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

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

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: January 22, 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 $f : E \to \mathbb{R}$,

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

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 $f \in \mathrm{b}\mathcal{E}$ (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 $\sup_{E \setminus K} |f| \leq \varepsilon$, 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 $(P_t)_{t \geq 0}$ 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* $s, t \geq 0$.
(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*

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

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 similar still works 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, 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*,[2] i.e.,

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

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] 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 $\mathbb{E}_x[f(X_t)] = (P_t f)(x)$ 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 $(E_\partial, \mathcal{E}_\partial)$. 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 $X_t(\omega) := \omega(t)$. 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

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

ζ is the *life-time* of the process.

Appendix A. Some properites 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 $f_n \leq f_{n+1}$ and $f_n \nearrow \mathbf{1}$ pointwise as $n \to \infty$. Such sequences are sometimes called *approximate units* or *approximate identities*.
- The 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[3] in the Banach space $(C_0(E), ||\cdot||)$ (supremum norm). Then there exists a uniquely defined finite signed measure[4] 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 $\mathbb{R} \to \mathbb{R} \cup \{\infty\}$—for $E = \mathbb{R}$, the one-point compactification does not distinguish between $+\infty$ and $-\infty$. 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 $f(\infty) = 0$.

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

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