e}^{-\alpha u} P_u g(x) \mathrm{d}u + \mathrm{e}^{\alpha t} \big(-\mathrm{e}^{-\alpha t} P_t g(x) \big).$$

We are only interested in $t = 0$ and note

$$\varphi_0 = \alpha U_\alpha g - g = \alpha f - g.$$

Remember from the proof of Proposition 8 that

$$P_t f(x) - f(x) = (\mathrm{e}^{\alpha t} - 1) \int_t^\infty \mathrm{e}^{-\alpha s} P_s g(x) \mathrm{d}s - \int_0^t \mathrm{e}^{-\alpha s} P_s g(x) \mathrm{d}s.$$

It follows that

$$\begin{aligned} P_t f(x) - f(x) - t\varphi_0(x) = \big(\mathrm{e}^{-\alpha t} - 1 - \alpha t \big) \int_t^\infty \mathrm{e}^{-\alpha s} P_s g(x) \mathrm{d}s - \alpha t \int_0^t \mathrm{e}^{-\alpha t} P_s g(x) \mathrm{d}s \\ - \int_0^t \mathrm{e}^{-\alpha s} P_s g(x) \mathrm{d}s + t g(x) \end{aligned}$$

and

$$||\frac{1}{t}(P_t f - f) - \varphi_0|| \leq \Big(\frac{\mathrm{e}^{\alpha t} - 1}{t} - \alpha \Big) ||U_\alpha g|| + \alpha t ||g|| + t \sup_{s \in [0,t]} ||\mathrm{e}^{-\alpha s} P_s g - g||,$$

which goes to zero as $t \to 0$ because of the strong continuity of $(P_t)_{t \geq 0}$. It follows that $f \in \mathscr{D}(L)$ and $Lf = \varphi_0$, moreover $(\alpha - L)f = \alpha f - \varphi_0 = g$.

Step 3: $\mathscr{D}(L) \subset \mathscr{R}(U_\alpha)$. Let $f \in \mathscr{D}(L)$. Set $g := (\alpha - L)f$. We would like to show $U_\alpha g = f$. By Step 1, the operator L satisfies the positive maximum principle. It follows that L is dissipative and $\alpha - L$ is injective. Consequently it is enough to s check $(\alpha - L)U_\alpha g = (\alpha - L)f$ i.e. $(\alpha - L)U_\alpha g = g$. Let $h := U_\alpha g$. We already know from Step 2 that $h \in \mathscr{D}(L)$ and $(\alpha - L)h = g$. Thus $g = (\alpha - L)h = (\alpha - L)U_\alpha g$, which is the inequality that we wanted to have. It follows that $\mathscr{D}(L) \subset \mathscr{R}(U_\alpha)$.

Step 4: The equivalence (4) *holds true.* Let $f \in \mathscr{D}(L)$ and $\alpha > 0$. If $f = U_\alpha g$, then $(\alpha - L)f = g$ by the considerations from Step 2. Conversely, if $g = (\alpha - L)f$, then $f = U_\alpha g$ by the considerations from Step 3. □

Proof of "⇒" in Theorem 5. By Proposition 9, we have $\mathscr{D}(L) = \mathscr{R}(U_\alpha)$, which is dense in $C_0(E)$ because of the strong continuity of $(P_t)_{t \geq 0}$ and Proposition 8(iv). Therefore $\mathscr{D}(L)$ is dense in $C_0(E)$. The positive maximum principle has been proven in Proposition 9. For the surjectivity of $\alpha I - L$, given $g \in C_0(E)$, let $f := U_\alpha g$. Then by Eq. (4), we have $g = (\alpha - L)f \in \mathscr{R}(\alpha I - L)$. Thus $\alpha I - L$ is surjective, actually for all $\alpha > 0$. □

Proof of "⇒" in Proposition 6. Let $(P_t)_{t \geq 0}$ be a FD semi-group with generator L. Assume that $(P_t f)(x) = \int_E T_t(x, \mathrm{d}y) f(y)$ for some Markovian transition function $(T_t)_{t \geq 0}$, so that $T_t(x, E) = 1$. If E is compact, then the constant function $\mathbf{1}$ is in $C_0(E) = C_b(E) = C(E)$ and $(P_t \mathbf{1})(x) = \int_E T_t(x, \mathrm{d}y) 1 = T_t(x, E) = 1$ hence $P_t \mathbf{1} = \mathbf{1}$. If E is not compact, then $\mathbf{1} \notin C_0(E)$ but by the properties of lccb spaces listed in Appendix A, there exists a sequence $(f_n)_{n \in \mathbb{N}}$ of non-negative functions in $C_0(E)$ such that $f_n \nearrow 1$ pointwise on E. Such a sequence is necessarily uniformly bounded since $\sup_n ||f_n|| \leq 1 < \infty$. Moreover $(P_t f_n)(x) = \int_E T_t(x, \mathrm{d}y) f_n(y) \to \int_E T_t(x, \mathrm{d}y) = 1$ by dominated convergence. □

Next we prove an analogue of the Kolmogorov backward and forward equations, and show how to go directly from resolvent to generator. This is not needed for the proofs of Theorem 5 but is of interest on its own.

Proposition 10. *Let $(P_t)_{t\geq 0}$ be a FD semi-group and L its infinitesimal generator. Then for all $f \in \mathscr{D}(L)$, the function $\mathbb{R}_+ \ni t \mapsto P_t f$ is differentiable, takes values in $\mathscr{D}(L)$, and satisfies*

$$\frac{\mathrm{d}}{\mathrm{d}t} P_t f = P_t L f = L P_t f.$$

Moreover for all $f \in C_0(E)$ and $t > 0$, we have

$$P_t f = \lim_{n\to\infty} \Big(I - \frac{t}{n} L\Big)^{-n} f = \lim_{n\to\infty} \Big(\frac{n}{t} U_{n/t}\Big)^n f \tag{5}$$

and the semi-group is uniquely determined by its generator.

Proof. Because of the semi-group property $P_{t+h} = P_t P_h = P_h P_t$, we have for all $t \geq 0$ and $h > 0$

$$\frac{1}{h}(P_{t+h} f - P_t f) = P_t\Big(\frac{1}{h}(P_h f - f)\Big) = \frac{1}{h}\big(P_h(P_t f) - P_t f\big).$$

By the strong continuity, the middle expression converges in norm to $P_t L f$ as $h \searrow 0$. Therefore the limit of the right expression exists and is equal to $P_t L f$. It follows that $P_t f \in \mathscr{D}(L)$ and $L P_t f = P_t L f$. and the right expression converges to $L P_t f$. The limit of the left expression exists as well. Consequently the map $\mathbb{R}_+ \ni t \mapsto P_t f$ has right derivatives everywhere, and the right derivative is given by the continuous map $t \mapsto L P_t f = P_t L f$. Proceeding as in Exercise 1(a) from Sheet 10 (or Lemma 3.3.2 in [Scheutzow, StochMod]), we deduce that $t \mapsto P_t f$ is in fact differentiable with derivative $P_t L f = L P_t f$.

An induction over $n \in \mathbb{N}$ shows that for all $\alpha > 0$ and $n \in \mathbb{N}$, we have

$$\Big(I - \frac{1}{\alpha} L\Big)^n f = \alpha^n U_\alpha^n f = \int_0^\infty \frac{\alpha^n s^{n-1}}{(n-1)!} \mathrm{e}^{-\alpha s} P_s f \mathrm{d}s. \tag{6}$$

The function $\mathbb{R}_+ \ni s \mapsto \frac{\alpha^n s^{n-1}}{(n-1)!} \exp(-\alpha s)$ is the probability density function of the *Gamma distribution* with parameters n and α, which is equal to the distribution of the sum of n i.i.d. $\mathrm{Exp}(\alpha)$-variables $\tau_1, \ldots, \tau_n$ (exercise!). For $\alpha = n/t$, we may write $\tau_i = \frac{t}{n} T_i$ with $T_1, \ldots, T_n$ i.i.d. $\mathrm{Exp}(1)$ variables. The probabilistic interpretation of (6) is that

$$(I - \tfrac{t}{n} L)^n f(x) = \mathbb{E}\Big[P_{\frac{t}{n}(T_1 + \cdots + T_n)} f(x)\Big]$$

or, if we are already given a Markov family $\mathbb{X}$ that goes with (P_t) and assume that $T_1, \ldots, T_n$ live on the same space Ω and are independent of $(X_t)_{t\geq 0}$,

$$(I - \tfrac{t}{n} L)^n f(x) = \mathbb{E}_x\Big[f\big(X_{\frac{t}{n}(T_1 + \cdots + T_n)}\big)\Big].$$

Assume that $f \in \mathscr{D}(L)$. Then by the differentiability of $t \mapsto P_t f$ and $||\frac{\mathrm{d}}{\mathrm{d}t} P_t f|| = ||P_t L f|| \leq ||L f||$, we have

$$||P_t f - P_s f|| \leq ||L f||\, |t - s|$$

for all $s, t \geq 0$. It follows that

$$||(I - \tfrac{t}{n} L)^n f - P_t f|| \leq ||L f||\, |t| \Big|\mathbb{E}\Big[\frac{1}{n}(T_1 + \cdots + T_n\Big] - 1\Big|$$

which goes to zero by the law of large numbers. This proves the required convergence in the case $f \in \mathscr{D}(L)$. The convergence for $f \in C_0(E)$ follows by an $\varepsilon/3$-argument, exploiting the density of $\mathscr{D}(L)$, the inequality

$$||(I - \tfrac{t}{n} L)^n f - P_t f|| \leq ||(\tfrac{n}{t} U_{n/t})^n (f - g)|| + ||(I - \tfrac{t}{n} L)^n g - P_t g|| + ||P_t (g - f)||$$

valid for all $f, g \in C_0(E)$, and the contractivity of αU_α and P_t.

Finally if $(\tilde{P}_t)_{t\geq 0}$ is another FD semi-group with generator $\tilde{L} = L$, then Eq. (5) shows that $\tilde{P}_t = P_t$ for all $t > 0$. □

6. From generator to semigroup

From now on let $L : \mathscr{D}(L) \to C_0(E)$ be an operator that satisfies conditions (i), (ii), and (iii) from Theorem 5. Our first task is to verify that L is associated with a strongly continuous contraction resolvent.

6.1. **From generator to resolvent.** As a preliminary observation, we note that the dissipativity (2) shows that for all $\alpha > 0$, the operator $\alpha - L : \mathscr{D}(L) \to C_0(E)$ is injective, moreover the inverse operator $(\alpha - L)^{-1} : \mathscr{R}(\alpha - L) \to \mathscr{D}(L)$ satisfies

$$\alpha||(\alpha - L)^{-1}g|| \leq ||g||. \tag{7}$$

We would like to define $U_\alpha g := (\alpha - L)^{-1}$ and have $U_\alpha g$ well-defined for all $g \in C_0(E)$, so we have to check that $\mathscr{R}(\alpha - L) = C_0(E)$, i.e., $\alpha - L$ is surjective for *all* $\alpha > 0$. So far we only know this is true for (at least) *one* $\alpha > 0$!

Lemma 11. *The operator L is* closed*, i.e., for every sequence $(f_n)_{n\in\mathbb{N}}$ in $\mathscr{D}(L)$ with the property that $||f_n - f|| \to 0$ and $||Lf_n - g|| \to 0$ for some $f, g \in C_0(E)$, we must have $f \in \mathscr{D}(L)$ and $g = Lf$.*

Proof. The proof only uses the surjectivity of $(\lambda I - L)$ for some $\lambda > 0$ and the dissipativity of L. Let $\lambda > 0$ such that $\mathscr{R}(\lambda - L) = C_0(E)$. If $f_n \to f$ and $Lf_n \to g$, then $\lambda f_n - Lf_n \to \lambda f - g$. By the surjectivity of $\lambda - L$, there exists some $\varphi \in \mathscr{D}(L)$ such that $\lambda f - g = \lambda\varphi - L\varphi$. By the dissipativity (2),

$$\lambda||f_n - \varphi|| \leq ||(\lambda - L)(f_n - \varphi)|| = ||(\lambda f_n - Lf_n) - (\lambda f - g)|| \to 0$$

hence $||f_n - \varphi|| \to 0$. It follows that $f = \varphi$ and $g = L\varphi = Lf$. □

Lemma 12. *For all $\alpha > 0$, the space $\mathscr{R}(\alpha - L)$ is a closed subspace of $C_0(E)$.*

Proof. The proof only uses that L is closed (by Lemma 11) and dissipative. Let $(g_n)_{n\in\mathbb{N}}$ be a sequence in $\mathscr{R}(\alpha - L)$ that converges to some $g \in C_0(E)$. Write $g_n = (\alpha - L)f_n$ with $f_n \in \mathscr{D}(L)$. By the dissipativity of L, we have

$$||g_n - g_m|| = ||(\alpha - L)(f_n - f_m)|| \geq \alpha||f_n - f_m||$$

for all $m, n \in \mathbb{N}$. It follows that $(f_n)_{n\in\mathbb{N}}$ is a Cauchy sequence and admits a limit $f \in C_0(E)$. Since L is closed, we conclude that the limit f is in $\mathscr{D}(L)$ and $g = (\alpha - L)f \in \mathscr{R}(\alpha - L)$. □

Lemma 13. $\mathscr{R}(\alpha - L) = C_0(E)$ *for all* $\alpha > 0$.

Proof. Let

$$J := \{\alpha \in (0, \infty) \mid \mathscr{R}(\alpha - L) = C_0(E)\}.$$

We have to show $J = (0, \infty)$ and do this by proving that J is both open and closed in $(0, \infty)$. "Closed in $(0, \infty)$" means "for every sequence $(\alpha_n)_{n\in\mathbb{N}}$ in J that has a limit $\alpha \in (0, \infty)$, we must have $\alpha \in J$." For example, the set $(0, 1]$ is closed in $(0, \infty)$!

Step 1: J is closed. Let $(\alpha_n)_{n\in\mathbb{N}}$ be a sequence of strictly positive numbers in J that converges to some $\alpha > 0$. For $g \in C_0(E)$, let $g_n := (\alpha - L)(\alpha_n - L)^{-1}g$. Notice that g_n is well-defined because $(\alpha_n - L)^{-1}$ is an operator from $\mathscr{R}(\alpha_n - L) \to \mathscr{D}(L)$. Clearly $g_n \in \mathscr{R}(\alpha - L)$. In view of $\alpha_n \to \alpha$, it seems plausible that $g_n \to g$. For a proof, we exploit the dissipativity of L: first note

$$g_n - g = \big[(\alpha - \alpha_n) + (\alpha_n - L)\big](\alpha_n - L)^{-1}g - g = (\alpha - \alpha_n)(\alpha_n - L)^{-1}g$$

and then by (7) and $\alpha_n \to 0$,

$$||g_n - g|| \leq \frac{|\alpha - \alpha_n|}{\alpha}||g|| \to 0.$$

It follows that each $g \in C_0(E)$ is the limit of a sequence g_n of elements in $\mathscr{R}(\alpha - L)$, i.e., $\mathscr{R}(\alpha - L)$ is dense in $C_0(E)$. By Lemma 12, the range is closed, so we find $\mathscr{R}(\alpha - L) = \overline{\mathscr{R}(\alpha - L)} = C_0(E)$ and $\alpha \in J$.

Step 2: J is open. Because of the injectivity of $\alpha - L$ and the bound (7), an element $\alpha > 0$ is in J if and only if $(\alpha - L)$ is a bijective map from $\mathscr{D}(L)$ to $C_0(E)$ with continuous inverse, i.e., if it is in the *resolvent set* $\rho(L)$ of L: we have $J = \rho(A) \cap (0, \infty)$. It is known from functional analysis that the resolvent set is open, so J must be open as well.

Step 3: $J = (0,\infty)$. By steps 1 and 2, J is both open and closed in $(0,\infty)$. Condition (iii) from Theorem 5 guarantees that J is not empty. But because $(0,\infty)$ is connected, the only non-empty subset that is both open and closed is $(0,\infty)$ itself.[7] □

Now we have everything we need to check that $U_\alpha := (\alpha - L)^{-1}$ forms a strongly continuous sub-Markov contraction resolvent.

Proposition 14. *For $\alpha > 0$, define $U_\alpha : C_0(E) \to \mathscr{D}(L) \subset C_0(E)$ by $U_\alpha := (\alpha - L)^{-1}$. Then $(U_\alpha)_{\alpha>0}$ satisfies properties (i)—(iv) from Proposition 8.*

Proof. (i) Clearly each U_α is a linear map from $C_0(E)$ to $C_0(E)$.

(ii) The resolvent identity is a fact from functional analysis (exercise: verify it by hand!).

(iii) Let $f \in C_0(E)$ with $0 \le f \le 1$ and $g := (\alpha - L)^{-1} f \in \mathscr{D}(L)$. We want to prove that $0 \le \alpha g \le 1$. Assume first that g has a maximizer x_0. Then $(Lg)(x_0) \le 0$, hence

$$1 \ge \inf f \ge f(x_0) = \alpha g(x_0) - Lg(x_0) \ge \alpha g(x_0) \ge \alpha \sup g.$$

If g has no maximizer, then because of $g \in C_0(E)$ we must have $\sup g = 0$ and the inequality $1 \ge \inf f \ge \alpha \sup g$ holds true as well. Thus $g \le 1$. A similar argument applied to $-g$ instead of g yields

$$0 \le \sup f \le \alpha \inf g,$$

hence $g \ge 0$.

(iv) For the strong continuity $||\alpha U_\alpha g - g|| \to 0$, we exploit again the dissipativity. If $g \in \mathscr{D}(L)$, then

$$||\alpha U_\alpha g - g|| = \alpha ||U_\alpha g - \tfrac{1}{\alpha} g|| \le ||(\alpha - L)(U_\alpha g - \tfrac{1}{\alpha} g)|| = \tfrac{1}{\alpha} ||Lg|| \to 0 \quad (\alpha \to \infty).$$

For general $g \in C_0(E)$, the claim follows from the density of $\mathscr{D}(L)$, the contractivity of αU_α (which in turn follows from (iii)) and an $\varepsilon/3$-argument. □

6.2. **Yosida approximation.** If L is a *bounded operator*, i.e., $\mathscr{D}(L) = C_0(E)$ and $||Lf|| \le c||f||$ for some $c > 0$ and all $f \in C_0(E)$, we can define a semigroup by the exponential series,

$$P_t f := f + \sum_{n=1}^\infty \frac{t^n}{n!} L^n f = \exp(tL) f.$$

The exponential series is absolutely convergent because $||L^n f|| \le c^n ||f||$ for all n and

$$||f|| + \sum_{n=1}^\infty \frac{t^n}{n!} ||L^n f|| \le \exp(tc) ||f||.$$

The trouble is that in general, L is not bounded. The idea is to approximate L by a family of bounded operators L_α. The *Yosida approximation* is

$$L_\alpha := \alpha L (\alpha - L)^{-1} = L\big(I - \frac{1}{\alpha} L\big)^{-1}.$$

Notice that

$$L_\alpha = (\alpha^2 + \alpha L - \alpha^2)(\alpha - L)^{-1} = \alpha(\alpha U_\alpha - I).$$

Lemma 15.

(a) *For each $\alpha > 0$, the operator L_α is bounded and $(\exp(tL_\alpha))_{t\ge0}$ is a FD semi-group.*
(b) *For all $\alpha, \beta > 0$, $L_\alpha L_\beta = L_\beta L_\alpha$.*
(c) *For each $f \in \mathscr{D}(L)$, $||L_\alpha f - Lf|| \to 0$ as $\alpha \to \infty$.*
(d) *For all $\alpha, \beta > 0$, $t \ge 0$, and $f \in C_0(E)$: $||\exp(tL_\alpha) f - \exp(tL_\beta) f|| \le t\, ||L_\alpha f - L_\beta f||$.*

[7] This is in fact the general topological *definition* of connectedness. If it feels too abstract to you, you can also check this by hand: let $m := \inf J$ and $M := \sup J$. Show that if J is both closed and open in $(0,\infty)$, then $m > 0$ or $M < \infty$ lead to a contradiction (exercise!).

Proof. (a) $\mathscr{D}(L_\alpha) = C_0(E)$ holds true by definition, moreover $||L_\alpha f|| \leq \alpha(||\alpha U_\alpha f|| + ||f||) \leq 2\alpha ||f||$ for all $f \in C_0(E)$. Thus L_α is bounded. Therefore we can define $\exp(tL_\alpha)$ by the exponential series. The family $(\exp(tL_\alpha))_{t\geq 0}$ clearly satisfies (i) and (ii) in Definition 1, we leave the proof of the strong continuity (iv) as an exercise. For the sub-Markov property (iii): Because of $L_\alpha = \alpha^2 U_\alpha - \alpha I$, using $\exp(A+B) = \exp(A)\exp(B) = \exp(B)\exp(A)$ whenever A, B are bounded operators with $AB = BA$, we get

$$\exp(tL_\alpha) = \exp(-\alpha t)\exp(t\alpha^2 U_\alpha).$$

Let $f \in C_0(E)$ with $0 \leq f \leq 1$, then $0 \leq \alpha U_\alpha \leq 1$ by Proposition 14. An induction over n yields $0 \leq (\alpha^2 U_\alpha)^n f \leq \alpha^n$ for all n and then

$$0 \leq \exp(tL_\alpha) f \leq \exp(-\alpha t)\sum_{n=0}^\infty \frac{(\alpha t)^n}{n!} = 1.$$

This concludes the proof of (a).

The commutativity (b) follows because, by the resolvent identity, $U_\alpha U_\beta = U_\beta U_\alpha$ for all $\alpha, \beta > 0$. For the convergence (c), we note that $L_\alpha f = \alpha U_\alpha L f$ whenever $f \in \mathscr{D}(L)$ and conclude with the strong continuity of the resolvent $||\alpha U_\alpha g - g|| \to 0$, applied to $g = Lf$. For (d) we note

$$\exp(tL_\alpha)f - \exp(tL_\beta)f = \int_0^t \frac{\mathrm{d}}{\mathrm{d}s} \mathrm{e}^{sL_\alpha}\mathrm{e}^{(t-s)L_\beta} f \mathrm{d}s = \int_0^t \mathrm{e}^{sL_\alpha}(L_\alpha - L_\beta)\mathrm{e}^{(t-s)L_\beta} f \mathrm{d}s$$

and the claim follows from the triangle inequality and the contractivity of $\exp(tL_\alpha)$ and $\exp(tL_\beta)$. □

Now we have everything we need to show how to go from generator to semi-group.

Proof of "$\Leftarrow$" in Theorem 5. Let L be a densely defined operator that satisfies the positive maximum principle and such that $(\alpha - L)$ is surjective for at least one α (hence by Lemma 13, all $\alpha > 0$). Define $U_\alpha := (\alpha - L)^{-1}$ and let L_α be the Yosida approximation defined above. We show that the limit

$$P_t f := \lim_{\alpha\to\infty} \exp(tL_\alpha) f \tag{8}$$

exists for all $f \in C_0(E)$ and $t > 0$, and that $(P_t)_{t\geq 0}$ defines a FD semi-group with infinitesimal generator L.

The existence of the limit (8) for $f \in \mathscr{D}(L)$ follows from Lemma 15(d). The limit inherits the contractivity $||P_t f|| \leq ||f||$ from $\exp(tL_\alpha)$. It follows that $||P_t f|| \leq ||f||$ on $\mathscr{D}(L)$. An $\varepsilon/3$-argument, together with the density of $\mathscr{D}(L)$ in $C_0(E)$, allows us to conclude the existence of the limit (8) for all $f \in C_0(E)$. P_t is a linear map in $C_0(E)$ by definition, the properties $P_0 = I$, $P_{t+s} = P_t P_s$, and $0 \leq f \leq 1 \;\Rightarrow 0 \leq P_t f \leq 1$ are inherited from $\exp(tL_\alpha)$. For the strong continuity, we bound

$$||P_t f - f|| \leq ||P_t f - \mathrm{e}^{tL_\alpha} f|| + ||\mathrm{e}^{tL_\alpha} f - f||.$$

If $f \in \mathscr{D}(L)$, we get from Lemma 15(d) that

$$||P_t f - \exp(tL_\alpha) f|| \leq t\,||Lf - L_\alpha f||.$$

Given $\varepsilon > 0$ and $f \in \mathscr{D}(L)$, we first choose α large enough so that $||Lf - L_\alpha f|| \leq \varepsilon$ and then $t_0 \in (0,1)$ small enough so that $||\exp(tL_\alpha)f - f|| \leq \varepsilon$ for all $t \in [0, t_0]$. Then $||P_t f - f|| \leq 2\varepsilon$ for all $t \in [0, t_0]$. Hence $||P_t f - f|| \to 0$. For general $f \in C_0(E)$, the statement is deduced from an $\varepsilon/3$-argument and the density of $\mathscr{D}(L)$ in $C_0(E)$. Thus $(P_t)_{t\geq 0}$ is a FD semi-group.

It remains to show that the infinitesimal generator of $(P_t)_{t\geq 0}$ is L. We start from the identity

$$\exp(tL_\alpha) f = f + \int_0^t \exp(sL_\alpha) L_\alpha f \mathrm{d}s.$$

The left side converges to $P_t f$. For $f \in \mathscr{D}(L)$, the right side converges to $\int_0^t P_s L f \mathrm{d}s$ (check!). It follows that

$$\frac{1}{t}(P_t f - f) = \frac{1}{t}\int_0^t P_s L f \mathrm{d}s \to Lf \quad (t \searrow 0).$$

Let us write G (instead of L) for the infinitesimal generator of (P_t). We have just checked that $\mathscr{D}(L) \subset \mathscr{D}(G)$ and for all $f \in \mathscr{D}(L)$, $Gf = Lf$. The proof is complete once we show $\mathscr{D}(G) = \mathscr{D}(L)$. To that aim we note that the considerations from Section 5 apply to G and yield that $\alpha - G : \mathscr{D}(G) \to C_0(E)$ is bijective. The map $\alpha - L : \mathscr{D}(L) \to C_0(E)$ is bijective as well because of the dissipativity and by Lemma 13. Let $f \in \mathscr{D}(G)$ and $\tilde{f} := (\alpha - L)^{-1}(\alpha - G)f \in \mathscr{D}(L)$. Then $(\alpha - G)\tilde{f} = (\alpha - L)\tilde{f} = (\alpha - G)f$ hence by the injectivity of $\alpha - G$, $\tilde{f} = f$. Therefore $f = \tilde{f} \in \mathscr{D}(L)$. This proves $\mathscr{D}(G) \subset \mathscr{D}(L)$. □

Proof of "⇐" in Proposition 6. If E is compact and $\mathbf{1} \in \mathscr{D}(L)$, $L\mathbf{1} = 0$: then $\frac{\mathrm{d}}{\mathrm{d}t} P_t \mathbf{1} = P_t L \mathbf{1} = 0$ by Proposition 10, hence $P_t \mathbf{1} = P_0 \mathbf{1} = \mathbf{1}$ for all $t \geq 0$. Notice that $\alpha U_\alpha \mathbf{1} = \mathbf{1}$ for all $\alpha > 0$.

If E is lccb but not compact: Fix $\alpha > 0$. Let $(f_n)_{n\in\mathbb{N}}$, $(g_n)_{n\in\mathbb{N}}$ be sequences in $\mathscr{D}(L)$ and $C_0(E)$, respectively, such that $g_n = \frac{1}{\alpha}(\alpha - L) f_n$, $\sup_n ||g_n|| < \infty$, and (f_n) and (g_n) both converge pointwise to $\mathbf{1}$. Then $\alpha U_\alpha g_n = f_n \to 1$.

On the other hand, let $(T_t)_{t\geq 0}$ be the sub-Markov transition function such that $(P_t f)(x) = \int_E T_t(x, \mathrm{d}y) f(y)$ for all $f \in C_0(E)$ and $t \geq 0$. Then

$$\alpha U_\alpha g_n = \int_0^\infty \alpha \mathrm{e}^{-\alpha s} \left(\int_E T_s(x, \mathrm{d}y) g_n(y) \right) \mathrm{d}s \to \int_0^\infty \alpha \mathrm{e}^{-\alpha s} T_s(x, E) \mathrm{d}s$$

by dominated convergence and because $g_n \to \mathbf{1}$. It follows that

$$1 = \int_0^\infty \alpha \mathrm{e}^{-\alpha s} T_s(x, E) \mathrm{d}s$$

for all $x \in E$. Because of $T_s(x, E) \leq 1$, it follows that $T_s(x, E) = 1$ for Lebesgue-almost all $s \geq 0$. Suppose by contradiction that there exists some $s_0 > 0$ such that $T_{s_0}(x, E) < 1$. Then by the Chapman-Kolmogorov equation, for all $h > 0$,

$$T_{s_0+h}(x, E) = \int_E T_{s_0}(x, \mathrm{d}y) T_h(y, E) \leq \int_E T_{s_0}(x, \mathrm{d}y) = T_{s_0}(x, E) < 1$$

hence $T_s(x, E) < 1$ for all $s \geq s_0$, in contradiction with $T_s(x, E) = 1$ for Lebesgue-almost all $s \geq 0$. Hence $T_s(x, E) = 1$ for all $s \geq 0$ and $x \in E$, from which one readily deduces that $(T_t)_{t\geq 0}$ and $(P_t)_{t\geq 0}$ are Markovian. □

6.3. **Probabilistic interpretation of the Yosida approximation.** The construction of the previous section admit a nice probabilistic interpretation. Assume for simplicity that E is compact (so that $\mathbf{1} \in C_0(E)$) and that we are in the Markovian situation $L\mathbf{1} = 0$. Then $\alpha U_\alpha \mathbf{1} = \alpha(\alpha - L)^{-1} \mathbf{1} = \mathbf{1}$. By the Riesz-Markov theorem, there exists a kernel $\Pi_\alpha(x, \mathrm{d}y)$ with $K_\alpha(x, E) = 1$ for all $x \in E$ such that

$$\alpha U_\alpha f(x) = \int_E \Pi_\alpha(x, \mathrm{d}y) f(y).$$

Notice

$$(L_\alpha f)(x) = \alpha(\alpha U_\alpha - I) f(x) = \alpha \int_E \Pi_\alpha(x, \mathrm{d}y)(f(y) - f(x)),$$

which should remind you of the form of generator of jump processes discussed earlier. If E is finite, then $L_\alpha = \alpha(\Pi_\alpha - I)$ is a conservative Q-matrix. Define

$$P_t^\alpha := \exp(t L_\alpha).$$

By the considerations in the proof of Lemma 15(a), we have

$$\begin{aligned} P_t^\alpha f(x) &= \mathrm{e}^{-t\alpha} \sum_{n=0}^\infty \frac{(\alpha t)^n}{n!} (\alpha U_\alpha)^n f(x) \\ &= \mathrm{e}^{-\alpha t} \Big\{ f(x) + \sum_{n=1}^\infty \frac{(\alpha t)^n}{n!} \int_{E^n} \Pi_\alpha(x, \mathrm{d}x_1) \Pi_\alpha(x_1, \mathrm{d}x_2) \cdots \Pi_\alpha(x_{n-1}, \mathrm{d}x_n) f(x_n) \Big\}. \end{aligned}$$

Since Π_α is a Markov kernel, the n-fold integral is best interpreted with a discrete-time Markov chain: Let $(Z_n^\alpha)_{n\geq 0}$ be a Markov chain with transition kernel Π_α, assume that there is a family of probability measures $\mathbb{P}_x$ with $Z_n^\alpha = x$, $\mathbb{P}_x$-a.s. Then

$$P_t^\alpha f(x) = \mathrm{e}^{-\alpha t} \sum_{n=0}^\infty \frac{(\alpha t)^n}{n!} \mathbb{E}_x\big[f(Z_n^\alpha)\big].$$

The summation over n is best interpreted as a sum over a number of jumps in $[0,t)$. In fact we recognize the distribution of a Poisson process $(N_t^\alpha)_{t\geq 0}$, whose holding times are i.i.d. $\mathrm{Exp}(\alpha)$-variables T_i^α.

Therefore $(P_t^\alpha)_{t\geq 0}$ is the transition function of a Markov process constructed as follows: let $T_1^\alpha, T_2^\alpha, \ldots$ be i.i.d. $\mathrm{Exp}(\alpha)$-variables T_i, and $(Z_n^\alpha)_{n\in\mathbb{N}_0}$ a discrete-time Markov process with transition kernel Π_α, independent of the T_i^α's. Then we define $X_t^\alpha := Z_0$ on $t \in [0, T_1^\alpha)$ and

$$X_t^\alpha := Z_n^\alpha, \quad t \in T_1^\alpha + \cdots + T_n^\alpha \leq t < T_1^\alpha + \cdots + T_{n+1}^\alpha \qquad (n \in \mathbb{N})$$

and obtain a Markov process with transition function (P_t^α). Feller speaks of processes of *pseudo-Poisson type* [Fel71, Chapter X.2].

Remark. The convergence of semi-groups $P_t^\alpha f \to P_t f$ proven for the implication "$\Leftarrow$" of Theorem 5 can actually be pushed to a notion convergence of processes $X_t^\alpha \Rightarrow X_t$, see [EK86, Chapter 4.2, Theorem 2.5]. This provides an alternative to what we do in Section 7 below: Instead of proving existence of a Markov family with transition function (P_t) via Kolmogorov extension theorem & path regularization, we could start from the jump processes (X_t^α), prove that they converge in some sense as $\alpha \to \infty$, and that the limit provides us with a Markov family with the required properties. See [EK86, Chapter 4.2, Theorem 2.7].

Conversely, suppose that we are already given a FD family $\mathbb{X}$ with transition function $(P_t)_{t\geq 0}$. Assume that the underlying probability space Ω is big enough to carry variables $T_n^\alpha : \Omega \to \mathbb{R}_+$, $\alpha > 0$, $n \in \mathbb{N}$, such that for each $\alpha > 0$, and for each $x \in E$, the random variables $T_1^\alpha, T_2^\alpha, \ldots$ on the probability space $(\Omega, \mathcal{F}, \mathbb{P}_x)$ are independent $\mathrm{Exp}(\alpha)$-variables, independent from the process $(X_t)_{t\geq 0}$. The kernel Π_α introduced above satisfies

$$\Pi_\alpha(x, A) = (\alpha U_\alpha \mathbb{1}_A)(x) = \mathbb{P}_x(X_{T_1^\alpha} \in A) \qquad (x \in E, A \in \mathcal{E}).$$

More generally, defining $Z_0^\alpha := X_0$ and

$$Z_n^\alpha := X_{T_1^\alpha + \cdots + T_n^\alpha} \qquad (n \in \mathbb{N}),$$

we obtain a discrete-time Markov process with transition kernel Π_α. The process (X_t^α) defined by

$$X_t^\alpha := Z_n^\alpha = X_{T_1^\alpha + \cdots + T_n^\alpha}, \quad T_1^\alpha + \cdots + T_n^\alpha \leq t < T_1^\alpha + \cdots + T_{n+1}^\alpha \qquad (n \in \mathbb{N})$$

and $X_t^\alpha = Z_0^\alpha = X_0$ on $[0, T_1^\alpha)$ is a jump process with transition function (P_t^α). The sample paths are piecewise constant càdlàg functions that approximate the sample paths of (X_t), you should try to draw a picture that reminds you of Riemann integration.

Again the convergence of semi-groups can be pushed to a convergence of processes: By [EK86, Chapter 4.2, Theorem 2.5], for each $x \in E$, the distribution of $(X_t^\alpha)_{t\geq 0}$ under $\mathbb{P}_x$ converges weakly to the distribution of $(X_t)_{t\geq 0}$ under $\mathbb{P}_x$, where weak convergence refers to a suitable topology on the Skorokhod space $\mathcal{D}_E[0,\infty)$.

7. Path regularization. From semigroup to process

Here we sketch some ingredients for the proof of Theorem 7.

Step 1: From semi-group to transition function. We have already observed that, by the Riesz-Markov theorem, there is a unique transition function associated with the semi-group. By a slight abuse of language, we use the same letter P_t for the map $P_t : C_0(E) \to C_0(E)$ and the kernels $P_t(x, \mathrm{d}y)$. The transition function is normal, i.e., $P_0(x, \cdot) = \delta_x(\cdot)$ for all $x \in E$.

Step 2: Construction of a Markov family with the correct finite-dimensional distributions (fidis). Let $\tilde{\Omega} := E^{\mathbb{R}_+}$ be the set of *all* maps $\omega : \mathbb{R}_+ \to E$ (càdlàg is *not* assumed). Let $Y_t(\omega) := \omega(t)$. The space $\tilde{\Omega}$ is equipped with the product σ-algebra $\tilde{\mathcal{F}} = \sigma(Y_t, t \geq 0)$. Let $x \in E$. For $I = \{t_1, \ldots, t_n\}$ with $n \in \mathbb{N}$, $0 \leq t_1 < \ldots < t_n$, define the measure $\tilde{\mu}^x_I$ on $E^{\{t_1,\ldots,t_n\}}$ by

$$\begin{aligned}\tilde{\mu}^x_{\{t_1,\ldots,t_n\}}&(A_1 \times \cdots \times A_n) \\ &:= \int_{E^n} P_{t_1}(x, \mathrm{d}x_1)\mathbb{1}_{A_1}(x_1)P_{t_2-t_1}(x_1, \mathrm{d}x_2)\mathbb{1}_{A_2}(x_2)\cdots P_{t_n-t_{n-1}}(x_{n-1}, \mathrm{d}x_n)\mathbb{1}_{A_n}(x_n) \quad (9)\end{aligned}$$

for all $A_1, \ldots, A_n \in \mathcal{E}$. (You may ask whether a case distinction is needed for t_1, but it is not: if $t_1 = 0$, then $P_0(x, \mathrm{d}y) = \delta_x(\mathrm{d}y)$ because the transition function is normal, and the formula above agrees with the definition one would like to give.) The family $(\tilde{\mu}^x_I)_{I\subset\mathbb{R}_+,\#I<\infty}$ forms a consistent family (exercise!) so by the Kolmogorov extension theorem, there exists a unique probability measure $\tilde{\mathbb{P}}_x$ on $(\tilde{\Omega}, \tilde{\mathcal{F}})$ with fidis $\tilde{\mu}^x_I$. It is not difficult to check that $(\tilde{\mathbb{P}}_x)_{x\in E}$, $(Y_t)_{t\geq 0}$, with the canonical filtration and the usual definition of shifts, define a Markov family with transition function (P_t).

Step 3: Path regularization. Our next task is to find càdlàg modifications of $(Y_t)_{t\geq 0}$. This step is analogous to the application of Kolmogorov's continuity theorem in the construction of Brownian motion, however sample paths in general are not continuous so we need something else. The presentation given here is adapted from [RW94, Chapter III.7] and [Lig10, Chapter 3.3]. Relevant background on continuous-time supermartingales can be found in [RW94, Chapter II.5] and [Lig10, Chapter 1.9, Propositions 1.113 and 1.114].

Set $\mathbb{Q}_+ := \mathbb{Q} \cap [0, \infty)$. A map $\gamma : \mathbb{Q}_+ \to \mathbb{R}$ is called *regularizable* if the following two conditions are met:

(i) For each $t \geq 0$, the limit $\lim_{q\downarrow t, q\in\mathbb{Q}_+} \gamma(q) \in \mathbb{R}$ exists.
(ii) For each $t > 0$, the limit $\lim_{q\uparrow t, q\in\mathbb{Q}_+} \gamma(q) \in \mathbb{R}$ exists.

The notation $q \downarrow t$ and $q \uparrow t$ stand for $q \to t$ along $q > t$ and $q < t$, respectively. In particular, the value at $q = t$ is irrelevant even if t is rational. If γ is regularizable, we define $\gamma^{\mathrm{reg}} : \mathbb{R}_+ \to \mathbb{R}$ by

$$\gamma^{\mathrm{reg}}(t) := \lim_{q\downarrow t, t\in\mathbb{Q}_+} \gamma(q).$$

The regularized function γ^{reg} is càdlàg [RW94, Chapter II. 62, Theorem 62.13]. The following theorem, called *Doob's regularity theorem: part 1* by Rogers and Williams, is extremely useful.

Theorem 16. [RW94, Theorem 65.1] *Let $(M_t)_{t\in\mathbb{R}_+}$ be a supermartingale defined on some filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{R}_+}, \mathbb{P})$. Let*

$$G := \{\omega \in \Omega \mid \text{the map } \mathbb{Q}_+ \ni q \mapsto M_q(\omega) \text{ is regularizable}\}.$$

Then $G \in \mathcal{F}$ and $\mathbb{P}(G) = 1$. For $t \in \mathbb{R}_+$, define

$$Z_t(\omega) := \begin{cases} \lim_{q\downarrow t, q\in\mathbb{Q}_+} M_q(\omega), & \omega \in G, \\ 0, & \omega \notin G. \end{cases}$$

Then $(Z_t)_{t\geq 0}$ has càdlàg sample paths.

We would like to apply the theorem to the process $(Y_t)_{t\geq 0}$. The trouble is, of course, that (Y_t) need not be a supermartingale. The idea instead is to find a family of functions $h \in \mathscr{H}$ such that $(h(Y_t))_{t\geq 0}$ is a supermartingale. Then hopefully if the limit $\lim_{s\downarrow t, s\in\mathbb{Q}_+} h(Y_s)$ exists for all $h \in \mathscr{H}$, almost surely, the limit $\lim_{s\downarrow t, s\in\mathbb{Q}_+} Y_s$ exists as well, almost surely.

3.1 Supermartingales from resolvent. Fix $\alpha > 0$ (for example, $\alpha = 1$). Suppose that $h \in \mathscr{R}(U_\alpha)$ is of the form $h = U_\alpha g$ with $g \geq 0$. Then h is *α-supermedian*, i.e.,

$$\forall s \geq 0 : \quad 0 \leq \mathrm{e}^{-\alpha s} P_s h \leq h.$$

Indeed,

$$\mathrm{e}^{-\alpha s} P_s U_\alpha g = \int_0^\infty \mathrm{e}^{-\alpha(s+t)} P_{s+t} g \mathrm{d}t = \int_s^\infty \mathrm{e}^{-\alpha u} P_u g \mathrm{d}u \leq U_\alpha g = h.$$

As a consequence, for every $x \in E$, and all $0 \leq s < t$,

$$\tilde{\mathbb{E}}_x\big[\mathrm{e}^{-\alpha t} h(Y_t) \mid \tilde{\mathcal{F}}_s\big] = \mathrm{e}^{-\alpha t}(P_{t-s}h)(Y_s) \leq \mathrm{e}^{-\alpha s} h(Y_s) \quad \tilde{P}_x\text{-a.s.}$$

with $\tilde{\mathcal{F}}_t$ the canonical filtration for (Y_t). Thus $(\mathrm{e}^{-\alpha t}h(Y_t))_{t\geq 0}$ is a $(\tilde{\mathbb{P}}_x, (\tilde{\mathcal{F}}_t)_{t\geq 0})$-supermartingale. As a consequence, there exists a $\tilde{\mathbb{P}}_x$-null set $\mathscr{N}_{x,h}$ such that

$$\forall \omega \in \tilde{\Omega} \setminus \mathscr{N}_{x,h}: \quad \lim_{\substack{s \downarrow t \\ s \in \mathbb{Q}_+}} h(Y_s(\omega)) \quad \text{exists for all } t \in \mathbb{R}_+$$

and the analogous statement for limits $s \uparrow t$ holds true as well.

3.2 Countable convergence determining set $\mathscr{H}$. Let $g_0, g_1, \dots$ be a countable dense subset of $\{f \in C_0(E) \mid f \geq 0\}$ with $g_0 > 0$ on E (such a set $\{g_n\}_{n\in\mathbb{N}_0}$ exists in lccb spaces, the reason for the choice $g_0 > 0$ will become clear later on). Set $h_n := U_\alpha g_n$, and $\mathscr{H} := \{h_n \mid n \in \mathbb{N}_0\}$. Every function $f \in \mathscr{D}(L)$ can be written as a limit of functions $U_\alpha(h_{n_j} - h_{m_j})$ with $j \to \infty$, therefore if $h_n(x) = h_n(y)$ for all $n \in \mathbb{N}_0$, then $f(x) = f(y)$ for all $f \in \mathscr{D}(L)$ and then for all $f \in C_0(E)$ because $\mathscr{D}(L)$ is dense in $C_0(E)$. As a consequence, $\mathscr{H}$ *separates points*, i.e., for all $x, y \in E$, we have the implication

$$\Big(h(x) = h(y) \text{ for all } h \in \mathscr{H}\Big) \Rightarrow x = y.$$

Lemma 17. *Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in E such that for each $h \in \mathscr{H}$, the limit $y_h := \lim_{n\to\infty} h(x_n)$ exists. Assume in addition that E is compact or that there exists $h \in \mathscr{H}$ such that $\lim_{n\to\infty} h(x_n) \neq 0$. Then $x = \lim_{n\to\infty} x_n$ exists as well.*

Proof. Let $x, x' \in E$ be two accumulation points of the sequence. Then $h(x) = h(x')$ for all $h \in \mathscr{H}$ and therefore $x = x'$. Thus $(x_n)_{n\in\mathbb{N}}$ can have at most one accumulation point in E.

If E is compact, then every sequence has a convergent subsequence so $(x_n)_{n\in\mathbb{N}}$ has at least one accumulation point. So $(x_n)_{n\in\mathbb{N}}$ has exactly one accumulation point x. Suppose by contradiction that $\lim_{n\to\infty} x_n \neq x$. Then $(x_n)_{n\in\mathbb{N}}$ has a subsequence (x_{n_k}) staying in $E \setminus \mathcal{O}$ with $\mathcal{O}$ some open neighborhood of x. By compactness, (x_{n_k}) in turn has a convergent subsequence with limit $y \in E \setminus \mathcal{O}$. Thus $y \neq x$ is another accumulation point of (x_n), contradicting uniqueness of accumulation points.

If E is lccb but not compact, for example, $E = \mathbb{R}$, we need to rule out subsequences going to infinity. Let $h^* \in \mathscr{H}$ with $y^* = \lim_{n\to\infty} h^*(x_n) \neq 0$. Let $K := \{x \in E \mid |h^*(x)| \geq \frac{1}{2}|y^*|\}$. Then K is compact and there exists some n^* such that for all $n \geq n^*$, we have $x_n \in K$. From here on we can repeat the argument given above for compact E. $\square$

Concerning the additional condition in the lemma, we note that $h_0 > 0$: we already know $h_0 \geq 0$ (because resolvents map non-negative functions to non-negative functions), and $h_0 = U_\alpha g_0 = 0$ would imply $g_0 = 0$ because the resolvent is injective. Therefore the supermartingale $M_t := \mathrm{e}^{-\alpha t} h_0(Y_t)$ is strictly positive and the following proposition comes in handy.

Proposition 18. *Suppose $(M_t)_{t\geq 0}$ is a non-negative supermartingale. Then for every $t \in \mathbb{Q}_+$, the probability of the event $\{M_t > 0, \inf_{s\in\mathbb{Q}_+\cap[0,t]} M_s = 0\}$ vanishes.*

See [Lig10, Proposition 1.114].

For the strictly positive supermartingale that we are dealing with, the proposition implies that the event $\{\inf_{s\in\mathbb{Q}_+\cap[0,t]} M_s = 0\}$ has probability zero, so any limit along subsequences must be strictly positive.

3.3 Definition of a modified process $(\tilde{X}_t)_{t\geq 0}$. The set $\mathscr{N}_x := \bigcup_{h\in\mathscr{H}} \mathscr{N}_{x,h}$ is a countable union of $\tilde{\mathbb{P}}_x$-null sets, hence, a $\tilde{\mathbb{P}}_x$-null set itself. Combining the considerations above, we see that for every $\omega \in \{X_0 = x\} \setminus \mathscr{N}_x$, and all $t \geq 0$, the limit

$$\tilde{X}_t(\omega) := \lim_{s\downarrow t, s\in\mathbb{Q}_+} Y_t(\omega)$$

exists, and for all $t > 0$, the limit $\lim_{s\uparrow t, s\in\mathbb{Q}_+} Y_t(\omega)$ exists as well. Let

$$\mathscr{N} := \tilde{\Omega} \setminus \bigcup_{y\in E} \Big(\{Y_0 = y\} \setminus \mathscr{N}_y \Big) = \bigcup_{y\in E} \big(\mathscr{N}_y \cap \{Y_0 = y\} \big).$$

Notice that for each $x \in E$, $\mathscr{N}$ is contained in a $\tilde{\mathbb{P}}_x$-null set, since $\mathscr{N} \subset \mathscr{N}_x \cup \{Y_0 \neq x\}$ and $\tilde{\mathbb{P}}_x(Y_0 \neq x) = 0$. We extend the definition of $(\tilde{X}_t)$ to all of $\tilde{\Omega}$ by defining $\tilde{X}_t(\omega) := x_0$ for $\omega \in \mathscr{N}$, with $x_0 \in E$ arbitrary dummy element not depending on x, ω, or t (for example, $x_0 = 0$ when $E = \mathbb{R}$). The process $(\tilde{X}_t)_{t\geq 0}$ has càdlàg paths.

3.4 Show that $(\tilde{X}_t)_{t\geq 0}$ is a modification of $(Y_t)_{t\geq 0}$. Fix $x \in E$. We have, for all $t \geq 0$ and $f_1, f_2 \in C_0(E)$,

$$\tilde{\mathbb{E}}_x\big[f_1(Y_t) f_2(\tilde{X}_t)\big] = \lim_{\substack{s\downarrow t\\ s\in\mathbb{Q}_+}} \tilde{\mathbb{E}}_x\big[f_1(Y_t) f_2(Y_s)\big] = \lim_{\substack{s\downarrow t\\ s\in\mathbb{Q}_+}} \tilde{\mathbb{E}}_x\big[f_1(Y_t)(P_{s-t} f_2)(Y_s)\big] = \mathbb{E}_x\big[f_1(Y_t) f_2(Y_t)\big].$$

A functional monotone class theorem shows that $\tilde{\mathbb{E}}_x[f(Y_t, \tilde{X}_t)] = \tilde{\mathbb{E}}_x[f(Y_t, Y_t)]$ for all bounded measurable $f : E \times E \to \mathbb{R}$. It follows that

$$\forall x \in E\ \forall t \geq 0 : \quad \tilde{\mathbb{P}}_x\big(\tilde{X}_t = Y_t\big) = 1.$$

Thus $(\tilde{X}_t)_{t\geq 0}$ is a modification of $(Y_t)_{t\geq 0}$, relative to each $\tilde{\mathbb{P}}_x$.

As a consequence, $(\tilde{X}_t)_{t\geq 0}$ has the same finite-dimensional distributions as $(Y_t)_{t\geq 0}$, hence $\tilde{X}_0 = Y_0 = x$, $\tilde{\mathbb{P}}_x$-a.s., and $(\tilde{X}_t)_{t\geq 0}$ is a Markov process (relative to its own natural filtration) with transition function $(P_t)_{t\geq 0}$.

Step 4: Clean-up. To conclude, we come back to the smaller space Ω of càdlàg paths. The map

$$\tilde{X} : \quad \tilde{\Omega} \to \Omega, \quad \tilde{\omega} \mapsto (\tilde{X}_t(\tilde{\omega}))_{t\geq 0}$$

is measurable, we define $\mathbb{P}_x$ as the image of $\tilde{\mathbb{P}}_x$ under this map. This definition is used to prove existence of $(\mathbb{P}_x)_{x\in E}$ as claimed in the theorem.

The uniqueness is a consequence of the definition of the σ-algebra: let $(\mathbb{P}_x)_{x\in E}$, $(\mathbb{Q}_x)_{x\in E}$ be two families of probability measures such that $(\Omega, \mathcal{F}, (\mathcal{F}_t^0)_{t\geq 0}, (\mathbb{P}_x)_{x\in E}, (X_t)_{t\geq 0}, (\theta_s)_{s\geq 0})$ and the same tuple with $(\mathbb{Q}_x)_{x\in E}$ instead of $(\mathbb{P}_x)_{x\in E}$ become Markov families with transition function $(P_t)_{t\geq 0}$. Then for each $x \in E$, the measures $\mathbb{P}_x$ and $\mathbb{Q}_x$ have the same fidis, i.e.,

$$\mathbb{P}_x(X_{t_1} \in A_1, \ldots, X_{t_n} \in A_n) = \mathbb{Q}_x(X_{t_1} \in A_1, \ldots, X_{t_n} \in A_n)$$

for all $x \in E$, $n \in \mathbb{N}$, $0 \leq t_1 < \cdots < t_n$, and $A_1, \ldots, A_n \in \mathcal{E}$. The family $\mathcal{C}$ of sets of the form $\{X_{t_1} \in A_1, \ldots, X_{t_n} \in A_n\}$ is a π-system that generates $\mathcal{F} = \sigma(X_t, t \geq 0)$, it follows that $\mathbb{P}_x(B) = \mathbb{Q}_x(B)$ for all $B \in \mathcal{F}$ hence $\mathbb{P}_x = \mathbb{Q}_x$. This concludes the proof of Theorem 7.

Appendix A. Some properties of lccb spaces

Every lccb space E satisfies a number of useful properties, some of them not easy to prove (some proofs build on theorems from set theoretic topology such as metrization lemmas etc.):

- E is Polish, i.e., separable and metrizable with complete metric.
- E is *σ-compact*: there is a sequence of compact sets $(K_n)_{n\in\mathbb{N}}$ such that $E = \cup_{n\in\mathbb{N}} K_n$.
- If E is compact, then the constant function $\mathbf{1}$ is in $C_0(E)$. If E is not compact, then $\mathbf{1} \notin C_0(E)$ but there exists a sequence of functions $(f_n)_{n\in\mathbb{N}}$ in $C_0(E)$ such that $f_n \le f_{n+1}$ and $f_n \nearrow \mathbf{1}$ pointwise as $n \to \infty$. Such sequences are sometimes called *approximate units* or *approximate identities*.
- The space $C_0(E)$ is separable.
- Every function $f \in C_0(E)$ is uniformly continuous.
- The functions in $C_0(E)$ generate the Borel-σ-algebra: $\sigma(f,\ f \in C_0(E)) = \mathcal{B}(E)$. This is useful in applications of the functional monotone class theorem.
- The *Riesz-Markov theorem* holds true: Let $\varphi : C_0(E) \to \mathbb{R}$ be a bounded linear map[8] in the Banach space $(C_0(E), ||\cdot||)$ (supremum norm). Then there exists a uniquely defined finite signed measure[9] μ on E (equipped with the Borel σ-algebra) such that $\varphi(f) = \int_E f \mathrm{d}\mu$.

Alexandroff one-point compactification. When E is not compact, it can be useful to compactify it by adding one point; think $\mathbb{R} \to \mathbb{R} \cup \{\infty\}$ with $+\infty$ and $-\infty$ identified. This is done as follows: Let $\mathcal{T}$ be the collection of open sets of E. Let ∞ be some point that is not in E. Define $E^* := E \cup \{\infty\}$. On E^* we define a topology $\mathcal{T}^*$ as the coarsest topology for which: (i) $\mathcal{T} \subset \mathcal{T}^*$, and (ii) $\mathcal{T}^*$ contains all sets of the form $E^* \setminus K$ with $K \subset E$ compact in $(E, \mathcal{T})$ (every such set is considered an open neighborhood of ∞). Then $(E^*, \mathcal{T}^*)$ is a compact lccb space, moreover we can identify $C_0(E)$ with the set of functions $f : E^* \to \mathbb{R}$ that are continuous and satisfy $f(\infty) = 0$.

References

[Chu67] K. L. Chung, *Markov chains with stationary transition probabilities*, Second edition. Die Grundlehren der mathematischen Wissenschaften, Band 104, Springer-Verlag New York, Inc., New York, 1967. MR 0217872

[EK86] S. N. Ethier and T. G. Kurtz, *Markov processes. characterization and convergence*, Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, John Wiley & Sons, Inc., New York, 1986.

[Fel50] W. Feller, *An introduction to probability theory and ots applications. Vol. I*, John Wiley & Sons, Inc., New York, N.Y., 1950. MR 0038583

[Fel71] ———, *An introduction to probability theory and its applications. Vol. II*, Second edition, John Wiley & Sons, Inc., New York-London-Sydney, 1971.

[KS91] I. Karatzas and S. E. Shreve, *Brownian motion and stochastic calculus*, second ed., Graduate Texts in Mathematics, vol. 113, Springer-Verlag, New York, 1991. MR 1121940

[Lig10] Thomas M. Liggett, *Continuous time Markov processes*, Graduate Studies in Mathematics, vol. 113, American Mathematical Society, Providence, RI, 2010, An introduction.

[MT09] S. Meyn and R. L. Tweedie, *Markov chains and stochastic stability*, second ed., Cambridge University Press, Cambridge, 2009, With a prologue by Peter W. Glynn. MR 2509253

[RW94] L. C. G. Rogers and D. Williams, *Diffusions, Markov processes, and martingales. Vol. 1*, second ed., Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, John Wiley & Sons, Ltd., Chichester, 1994, Foundations.

[Sch11] M. Scheutzow, *Stochastische Modelle*, Online lecture notes, TU Berlin, 2011.

[Sch18] ———, *Stochastic Processes II / Wahrscheinlichkeitstheorie III*, Online lecture notes, TU Berlin, 2018.

[8]Meaning: $|\varphi(f)| \le C||f||$ for some $C \in \mathbb{R}_+$ and all $f \in C_0(E)$.

[9]Every such measure can be written as $\mu = \alpha\mu_1 - \beta\mu_2$ with $\alpha, \beta \ge 0$ and μ_1, μ_2 probability measures on E.