MARKOV JUMP PROCESSES

SABINE JANSEN

Contents

The present notes collect some background, mostly without proofs, on Markov jump processes in continuous time and with general state spaces. They serve as a preparation for the study of spatial birth and death processes where the state space consists of finite point configurations and the "jumps" correspond to addition or removal of a point.

Jump processes with discrete, countable state spaces, often called *Markov chains*, are treated in [Lig10, Nor98, Sch11]. An in-depth treatment is given in [Chu67]. For jump processes with possibly uncountable state spaces, see [Fel71, Chapter X.3], [Kal97, Chapter 10], [EK86, Chapters 4.2 and 8.3] and [BG68, Chapter I.12].

1. Markov processes

Let $(E, \mathcal{E})$ be some measurable space and $(X_t)_{t\geq 0}$ a stochastic process with state space E, defined on some underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$. To avoid pathologies, we assume that all singletons $\{x\} \subset E$ are measurable, i.e., $\{x\} \in \mathcal{E}$. This is always true in metric spaces equipped with the Borel σ-algebra.

Date: January 26, 2020.

Definition 1. *A* Markov kernel *(also called* probability kernel *or* stochastic kernel*) on E is a map $P : E \times \mathcal{E} \to [0,1]$ such that:*

(i) *The map $E \to [0,1]$, $x \mapsto P(x,A)$ is measurable for all $A \in \mathcal{E}$.*
(ii) *For every $x \in E$, the map $A \mapsto P(x,A)$ is a probability measure on $(E,\mathcal{E})$.*

A sub-Markov kernel *satisfies instead (i) and*

(ii') *For every $x \in E$, the map $A \mapsto P(x,A)$ is a measure on $(E,\mathcal{E})$ with total mass $P(x,E) \leq 1$.*

A kernel *satisfies (i) and*

(ii") *For every $x \in E$, the map $A \mapsto P(x,A)$ is a measure on $(E,\mathcal{E})$.*

Definition 2. *A* normal transition function *on E is a family $(P_t)_{t\geq 0}$ such that*

(i) *Each P_t, $t \geq 0$, is a Markov kernel on E.*
(ii) *The family satisfies the* Chapman-Kolmogorov equations*, i.e., for all $x \in E$, $A \in \mathcal{E}$, and $s,t \geq 0$,*

$$\boxed{\int_E P_t(x,\mathrm{d}y) P_s(y,A) = P_{s+t}(x,A).} \tag{CK}$$

(iii) *$P_0(x,A) = \delta_x(A)$ for all $x \in E$ and $A \in \mathcal{E}$.*

The word "normal" refers to condition (iii). We will always assume that condition (iii) is satisfied and often drop the word "normal" in the sequel.

Definition 3. *The process $(X_t)_{t\geq 0}$ satisfies the* simple Markov property *with respect to the canonical filtration $\mathcal{F}_t^0 = \sigma(X_s, s \leq t)$ if for all $t,h \geq 0$ and $A \in \mathcal{E}$, any version of the conditional probability $\mathbb{P}[X_{t+h} \in A \mid X_t]$ is also a version of the conditional probability $\mathbb{P}(X_{t+h} \in A \mid X_s,\, s \leq t)$, which we write as*

$$\mathbb{P}(X_{t+h} \in A \mid X_s,\, s \leq t) = \mathbb{P}(X_{t+h} \in A \mid X_t) \quad \mathbb{P}\text{-}a.s. \tag{1}$$

The process is a (simple) Markov process with transition function $(P_t)_{t\geq 0}$ *if in addition $P_h(X_t,A)$ is a version of both these conditional expectations,*

$$\mathbb{P}(X_{t+h} \in A \mid X_s,\, s \leq t) = \mathbb{P}[X_{t+h} \in A \mid X_t] = P_h(X_t,A) \quad \mathbb{P}\text{-}a.s. \tag{2}$$

A useful mnemotechnic notation is

$$\boxed{P_h(x,A) = \mathbb{P}(X_{t+h} \in A \mid X_t = x).} \tag{3}$$

Theorem 4. *Let $(P_t)_{t\geq 0}$ be a normal transition function on E and μ a probability measure on E. Assume that E is Polish. Then there exists a simple Markov process $(X_t)_{t\geq 0}$, defined on some probability space $(\Omega,\mathcal{F},\mathbb{P}^\mu)$, with transition function $(P_t)_{t\geq 0}$ and initial law $\mathbb{P}^\mu(X_0 \in \cdot) = \mu(\cdot)$.*

For $\mu = \delta_x$ we use the notation $\mathbb{P}_x := \mathbb{P}^{\delta_x}$. Thus $X_0 = x$, $\mathbb{P}_x$-a.s.

The stochastic process in the theorem (more precisely, the tuple $(\Omega,\mathcal{F},\mathbb{P}^\mu,(X_t)_{t\geq 0})$) is not uniquely determined by the transition function but the finite-dimensional distributions are. If $(X_t)_{t\geq 0}$ is a simple Markov process with transition function $(P_t)_{t\geq 0}$ and initial law μ, then

$$\boxed{\begin{aligned}&\mathbb{P}^\mu\big(X_{t_1}\in A_1,\ldots,X_{t_n}\in A_n\big)\\ &\qquad=\int_{E^{n+1}}\mu(\mathrm{d}x_0)P_{t_1}(x_0,\mathrm{d}x_1)P_{t_2-t_1}(x_1,\mathrm{d}x_2)\cdots P_{t_n-t_{n-1}}(x_{n-1},\mathrm{d}x_n)\mathbb{1}_{A_1}(x_1)\cdots\mathbb{1}_{A_n}(x_n)\end{aligned}}$$

for all $n\in\mathbb{N}$, $0\leq t_1<\cdots<t_n$, $A_1,\ldots,A_n\in\mathcal{E}$. As a consequence, any two Markov processes with same initial law and same transition function are equal in distribution (meaning that they have the same finite-dimensional distributions). The proof of the theorem consists in applying Kolmogorov's extension theorem to the product space $\Omega=E^{[0,\infty)}$ equipped with the product σ-algebra $\mathcal{F}=\mathcal{E}^{[0,\infty)}$, the maps $X_t\big((\omega_s)_{s\geq 0}\big):=\omega_t$, and a well-chosen family of finite-dimensional distributions. In particular, the underlying measure space $(\Omega,\mathcal{F})$ can be chosen independent of the initial law μ.

Remark (Markov family). A slightly stronger version of the theorem holds true: there exists a tuple $(\Omega,\mathcal{F},(\mathbb{P}_x)_{x\in E},(X_t)_{t\geq 0})$ such that:

- $(\Omega,\mathcal{F})$ is a measurable space.
- Each $\mathbb{P}_x$, $x\in E$, is a probability measure on $(E,\mathcal{E})$.
- Each X_t, $t\geq 0$, is a measurable map from Ω to E (measurable with respect to $\mathcal{F}$ and $\mathcal{E}$).
- $X_0=x$, $\mathbb{P}_x$-almost surely.
- For every $A\in\mathcal{E}$ and $t\geq 0$, the map $E\to\mathbb{R}$, $x\mapsto\mathbb{P}_x(X_t\in A)$ is measurable.
- For all $x\in E$, $t,h\geq 0$, and $A\in\mathcal{E}$, we have
$$\mathbb{P}_x\big(X_{t+h}\in A\mid X_s,0\leq s\leq t\big)=P_h(X_t,A)\quad\mathbb{P}_x\text{-a.s.}$$

The extended version of the theorem emphasizes that there is actually a *family* of probability measures $(\mathbb{P}_x)_{x\in E}$ and paves the way for the abstract notion of *Markov family*, see [KS91, Definition 5.11] or [Sch18, Definition 4.10].

Invariant and reversible measure

Definition 5. *Let $(P_t)_{t\geq 0}$ be a Markov transition function on $(E,\mathcal{E})$. A probability measure μ on $(E,\mathcal{E})$ is* invariant *with respect to $(P_t)_{t\geq 0}$ if for all $t\geq 0$*

$$\int_E\mu(\mathrm{d}x)P_t(x,A)=\mu(A)\quad(A\in\mathcal{E}).$$

A convenient sufficient criterion for invariance is *reversibility.* For $t\geq 0$, consider the t-dependent measure $\mu\otimes P_t$ on $E\times E$ given by

$$\big(\mu\otimes P_t\big)(A\times B)=\int_A\mu(\mathrm{d}x)P_t(x,B).$$

A good way to remember this definition is to write

$$\big(\mu\otimes P_t\big)\big(\mathrm{d}(x,y)\big)=\mu(\mathrm{d}x)P_t(x,\mathrm{d}y).$$

A measure ν on $E\times E$ is *symmetric* if $\nu(A\times B)=\nu(B\times A)$ for all $A,B\in\mathcal{E}$.

Definition 6. *Let $(P_t)_{t\geq 0}$ be a transition function on $(E,\mathcal{E})$. A probability measure μ on $(E,\mathcal{E})$ is* reversible *with respect to $(P_t)_{t\geq 0}$ if for all $t\geq 0$, the measure $\mu\otimes P_t$ is symmetric, i.e.,*

$$\int_A \mu(\mathrm{d}x)P_t(x,B) = \int_B \mu(\mathrm{d}y)P_t(y,A)$$

for all $A,B\in\mathcal{E}$ and $t\geq 0$.

Remark. More generally, if $K(x,\mathrm{d}y)$ is a kernel and μ a measure and

$$\int_A \mu(\mathrm{d}x)K(x,B) = \int_B \mu(\mathrm{d}y)K(y,A)$$

for all $A,B\in\mathcal{E}$, then we say that μ *symmetrizes* K or it is a *symmetrizing measure.* We reserve the word "reversible" for probability measures and Markov transition functions.

A convenient mnemotechnic notation for the symmetry of $\mu\otimes P_t$ is

$$\boxed{\mu(\mathrm{d}x)P_t(x,\mathrm{d}y) = \mu(\mathrm{d}y)P_t(y,\mathrm{d}x).} \tag{4}$$

This relation is often referred to as *detailed balance equation.* If E is discrete, then physicists think of $\mu(\{x\})P_t(x,\{y\})$ as a current from x to y. Detailed balance then says that the current from x to y is the same as the current from y to x.

Lemma 7. *If a probability measure μ is reversible for the (Markov) transition function $(P_t)_{t\geq 0}$, then it is invariant with respect to $(P_t)_{t\geq 0}$.*

For the lemma it is important that $P_t(x,E)=1$—symmetrizing measures for sub-Markov transition functions are not necessarily invariant!

2. Jump processes

2.1. **Setting.** Let $\Pi(x,\mathrm{d}y)$ be a Markov kernel on $(E,\mathcal{E})$ and $\lambda: E\to[0,\infty)$ a non-negative measurable function. Define the *jump kernel*

$$K(x,\mathrm{d}y) = \lambda(x)\Pi(x,\mathrm{d}y)$$

(more precisely, K is the kernel given by $K(x,A)=\lambda(x)P(x,A)$). We assume that each singleton is measurable and

$$\Pi(x,E\setminus\{x\}) = 1 \qquad (x\in E).$$

We would like to define a Markov process $(X_t)_{t\geq 0}$ with state space E and piecewise constant sample paths. The behavior, loosely speaking, should be as follows.

Suppose the process starts in $X_0=x$. Then it waits for random time with exponential distribution $\tau_1\sim \mathrm{Exp}(\lambda(x))$—this is often referred to as an exponential clock. At $t=\tau_1$—i.e. when the exponential clock rings—the process jumps to a new location $X_{\tau_1}=y$, chosen according to the Markov kernel $\Pi(x,\mathrm{d}y)$. Then it waits again until an exponential clock with parameter $\lambda(y)$ rings and jumps to a new location z chosen according to the kernel $\Pi(y,\mathrm{d}z)$, etc.

In the context of spatial birth and death processes, the state space E consists of configurations of points e.g. in some bounded domain in $\mathbb{R}^d$ and the "jumps" correspond to addition or removal of some points of the configuration.

Jump processes can be approached in two complementary ways. They can be constructed directly with the jump-hold construction, using an auxiliary embedded discrete-time Markov chain and exponential random variables. A more analytic angle of attack is to define first the associated transition function.

2.2. **Jump-hold construction.** Let us fix some initial law μ. Suppose we are given a probability space $(\Omega, \mathcal{F}, \mathbb{P}^\mu)$ and random variables $Z_n : \Omega \to E$, $n \in \mathbb{N}_0$, and $\tau_k : \Omega \to \mathbb{R}_+ \cup \{\infty\}$, $k \in \mathbb{N}$, such that:

(i) $(Z_n)_{n\in\mathbb{N}_0}$ is a discrete-time Markov process with transition kernel Π and initial law $Z_0 \sim \mu$. Thus

$$\mathbb{P}^\mu\bigl(Z_0 \in A_0, \dots, Z_n \in A_n\bigr) = \int_{E^{n+1}} \mu(\mathrm{d}z_0)\Pi(z_0, \mathrm{d}z_1)\cdots\Pi(z_{n-1}, \mathrm{d}z_n)\mathbb{1}_{A_0}(z_0)\cdots\mathbb{1}_{A_n}(z_n),$$

for all $n \in \mathbb{N}_0$ and $A_0, \dots, A_n \in \mathcal{E}$. The chain $(Z_n)_{n\in\mathbb{N}_0}$ is called *embedded jump chain.*

(ii) Conditional on $(Z_k)_{k\in\mathbb{N}_0}$, the variables τ_k, $k \in \mathbb{N}$ are independent exponential random variables[1] with parameters $\lambda(Z_{k-1})$,, i.e.,

$$\mathbb{P}^\mu(\tau_1 \geq s_1, \dots, \tau_n \geq s_n \mid Z_k,\, k \in \mathbb{N}_0) = \prod_{i=1}^n \mathrm{e}^{-\lambda(Z_{i-1})s_i}.$$

for all $n \in \mathbb{N}$ and all $s_1, \dots, s_n \geq 0$. The τ_i's are called *holding times.*

Remark. If E is Polish, the existence of a tuple $(\Omega, \mathcal{F}, \mathbb{P}^\mu, (Z_n)_{n\in\mathbb{N}_0}, (\tau_n)_{n\in\mathbb{N}})$ as above can be checked with Kolmogorov's extension theorem. The existence can also be checked with the *Ionescu-Tulcea theorem* [Kle08, Theorem 14.32], which is applicable for every measurable space $(E, \mathcal{E})$—the space E need not be Polish.

Define

$$T_0 := 0, \quad T_k := \tau_1 + \cdots + \tau_k \quad (k \in \mathbb{N}), \quad \zeta := \sum_{i=1}^\infty \tau_i.$$

The sum ζ is always well-defined with values in $\mathbb{R}_+ \cup \{\infty\}$. It will play the role of the *life-time* of our jump process. The random time interval $[0, \zeta)$ can be partitioned as

$$[0, \zeta) = \bigcup_{k=0}^\infty [T_k, T_{k+1}).$$

To help us deal with the case $\zeta < \infty$, we enlarge the state space E by adding a state $\partial \notin E$, called *coffin* or *cemetery.* Then we set

$$E_\partial := E \cup \{\partial\}$$

[1] If $\lambda = 0$, we define $\mathrm{Exp}(\lambda)$ as the measure on $\mathbb{R}_+ \cup \{\infty\}$ that puts full mass on $\{\infty\}$. Thus $T \sim \mathrm{Exp}(\lambda)$ if and only if $T = \infty$ almost surely. This is consistent with the equality $\mathbb{P}(T \geq t) = \exp(-\lambda t)$ for all $t \geq 0$.

and let $\mathcal{E}_\partial$ be the smallest σ-algebra containing all sets $E \in \mathcal{E}$ and the singleton $\{\partial\}$. We define a process $(X_t^\partial)_{t\geq 0}$ with state space E_∂ by

$$X_t^\partial := \begin{cases} Z_k, & t \in [T_k, T_{k+1}), \\ \partial, & t \geq \zeta. \end{cases}$$

Thus $X_t^\partial = Z_0 = X_0^\partial$ on $[0, \tau_1)$, $X_t^\partial = Z_1$ on $[\tau_1, \tau_1 + \tau_2)$, etc. Then

$$\zeta = \min\{t \geq 0 \mid X_t^\partial = \partial\}.$$

The random time is called the *life-time* of the process. Finally we define the number of jumps up to time t as

$$N_t := \max\{k \in \mathbb{N}_0 \mid \tau_1 + \cdots + \tau_k \leq t\} = \#\{k \in \mathbb{N}_0 \mid T_k \leq t\}$$

and note

$$\zeta > t \ \Leftrightarrow \ N_t < \infty.$$

Definition 8. *The process $(X_t^\partial)_{t\geq 0}$ defined on $(\Omega, \mathcal{F}, \mathbb{P}^\mu)$ is* non-explosive *if its life-time is infinite, $\mathbb{P}^\mu$-almost surely, and* explosive *otherwise.*

Thus $(X_t^\partial)_{t\geq 0}$ is non-explosive if and only if

$$\sum_{k=1}^\infty \tau_k = \infty, \quad \mathbb{P}^\mu\text{-a.s.}$$

If the process is non-explosive, there is no need to extend the state space from E to E^∂ and we can define instead a process with state space E as follows. Let $N \subset \Omega$ be a measurable set with $\mathbb{P}^\mu(N) = 0$ such that $\sum_{k=1}^\infty \tau_k = \infty$ (hence $\zeta = \infty$) on $\Omega \setminus N$. Let $x_0 \in E$ be an arbitrary element. We define

$$X_t(\omega) := \begin{cases} Z_k(\omega), & \omega \in \Omega \setminus N \text{ and } t \in [T_k(\omega), T_{k+1}(\omega)), \\ x_0, & \omega \in N. \end{cases}$$

Then $(X_t)_{t\geq 0}$ is a stochastic process with state space E, moreover $X_t(\omega) = X_t^\partial(\omega)$ for all $t \geq 0$ and all $\omega \in \Omega \setminus N$ (the processes are *indistinguishable*).

2.3. **Transition function. Kolmogorov backward equation.**

2.3.1. *Analytic construction.* Intuitively, a path from x_0 to $x_k \in A$ in time t that has exactly k jumps at the moments $0 < t_1 < \cdots < t_k < t$ and locations $x_1, \ldots, x_k$ right after the jumps should be associated with the weight

$$\lambda(x_0)\mathrm{e}^{-\lambda(x_0)t_1}\Pi(x_0, \mathrm{d}x_1)\mathrm{e}^{-\lambda(x_1)(t_2-t_1)}\Pi(x_1, \mathrm{d}x_2)\cdots\Pi(x_{k-1}, \mathrm{d}x_k)\lambda(x_{k-1})\mathrm{e}^{-\lambda(x_{k-1})(t-t_k)}.$$

This motivates the following definition. For $t > 0$, $n \in \mathbb{N}$, $x_0 \in E$, and $A \in \mathcal{E}$, set

$$\begin{aligned} P_t^{(n)}(x_0, A) &= \mathrm{e}^{-\lambda(x_0)t}\delta_{x_0}(A) \\ &\quad + \sum_{k=1}^{n} \int_{0<t_1<\cdots<t_k\le t} \mathrm{d}t_1 \cdots \mathrm{d}t_k \int_{E^k} \Pi(x_0, \mathrm{d}x_1)\Pi(x_1, \mathrm{d}x_2)\cdots\Pi(x_{k-1}, \mathrm{d}x_k)\mathbb{1}_A(x_k) \\ &\qquad \times \lambda(x_0)\mathrm{e}^{-\lambda(x_0)t_1}\lambda(x_1)\mathrm{e}^{-\lambda(x_1)(t_2-t_1)} \times\cdots\times \lambda(x_{k-1})\mathrm{e}^{-\lambda(x_{k-1})(t-t_k)}. \end{aligned} \tag{5}$$

For $n = 0$ we set

$$P_t^{(0)}(x_0, A) = \mathrm{e}^{-\lambda(x_0)t}\delta_{x_0}(A).$$

This defines a family of mappings $P_t^{(n)} : E \times \mathcal{E} \to [0, \infty)$, $t \ge 0$, $n \in \mathbb{N}$. We note the joint measurability in t and x: for every $n \in \mathbb{N}_0$ and $A \in \mathcal{E}$, the map

$$\mathbb{R}_+ \times E \to \mathbb{R}_+, \quad (t, x) \mapsto P_t^{(n)}(x, A)$$

is measurable.

Lemma 9. *The family $(P_t^{(n)})_{n\in\mathbb{N}_0, t\ge 0}$ satisfies*

$$P_t^{(n+1)}(x, A) = \mathrm{e}^{-\lambda(x)t}\delta_x(A) + \int_0^t \mathrm{d}s \int_E \lambda(x)\mathrm{e}^{-\lambda(x)s}\Pi(x, \mathrm{d}y)P_{t-s}^{(n)}(y, A), \tag{6}$$

for all $n \in \mathbb{N}_0$, $t \ge 0$, $x \in E$, and $A \in \mathcal{E}$.

The recurrence relation together with the explicit value of $P_t^{(0)}$ is often used as a definition of $P_t^{(n)}$, instead of the sum (5), see e.g. [Lig10, Chapter 2.5] for discrete countable space E, and [Fel71, Chapter X.3] for the general case.

Lemma 10. *We have $P_t^{(n)}(x, E) \le P_t^{(n+1)}(x, E) \le 1$, for all $n \in \mathbb{N}_0$ and $x \in E$.*

Let $P_t^{\langle k\rangle}(x_0, A)$ be the k-th summand in (5) so that

$$P_{s+t}^{(n)}(x, A) = \sum_{k=0}^{n} P_t^{\langle k\rangle}(x, A).$$

Lemma 11. *For all $s, t \ge 0$, $n \in \mathbb{N}_0$, $x \in E$, and $A \in \mathcal{E}$, we have*

$$P_{s+t}^{\langle n\rangle}(x, A) = \sum_{k=0}^{n} \int_E P_s^{\langle k\rangle}(x, \mathrm{d}y)P_t^{\langle n-k\rangle}(y, A).$$

Compare [Chu67, Chapter II.18, Eq. (5].

Clearly the pointwise limit

$$P_t^*(x, A) := \lim_{n\to\infty} P_t^{(n)}(x, A)$$

exists for all $t \ge 0$, $x \in E$, and $A \in \mathcal{E}$ and satisfies $P_t^*(x, A) \le P_t^*(x, E) \le 1$. The convergence is monotone, $P_t^{(n)}(x, A) \nearrow P_t^*(x, A)$.

Proposition 12. *The family $(P_t^*)_{t\ge 0}$ is a normal sub-Markov transition function, i.e.,*

(a) *Each* P_t^* *is a sub-Markov kernel.*
(b) $P_0^*(x,\cdot) = \delta_x(\cdot)$, *for all* $x \in E$.
(c) $(P_t^*)_{t\geq 0}$ *satisfies the Chapman-Kolmogorov equations.*

Proposition 13. *The sub-Markov transition function* $(P_t^*)_{t\geq 0}$ *solves the* integrated backward Kolmogorov equation, *i.e.,*

$$P_t^*(x,A) = \mathrm{e}^{-\lambda(x)t}\delta_x(A) + \int_0^t \mathrm{d}s \int_E \lambda(x)\mathrm{e}^{-\lambda(x)(t-s)}\Pi(x,\mathrm{d}y)P_s^*(y,A). \tag{7}$$

for all $t \geq 0$, $x \in E$, *and* $A \in \mathcal{E}$.

Proof. We pass to the limit $n \to \infty$ in the recurrence relation (6). On the right-hand side the exchange of limits and integration is justified by monotone convergence. We obtain the integral equation

$$P_t^*(x,A) = \mathrm{e}^{-\lambda(x)t}\delta_x(A) + \int_0^t \mathrm{d}s \int_E \lambda(x)\mathrm{e}^{-\lambda(x)s}\Pi(x,\mathrm{d}y)P_{t-s}^*(y,A). \tag{8}$$

A change of variables from s to $t-s$ in (8) yields Eq. (7). □

Differentiating on both sides of (7) yields an integro-differential equation.

Definition 14. *We say that* $(P_t)_{t\geq 0}$ *solves the* backward Kolmogorov equation *if* $(P_t)_{t\geq 0}$ *is a sub-Markov normal transition function and for all* $x \in E$ *and* $A \in \mathcal{E}$, *the map* $t \mapsto P_t(x,A)$ *is differentiable on* $\mathbb{R}_+$ *and*

$$\frac{\partial}{\partial t}P_t(x,A) = -\lambda(x)P_t(x,A) + \lambda(x)\int_E \Pi(x,\mathrm{d}y)P_t(y,A). \tag{9}$$

The backward Kolmogorov equation can be rewritten with the jump kernel $K(x,\mathrm{d}y) = \lambda(x)\Pi(x,\mathrm{d}y)$ as

$$\boxed{\frac{\partial}{\partial t}P_t(x,A) = \int_E K(x,\mathrm{d}y)\big(P_t(y,A) - P_t(x,A)\big).} \tag{10}$$

Theorem 15.

(a) *The family* $(P_t^*)_{t\geq 0}$ *solves the backward Kolmogorov equation.*
(b) *Every sub-Markov transition function* $(P_t)_{t\geq 0}$ *that solves the backward Kolmogorov equation satisfies* $P_t(x,A) \geq P_t^*(x,A)$ *for all* $t \geq 0$, $x \in E$, *and* $A \in \mathcal{E}$.
(c) *If* $P_t^*(x,E) = 1$ *for all* $x \in E$ *and* $t \geq 0$, *then* $(P_t^*)_{t\geq 0}$ *is the unique solution of the backward Kolmogorov equation.*

Parts (a) and (b) say that $(P_t^*)_{t\geq 0}$ is the *minimal solution* of the backward Kolmogorov equation, part (c) says that if the minimal solution is *stochastic* (another word for Markov, as opposed to sub-Markov), then it is the unique solution.

Before we prove Theorem 15, we check that $t \mapsto P_t^*(x,A)$ is continuous.

Lemma 16. *For every* $x \in E$ *and* $A \in \mathcal{E}$, *the map* $t \mapsto P_t^*(x,A)$ *is continuous on* $\mathbb{R}_+$.

Proof. We show first continuity in 0. Let $f : E \to \mathbb{R}_+$ be a measurable, bounded function and $(P_t^* f)(x) := \int_E P_t^*(x, \mathrm{d}y) f(y)$. Notice $P_t^* f \geq 0$ and

$$\sup_{x \in E} (P_t^* f)(x) \leq \sup_{x \in E} f(x).$$

By the integrated backward Kolmogorov equation in the form (8), we have

$$(P_t^* f)(x) = \mathrm{e}^{-\lambda(x)t} f(x) + \int_0^t \mathrm{d}s \int_E \lambda(x) \mathrm{e}^{-\lambda(x)s} \Pi(x, \mathrm{d}y) \big(P_{t-s}^* f\big)(y).$$

It follows that

$$\begin{aligned} \big|(P_t^* f)(x) - \mathrm{e}^{-\lambda(x)t} f(x)\big| &\leq \Big(\int_0^t \lambda(x) \mathrm{e}^{-\lambda(x)s} \mathrm{d}s\Big) \sup_{x \in E} f(x) \\ &= \big(1 - \mathrm{e}^{-\lambda(x)t}\big) \sup_{x \in E} f(x) \\ &\leq \lambda(x) t \sup_{x \in E} f(x). \end{aligned}$$

For the last bound we have applied the inequality $\exp(u) \geq 1 + u$ to $u = -\lambda(x)t$. It follows in particular that $(P_t^* f)(x) \to f(x)$ as $t \searrow 0$, for all $x \in E$.

Now let $t, h \geq 0$ and $A \in \mathcal{E}$. Set $f(x) := P_t^*(x, A)$. By the Chapman-Kolmogorov equation,

$$P_{t+h}^*(x, A) = \int_E P_h^*(x, \mathrm{d}y) P_t^*(y, A) = (P_h^* f)(x).$$

As a consequence,

$$\big|P_{t+h}^*(x, A) - P_t^*(x, A)\big| \leq h\lambda(x) + \big|\mathrm{e}^{-\lambda(x)h} - 1\big| \leq 2h\lambda(x)$$

and the map $t \mapsto P_t^*(x, A)$ is continuous (in fact, Lipschitz-continuous) on $\mathbb{R}_+$. □

Proof of Theorem 15. (a) Below we check that the right-hand side of (7) is differentiable in t, with derivative

$$-\lambda(x) \mathrm{e}^{-\lambda(x)t} \delta_x(A) - \int_0^t \mathrm{d}s \int_E \lambda(x)^2 \mathrm{e}^{-\lambda(x)(t-s)} \Pi(x, \mathrm{d}y) P_s^*(y, A) + \int_E \lambda(x) \Pi(x, \mathrm{d}y) P_t^*(y, A). \tag{11}$$

The first two terms can be regrouped by using (7), which gives $-\lambda(x) P_t^*(x, A)$. It follows that (11) is equal to the right-hand side of the backward Kolmogorov equation for P_t^*.

It remains to prove that the right-hand side of (7) is differentiable in t with derivative given by (11). The integral equation (7) yields, for $h > 0$,

$$\begin{aligned} \frac{1}{h}\big(P_{t+h}^*(x, A) - P_t(x, A)\big) = \frac{1}{h}\big(\mathrm{e}^{-\lambda(x)(t+h)} - \mathrm{e}^{-\lambda(x)t}\big) \delta_x(A) \\ + \int_0^t \mathrm{d}s \int_E \lambda(x) \frac{1}{h}\big(\mathrm{e}^{-\lambda(x)h} - 1\big) \mathrm{e}^{-\lambda(x)(t-s)} \Pi(x, \mathrm{d}y) P_s^*(y, A) \\ + \frac{1}{h} \int_t^{t+h} \mathrm{d}s \int_E \lambda(x) \mathrm{e}^{-\lambda(x)(t+h-s)} \Pi(x, \mathrm{d}y) P_s^*(y, A) \end{aligned} \tag{12}$$

The first term on the right-hand side clearly converges to $-\lambda(x)\mathrm{e}^{-\lambda(x)t}\delta_x(A)$ as $h \searrow 0$, which is the first term in (11). For the second term on the right-hand side of (12) we bound

$$\frac{1}{h}\big|\mathrm{e}^{-\lambda(x)h} - 1\big| = \frac{1}{h}\int_0^h \lambda(x)\mathrm{e}^{-\lambda(x)u}\mathrm{d}u \leq \lambda(x)$$

and

$$\int_0^t \mathrm{d}s \int_E \lambda(x)^2 \mathrm{e}^{-\lambda(x)(t-s)}\Pi(x,\mathrm{d}y)P_s^*(y,A) \leq t\lambda(x)^2 < \infty.$$

Dominated convergence therefore allows us to pass to the limit $h \to 0$ for the middle term on the right-hand side of (12), which gives the middle term in (11). The last integral in (12) is equal to

$$\lambda(x)\mathrm{e}^{-\lambda(x)(t+h)}\int_E \Big(\frac{1}{h}\int_t^{t+h} P_s^*(y,A)\mathrm{d}s\Big)\Pi(x,\mathrm{d}y).$$

The inner integral is bounded by 1 and converges to $P_t^*(y,A)$, for each y, because of the continuity of $t \mapsto P_t^*(x,A)$. It follows that altogether, the expression converges to last term in (11).

Thus we have proven that the limit of the difference quotient $\frac{1}{h}(P_{t+h}^*(x,A) - P_t^*(x,A))$ as $h \searrow 0$ exists and is given by (11). It is not difficult to check that the right-hand side is a continuous function of t. Thus $t \mapsto P_t^*(x,A)$ is continuous (by Lemma 16) and has right derivatives everywhere, and the right derivative is a continuous function of t. It follows from a general lemma (see [Sch11, Lemma 3.3.2]) that $t \mapsto P_t^*(x,A)$ is in fact differentiable and the derivative is given by (11).

(b) $(P_t^*)_{t\geq 0}$ is the minimal solution. Let $(P_t)_{t\geq 0}$ be a solution of the backward Kolmogorov equation. Then $(P_t)_{t\geq 0}$ also solves the integrated backward Kolmogorov equation,

$$P_t(x,A) = \mathrm{e}^{-\lambda(x)t}\delta_x(A) + \int_0^t \mathrm{d}s \int_E \lambda(x)\mathrm{e}^{-\lambda(x)s}\Pi(x,\mathrm{d}y)P_{t-s}(y,A)$$

because the equation holds true at $t = 0$ and the left and right-hand sides have the same derivatives with respect to t. An induction over n yields $P_t(x,A) \geq P_t^{(n)}(x,A)$ for all $n \in \mathbb{N}_0$, hence $P_t(x,A) \geq P_t^*(x,A)$.

(c) Uniqueness if the minimal solution is stochastic. Suppose $P_t^*(x,E) = 1$ for all $t \geq 0$ and $x \in E$. Let $(P_t)_{t\geq 0}$ be a solution of the backward Kolmogorov equation. Then $P_t(x,A) \geq P_t^*(x,A)$ because $(P_t^*)_{t\geq 0}$ is the minimal solution (by part (b) of the proof). Similarly, $P_t(x,E\setminus A) \geq P_t^*(x,E\setminus A)$. Moreover

$$1 \geq P_t(x,E) = P_t(x,A) + P_t(x,E\setminus A) \geq P_t^*(x,A) + P_t^*(x,E\setminus A) = P_t^*(x,E).$$

Because of $P_t^*(x,E)$, the previous inequalities must be equalities, which is only possible if $P_t(x,A) = P_t^*(x,A)$. It follows that $(P_t)_{t\geq 0}$ is equal to $(P_t^*)_{t\geq 0}$. $\square$

2.3.2. *Probabilistic interpretation.* The relation between the process $(X_t^\partial)_{t\geq 0}$ from Section 2.2 and the sub-Markov kernels $P_t^{(n)}$, P_t^* from Section 2.3.1 is summarized by the following lemma. Remember $\mathbb{P}_x = \mathbb{P}^{\delta_x}$ and $X_0 = x$, $\mathbb{P}_x$-a.s.

Lemma 17. *We have*

$$P_t^{(n)}(x, A) = \mathbb{P}_x\Big(X_t^\partial \in A,\ N_t \leq n\Big)$$

and

$$P_t^*(x, A) = \mathbb{P}_x\Big(X_t^\partial \in A,\ \zeta > t\Big),$$

for all $t \geq 0$, $x \in E$, $A \in \mathcal{E}$, and $n \in \mathbb{N}$.

Proposition 18. *Let $(\Omega, \mathcal{F}, \mathbb{P}^\mu)$ and $(X_t^\partial)_{t\geq 0}$ be as in Section 2.2. The following two conditions are equivalent:*

(i) *The process $(X_t^\partial)_{t\geq 0}$ is non-explosive with respect to $\mathbb{P}_x$, for every $x \in E$.*
(ii) *The minimal solution $(P_t^*)_{t\geq 0}$ of the backward Kolmogorov equation is stochastic.*

Theorem 19. *If the minimal solution $(P_t^*)_{t\geq 0}$ is stochastic, then for every $x \in E$, the process $(X_t)_{t\geq 0}$ on $(\Omega, \mathcal{F}, \mathbb{P}_x)$ defined with the jump-hold construction is a Markov process with state space E and transition function $(P_t^*)_{t\geq 0}$.*

Without the assumption that $(P_t^*)_{t\geq 0}$ is stochastic, we can still formulate a Markov property. First we extend the sub-Markov transition function $(P_t^*)_{t\geq 0}$ on $(E, \mathcal{E})$ to a *Markov* transition function $(P_t^\partial)_{t\geq 0}$ on $(E_\partial, \mathcal{E}_\partial)$ by asking that for all $t \geq 0$

$$\begin{aligned} P_t^\partial(x, A) &= P_t^*(x, A) && (x \in E,\ A \in \mathcal{E}), \\ P_t^\partial(x, \{\partial\}) &= 1 - P_t^*(x, E) && (x \in E), \\ P_t^\partial(\partial, \{\partial\}) &= 1. \end{aligned}$$

It is easily checked that there is a unique Markov transition function on $(E_\partial, \mathcal{E}_\partial)$ that satisfies the previous three equations. Then for the process $(X_t^\partial)_{t\geq 0}$ is a Markov process with state space E_∂ and transition function $(P_t^\partial)_{t\geq 0}$. In particular, for all $x \in E$, $t, h \geq 0$, and $A \in \mathcal{E}$,

$$\mathbb{P}_x(X_{t+h}^\partial \in A \mid X_s^\partial, s \leq t) = \mathbb{1}_{\{\zeta > t\}}\, P_{t+h}^\partial(X_t^\partial, A) = \mathbb{1}_{\{\zeta > t\}}\, P_{t+h}^*(X_t^\partial, A) \quad \mathbb{P}_x\text{-a.s.}$$

2.4. **Reversible measure.**

Proposition 20. *Let μ be a probability measure on $(E, \mathcal{E})$. Suppose that $\mu(\mathrm{d}x)K(x, \mathrm{d}y) = \mu(\mathrm{d}y)K(y, \mathrm{d}x)$. Then μ is a symmetrizing measure for $(P_t^*)_{t\geq 0}$, i.e.,*

$$\int_A \mu(\mathrm{d}x) P_t^*(x, B) = \int_B \mu(\mathrm{d}y) P_t^*(y, A)$$

for all $A, B \in \mathcal{E}$ and $t \geq 0$.

Corollary 21. *If the probability measure μ is a symmetrizing measure for the jump kernel $K(x, \mathrm{d}y)$ and the minimal solution $(P_t^*)_{t\geq 0}$ is stochastic, then μ is a reversible measure, hence also an invariant measure for $(P_t^*)_{t\geq 0}$.*

3. Coupling

Preston's criterion for non-explosion of spatial birth and death processes [Pre75] builds on a *coupling* between two processes, a spatial birth and death process $(\eta_t)_{t\geq 0}$ and a non-spatial birth and death process $(X_t)_{t\geq 0}$. "Coupling" roughly means that the processes $(\eta_t)_{t\geq 0}$, $(X_t)_{t\geq 0}$ are defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$. One way of trying to get such a coupling is to construct a Markov process $(\eta_t, X_t)_{t\geq 0}$ with bigger state space $\mathbf{N}_{<\infty}(\mathbb{X}) \times \mathbb{N}_0$ that is such that the component processes $(\eta_t)_{t\geq 0}$, $(X_t)_{t\geq 0}$ are spatial and non-spatial birth and jump processes, respectively.

Here we collect some relevant background for such couplings. The general setting is as follows [Che86]. Let $(E_1, \mathcal{E}_1)$ and $(E_2, \mathcal{E}_2)$ be two measurable spaces, and $(E, \mathcal{E}) = (E_1 \times E_2, \mathcal{E}_1 \otimes \mathcal{E}_2)$ the product space. Let K_1 and K_2 be jump kernels on $(E_1, \mathcal{E}_1)$ and $(E_2, \mathcal{E}_2)$ respectively. We look for ways of constructing a process $(X_t)_{t\geq 0} = ((X_{1,t}, X_{2,t}))_{t\geq 0}$ with state space E in such a way that the component processes $(X_{1,t})_{t\geq 0}$ and $(X_{2,t})_{t\geq 0}$ are jump processes with prescribed jump kernels K_1 and K_2.

We write L_i and $P^*_{i,t}$ for the formal generator and minimal solution of the backward Kolmogorov equations associated with the jump kernel K_i.

3.1. Independent processes.

The simplest—though often least useful—coupling consists in making the two component processes independent. Let ∂ be an element not contained in E and

$$\Big(\Omega, \mathcal{F}, (\mathbb{P}_x)_{x\in E}, (X^\partial_{1,t})_{t\geq 0}, (X^\partial_{2,t})_{t\geq 0}\Big)$$

a tuple consisting of: a measurable space $(\Omega, \mathcal{F})$, a family of probability measures $\mathbb{P}_x$, $x \in E$, measurable maps $X^\partial_{i,t} : \Omega \to E^\partial_i$, where $E^\partial_i = E_i \cup \{\partial\}$. We assume that for each $x = (x_1, x_2)$, under $\mathbb{P}_x$, the processes $(X^\partial_{1,t})_{t\geq 0}$ and $(X^\partial_{2,t})_{t\geq 0}$ are independent and they are the minimal jump processes with respective kernels K_1 and K_2, initial values x_1 and x_2, and cemetery ∂. Let ζ_1 and ζ_2 be the respective lifetimes of the processes and

$$\zeta := \min(\zeta_1, \zeta_2).$$

Define

$$\tilde{X}_t := \big(X^\partial_{1,t}, X^\partial_{2,t}\big) \quad (t \geq 0)$$

and

$$X^\partial_t := \begin{cases} \tilde{X}_t, & t < \zeta, \\ \partial, & t \geq \zeta. \end{cases}$$

Notice $\tilde{X}_t(\omega) = X^\partial_t(\omega) \in E$ for all $t < \zeta(\omega)$ and $\omega \in \Omega$. Set

$$P_t(x, A) := \mathbb{P}_x\big(X^\partial_t \in A,\, \zeta < t\big) \quad (A \in \mathcal{E}).$$

Then for $A_1 \subset E_1$ and $A_2 \subset E_2$, we have

$$\mathbb{P}_x\big(X^\partial_t \in A,\, \zeta < t\big) = \mathbb{P}_{(x_1,x_2)}\big(X^\partial_{1,t} \in A_1,\, t < \zeta_1\big)\mathbb{P}_{(x_1,x_2)}\big(X^\partial_{2,t} \in A_2,\, t < \zeta_2\big)$$

hence

$$P_t(x_1, x_2; A_1 \times A_2) = P^*_{1,t}(x_1, A_1) P^*_{2,t}(x_2, A_2).$$

Notice that

$$P_t(x_1,x_2;A_1\times E_2)\le P^*_{1,t}(x_1,A_1),\quad P_t(x_1,x_2;E_1\times A_2)\le P^*_{2,t}(x_2,A_2) \tag{13}$$

and the inequality may be strict. One checks that $(P_t)_{t\geq 0}$ is a sub-Markov transition function on $(E,\mathcal{E})$. It solves the backward Kolmogorov equation associated with the jump kernel uniquely defined by

$$K(x_1,x_2;A_1\times A_2):=K_1(x_1,A_1)\delta_{x_2}(A_2)+\delta_{x_1}(A_1)K_2(x_2,A_2).$$

The associated formal generator is

$$(Lf)(x_1,x_2)=\int_{E_1}K_1(x_1,\mathrm{d}y_1)\big(f(y_1,x_2)-f(x_1,x_2)\big)+\int_{E_2}K_2(x_2,\mathrm{d}y_2)\big(f(x_1,y_2)-f(x_1,x_2)\big).$$

Then $(P_t)_{t\geq 0}$ solves the backward Kolmogorov equation for the jump kernel K.

3.2. **Dependent coupling.** More generally, suppose that kernels K, K_1, and K_2 are given such that the associated generators L, L_1, L_2 satisfy the following condition.

Condition 1. *The following holds true for all $f\in\mathscr{L}^\infty(E,\mathcal{E})$:*

(i) *If $f(x_1,x_2)=f_1(x_1)$ for some $f_1\in\mathscr{L}^\infty(E_1,\mathcal{E}_1)$ and all $(x_1,x_2)\in E$, then $(Lf)(x_1,x_2)=(L_1f_1)(x_1)$ on E.*

(ii) *If $f(x_1,x_2)=f_1(x_1)$ for some $f_2\in\mathscr{L}^\infty(E_2,\mathcal{E}_2)$ and all $(x_1,x_2)\in E$, then $(Lf)(x_1,x_2)=(L_2f_2)(x_2)$ on E.*

A compact way of writing the condition uses the notation

$$(f\otimes g)(x_1,x_2):=f(x_1)g(x_2)$$

and $\mathbf{1}$ for the constant function 1. Then the condition becomes

$$\boxed{L(f_1\otimes\mathbf{1})=(L_1f_1)\otimes\mathbf{1},\quad L(\mathbf{1}\otimes f_2)=\mathbf{1}\otimes(L_2f_2).}$$

Write P^*_t, $P^*_{i,t}$ for the minimal solutions associated with the jump kernels K and K_i.

Theorem 22. [Che86, Theorem 13] *Suppose that $(P^*_t)_{t\geq 0}$ is stochastic and that Condition 1 is satisfied. Then $(P^*_{1,t})_{t\geq 0}$ and $(P^*_{2,t})_{t\geq 0}$ are stochastic as well, moreover*

$$P^*_t(x_1,x_2;A_1\times E_2)=P^*_{1,t}(x_1,A_1),\quad P^*_t(x_1,x_2;E_1\times A_2)=P^*_{2,t}(x_2,A_2) \tag{14}$$

for all $t\geq 0$, $x_1\in E_1$, $x_2\in E_2$, $A_1\in\mathcal{E}_1$, and $A_2\in\mathcal{E}_2$.

If Condition 1 is satisfied and $(P^*_{1,t})_{t\geq 0}$ and $(P^*_{2,t})_{t\geq 0}$ are both stochastic, then $(P^*_t)_{t\geq 0}$ is stochastic too [Che86, Theorem 37].

Remark. If $(P^*_t)_{t\geq 0}$ is not stochastic, then in general we can only expect an inequality similar to (13).

A variant of the theorem works for sub-Markov jump processes with bounded generator. Let K, K_1, K_2 be as before and $\tilde\lambda: E\to\mathbb{R}_+$, $\tilde\lambda_1: E_1\to\mathbb{R}_+$, $\tilde\lambda_2: E_2\to\mathbb{R}_+$ measurable with

$$\tilde\lambda(x)\geq K(x,E),\quad \tilde\lambda_1(x_1)\geq K_1(x_1,E_1),\quad \tilde\lambda_2(x_2)\geq K_2(x_2,E_2).$$

Define modified generators

$$(\tilde{L}f)(x) := -\tilde{\lambda}(x)f(x) + \int_E K(x,\mathrm{d}y)f(y)$$

similarly $\tilde{L}_1$, $\tilde{L}_2$. If the rate $\tilde{\lambda}(\cdot)$ is bounded, then $\tilde{P}_t := \exp(t\tilde{L})$ (see Appendix A.2) is the unique solution of the backward Kolmogorov equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\tilde{P}_t = \tilde{L}P_t.$$

Sub-Markov transition functions $\tilde{P}_{1,t}$ and $\tilde{P}_{2,t}$ are defined in a similar way.

Proposition 23. *Suppose that the rates $\tilde{\lambda}(\cdot)$, $\tilde{\lambda}_1(\cdot)$, $\tilde{\lambda}_2(\cdot)$ are bounded and that $\tilde{L}$, $\tilde{L}_1$, and $\tilde{L}_2$ satisfy Condition 1. Then*

$$\tilde{P}_t(x_1,x_2;A_1\times E_2) = \tilde{P}_{1,t}(x_1,A_1), \quad \tilde{P}_t(x_1,x_2;E_1\times A_2) = \tilde{P}_{2,t}(x_2,A_2) \tag{15}$$

for all $t\geq 0$, $x_1\in E_1$, $x_2\in E_2$, $A_1\in\mathcal{E}_1$, and $A_2\in\mathcal{E}_2$.

Proof. An induction over k yields $\tilde{L}^k(f_1\otimes\mathbf{1}) = (\tilde{L}_1^k f_1)\otimes\mathbf{1}$, for all $k\in\mathbb{N}$ and $f_1\in\mathscr{L}^\infty(E_1,\mathcal{E}_1)$. It follows that

$$\tilde{P}_t(f_1\otimes\mathbf{1}) = \exp(t\tilde{L})(f_1\otimes\mathbf{1}) = (\exp(t\tilde{L}_1)f_1)\otimes\mathbf{1} = (\tilde{P}_{1,t}f_1)\otimes\mathbf{1}.$$

We apply the equality to $f_1 = \mathbf{1}_{A_1}$ and evaluate the functions at the variable (x_1,x_2) and obtain the claim for the first marginal. The proof for the second marginal is similar. □

Let us return to the situation of Theorem 22. We drop the *-superscript because the minimal solutions are stochastic. Consider a tuple

$$\big(\Omega,\mathcal{F},(\mathbb{P}_x)_{x\in E},(X_t)_{t\geq 0}\big)$$

consisting of a measurable space $(\Omega,\mathcal{F})$, probability measures $\mathbb{P}_x$, and measurable maps $X_t:\Omega\to E$ such that under each $\mathbb{P}_x$, the process $(X_t)_{t\geq 0}$ is a jump process with transition function $(P_t)_{t\geq 0}$. Write $X_t = (X_{1,t},X_{2,t})\in E_1\times E_2$.

Proposition 24. *Under the conditions of Theorem 22: Under each $\mathbb{P}_{(x_1,x_2)}$, the component processes $(X_{1,t})_{t\geq 0}$ and $(X_{2,t})_{t\geq 0}$ are non-explosive jump processes with jump kernels K_1 and K_2 with initial values x_1 and x_2.*

Proof. Fix $x=(x_1,x_2)\in E$. Clearly $X_{1,t}$ is a measurable map from Ω to E_1 and $X_{1,0}=x_1$, $\mathbb{P}_{(x_1,x_2)}$-almost surely. Let $t,h\geq 0$ and $A_1\in\mathcal{E}_1$. The tower property of conditional expectations, the Markov property for $(X_t)_{t\geq 0}$, and Theorem 22 yield the $\mathbb{P}_x$-almost sure equalities

$$\begin{aligned}
\mathbb{P}_x\big[X_{1,t+h}\in A_1\times E\,\big|\,X_{1,s},s\leq t\big] &= \mathbb{E}_x\Big[\mathbb{E}_x\big[X_{t+h}\in A_1\times E\,\big|\,X_s,s\leq t\big]\Big|\,X_{1,s},s\leq t\Big]\\
&= \mathbb{E}_x\big[P_h(X_t,A_1\times E)\mid X_{1,s},s\leq t\big]\\
&= \mathbb{E}_x\big[P_{1,h}(X_{1,t},A_1)\mid X_{1,s},s\leq t\big]\\
&= P_{1,h}(X_{1,t},A_1).
\end{aligned}$$

Thus $(X_{1,t})_{t\geq 0}$ is a simple Markov process with transition function $(P_{1,t})_{t\geq 0}$. The proof for the second component $(X_{2,t})_{t\geq 0}$ is similar. $\square$

4. Truncation

Sometimes it is of interest to truncate a jump process and investigate transition functions where a part $H \subset E$ of the state space is taboo. The motivation is twofold. First, it can be genuinely of interest to investigate the probability

$$\mathbb{P}_x(X_t^\partial \in A, \forall s \leq t: \; X_s^\partial \in E \setminus H)$$

that the process goes from x to A in time time but without ever visiting the taboo set H. Equivalently, we may look at the process killed upon exiting $E \setminus H$,

$$Y_t^\partial = \begin{cases} X_t^\partial, & \forall s \leq t: \; X_s^\partial \in E \setminus H, \\ \partial, & \text{else.} \end{cases}$$

Second, truncation may allow us to deal with processes with bounded generators instead of the orginal process. For example, if $\sup_{x\in E} K(x,E) = \infty$, we may set $H_n := \{x \in E \mid K(x,E) > n\}$ and ask about processes defined in $E \setminus H_n$, and perhaps about convergence as $n \to \infty$.

Let K be a jump kernel, $(P_t^*)_{t\geq 0}$ the minimal solution of the backward Kolmogorov equation, and $\big(\Omega, \mathcal{F}, (\mathbb{P}_x)_{x\in E}, (X_t^\partial)_{t\geq 0}\big)$ the usual associated setup. Working with the jump-hold construction, we may assume without loss of generality that all sample paths $t \mapsto X_t(\omega)$ are piecewise constant. Here we agree to call a map $x : [0,\infty) \to E$ piecewise constant if for all $t \geq 0$, there exists $h > 0$ such that $x(\cdot)$ is constant on $[t, t+h)$. Fix $F \in \mathcal{E}$ and equip F with the trace of the σ-algebra $\mathcal{E}$. Then for every $t \geq 0$, the set

$$\{\omega \in \Omega \mid \forall s \in [0,t]: \; X_s^\partial(\omega) \in F\} = \{\omega \in \Omega \mid \forall s \in ([0,t) \cap \mathbb{Q}) \cup \{t\}: \; X_s^\partial(\omega) \in F\}$$

is measurable. For $x \in F$ and measurable $A \subset F$, define

$$Q_t(x,A) := \mathbb{P}_x\big(X_t \in A, \forall s \in [0,t]: \; X_s^\partial(\omega) \in F\big),$$

the probability of a transition from x to A in time t without ever leaving F. We may also think of Q_t as the transition function of a process killed upon exiting F. Remember $\lambda(x) = K(x,E)$.

Proposition 25. *$(Q_t)_{t\geq 0}$ defines a normal sub-Markov transition function on F. It solves the backward Kolmogorov equation*

$$\frac{\mathrm{d}}{\mathrm{d}t} Q_t(x,A) = -\lambda(x) Q_t(x,A) + \int_F K(x,\mathrm{d}y) Q_t(y,A)$$

for all $t \geq 0$, $x \in F$, and measurable $A \subset F$.

The proposition is complemented by the following approximation result. Let $E_n \nearrow E$, $E_n \in \mathcal{E}$. Let

$${}^{(n)}P_t(x,A) := \mathbb{P}_x\big(X_t \in A, \forall s \in [0,t]: \; X_s^\partial(\omega) \in E_n\big),$$

Then

$$\lim_{n\to\infty} {}^{(n)}P_t(x,A) = P_t^*(x,A),$$

for all $t \geq 0$, $x \in E$, and $A \in \mathcal{E}$. Choosing E_n in such a way that $\sup_{x\in E_n} K(x,E_n) < \infty$, we see that the minimal solution $(P_t^*)_{t\geq 0}$ is approximated by the (unique) solutions $({}^{(n)}P_t)_{t\geq 0}$ of backward equations with *bounded* generators.

Appendix A. Functional-analytic aspects

A.1. **Positivity-preserving contraction semi-group.** Let $(P_t)_{t\geq 0}$ be a normal Markov or sub-Markov transition function on $(E,\mathcal{E})$. Then each P_t induces a linear map from $\mathscr{L}^\infty(E,\mathcal{E})$, the space of bounded measurable functions, into itself. The linear operator is given by

$$(P_t f)(x) := \int_E P_t(x,\mathrm{d}y) f(y).$$

By a slight abuse of notation we use the same letter P_t for the kernel and for the operator $P_t : \mathscr{L}^\infty(E,\mathcal{E}) \to \mathscr{L}^\infty(E,\mathcal{E})$. We write $\mathbf{1}$ for the constant function that is everywhere equal to 1.

Proposition 26. *Let $(P_t)_{t\geq 0}$ be a normal Markov or sub-Markov transition function on $(E,\mathcal{E})$. Then the associated family of operators in $\mathscr{L}^\infty(E,\mathcal{E})$ satisfies the following:*

(a) $P_{t+s}f = P_tP_sf$ *for all* $s,t \geq 0$ *and all* $f \in \mathscr{L}^\infty(E,\mathcal{E})$.
(b) $P_0 f = f$, *for all* $f \in \mathscr{L}^\infty(E,\mathcal{E})$.
(c) *If* $0 \leq f \leq 1$ *pointwise on* E, *then also* $0 \leq P_t f \leq 1$ *pointwise on* E.
(d) *Let* $(f_n)_{n\in\mathbb{N}}$ *be a sequence in* $\mathscr{L}^\infty(E,\mathcal{E})$ *and* $f \in \mathscr{L}^\infty(E,\mathcal{E})$. *Then, if* $f_n \nearrow f$ *pointwise on* E, *then also* $P_t f_n \nearrow P_t f$ *pointwise on* E, *for all* $t \geq 0$.

Moreover $(P_t)_{t\geq 0}$ *is Markov if and only if* $P_t\mathbf{1} = \mathbf{1}$ *for all* $t \geq 0$.

Item (a) says that $(P_t)_{t\geq 0}$ defines a *semi-group* on E, it is inherited from the Chapman-Kolmogorov equations. Item (c) implies in particular that $(P_t)_{t\geq 0}$ is *positivity-preserving*, i.e., it maps non-negative functions to non-negative functions. Moreover, item (c) implies the inequality

$$\sup_{x\in E}\big|(P_t f)(x)\big| \leq \sup_{x\in E}|f(x)|$$

for all $f \in \mathscr{L}^\infty(E,\mathcal{E})$. As a consequence, each P_t defines a bounded linear operator in the Banach space $\mathscr{L}^\infty(E,\mathcal{E})$ of bounded functions equipped with the supremum norm

$$||f||_\infty := \sup_{x\in E}|f(x)|$$

moreover the operator norm is smaller or equal to 1: the family $(P_t)_{t\geq 0}$ is a *contraction semi-group*. It follows automatically that each P_t is continuous with respect to uniform convergence. Item (d) of the proposition says that in addition, each P_t is continuous with respect to pointwise monotone convergence of uniformly bounded, non-negative sequences.

Remark (Daniell-Stone theorem). Conversely, if $(P_t)_{t\geq 0}$ is a family of linear operators in $\mathscr{L}^\infty(E,\mathcal{E})$—a priori not necessarily given by a family of kernels—satisfying the items (a)-(d) from Proposition 26, then there exists a uniquely defined associated sub-Markov transition function. Here property (d) is very important. In combination with the *Daniell-Stone theorem*[2] [Bau68, Chapter VII.39] from measure theory, it allows to prove the existence of a measure $P_t(x,\mathrm{d}y)$ such that $(P_t f)(x) = \int_E P_t(x,\mathrm{d}y) f(y)$.

A.2. **Semi-groups with bounded generator. Dyson expansion.** Suppose that the jump kernel K is such that

$$\sup_{x\in E} \lambda(x) = \sup_{x\in E} K(x,E) =: C < \infty$$

(remember $\lambda(x) = K(x,E)$). Set $(Kf)(x) := \int_E K(x,\mathrm{d}y) f(y)$. Then for all $f \in \mathscr{L}^\infty(E,\mathcal{E})$,

$$||Kf||_\infty \leq \sup_{x\in E} \lambda(x) ||f||_\infty = C||f||_\infty,$$

hence K defines a bounded linear operator in $\mathscr{L}^\infty(E,\mathcal{E})$, which by a slight abuse of notation is designated with the same letter as the kernel K. Similarly,

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

defines a bounded operator $L : \mathscr{L}^\infty(E,\mathcal{E}) \to \mathscr{L}^\infty(E,\mathcal{E})$. As a consequence, the exponential series

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

converges in operator norm. This suggests to use

$$P_t := \exp(tL)$$

instead of the backward Kolmogorov equation in order to define a transition function. We shall see that $(P_t)_{t\geq 0}$ is actually exactly the semi-group associated with the minimal solution, moreover, the minimal solution is stochastic.

Let us pretend that we do not know anything about the minimal solution and examine $P_t = \exp(tL)$ as an object in its own right, as an alternative to what we did earlier. Clearly $P_0 f = f$ and $P_{t+s} f = P_t P_s f$ for all $s,t \geq 0$ and $f \in \mathscr{L}^\infty(E,\mathcal{E})$, i.e., we definitely have a semi-group. It is less immediate, however, that P_t is positivity-preserving. Indeed even when f is non-negative, because of the subtractions, Lf in general will not be non-negative and neither will $L^n f$, so a priori there seems to be no reason why the exponential series for $\exp(tL)f$ should yield a non-negative function.

The only reasonable way out is not to expand the negative part. Let M_λ be the multiplication operator

$$(M_\lambda f)(x) := \lambda(x) f(x).$$

[2]The theorem deals with the following question: If $I : \mathscr{F} \to \mathbb{R}$ is a linear map on a vector space of real-valued functions $\mathscr{F} \subset \{f \mid f : E \to \mathbb{R}\}$, can we find a σ-algebra $\mathcal{E}$ on E and a measure μ on $(E,\mathcal{E})$ such that (i) every $f \in \mathscr{F}$ is measurable with respect to $\mathcal{E}$ (and the Borel σ-algebra on $\mathbb{R}$), and (ii) $I(f) = \int_E f \mathrm{d}\mu$ for all $f \in \mathscr{F}$? Is the measure unique?

Notice

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

hence

$$L = K - M_\lambda.$$

If the function $\lambda(x)$ is constant, then M_λ is a scalar multiple of the identity operator, therefore $KM_\lambda = M_\lambda K$ and $P_t = \exp(tL) = \exp(-tM_\lambda)\exp(tK)$, which can be used to prove that P_t is indeed positivity-preserving.

In general, however, M_λ and K do not commute. But the following observation is useful: we have

$$\frac{\mathrm{d}}{\mathrm{d}t} \mathrm{e}^{tM_\lambda} \mathrm{e}^{tL} = \mathrm{e}^{tM_\lambda}\big(M_\lambda + L\big)\mathrm{e}^{tL} = \mathrm{e}^{tM_\lambda} K \mathrm{e}^{tL}.$$

Set

$$\boxed{A(t) := \mathrm{e}^{tM_\lambda}\mathrm{e}^{tL} = \mathrm{e}^{tM_\lambda} P_t, \quad K(t) := \mathrm{e}^{tM_\lambda} K \mathrm{e}^{-tM_\lambda}}$$

then

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}t} A(t) = K(t) A(t), \quad A(0) = \exp(0) = \mathrm{id}.}$$

Following the standard procedure for proving the existence of a solution to ordinary differential equations, we can write down the associated integral equation, interpret the integral equation as a fixed point problem, and then try a Picard iteration. Thus

$$A(t) = \mathrm{id} + \int_0^t K(s) A(s) \mathrm{d}s. \tag{16}$$

Multiplying with the operator $\exp(-tM_\lambda)$ from the left and inserting the definitions of $A(\cdot)$ and $K(\cdot)$, we obtain

$$P_t = \mathrm{e}^{-tM_\lambda} + \int_0^t \mathrm{e}^{-(t-s)M_\lambda} K P_s \mathrm{d}s. \tag{17}$$

This is precisely the integrated backward Kolmogorov equation from Proposition 13, written in a more functional-analytic way. The associated Picard iteration reads

$$P_t^{(0)} := \mathrm{e}^{-tM_\lambda}, \quad P_t^{(n+1)} = \mathrm{e}^{-tM_\lambda} + \int_0^t \mathrm{e}^{-(t-s)M_\lambda} K P_s^{(n)} \mathrm{d}s, \tag{18}$$

which is the recurrence relation from Lemma 9.

Lemma 27. *If $\sup_{x\in E} \lambda(x) = \sup_{x \in E} K(x, E) < \infty$, then for each $t \geq 0$, the sequence $(P_t^{(n)})_{n\in\mathbb{N}_0}$ converges to $P_t = \exp(tL)$ in operator norm, i.e., there is a sequnce $(\varepsilon_n(t))_{n\in\mathbb{N}_0}$ such that $\lim_{n\to\infty} \varepsilon_n(t) = 0$ and*

$$||P_t^{(n)} f - P_t f||_\infty \leq \varepsilon_n(t)\ ||f||_\infty \qquad \big(f \in \mathscr{L}^\infty(E, \mathcal{E})\big).$$

The proof is similar to existence proofs for ordinary differential equations.

Proposition 28. *If $\sup_{x\in E} \lambda(x) = \sup_{x\in E} K(x, E) < \infty$, then the family of operators $(P_t)_{t\geq 0} = (\exp(tL))_{t\geq 0}$ satisfies the properties (a), (b), (c) and (d) from Proposition 26, moreover $P_t \mathbf{1} = \mathbf{1}$ for all $t \geq 0$.*

Proof. The semigroup property $P_{t+s} = P_t P_s$ and the initial value $P_0 = \mathrm{id}$ follow from properties of the exponential. This proves (a) and (b). The constant function $\mathbf{1}$ satisfies

$$(L\mathbf{1})(x) = \int_E K(x, \mathrm{d}y)(1-1) = 0 \qquad (x \in E),$$

therefore

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

for all $t \geq 0$. For (c), let $f \in \mathscr{L}^\infty(E, \mathcal{E})$ be such that $0 \leq f \leq 1$ pointwise on E. Then $0 \leq P_t^{(n)} f \leq 1$ for all $t \geq 0$, $n \in \mathbb{N}_0$. Indeed this is clearly true for $n = 0$; and if it is true for n, then

$$\begin{aligned}\big(P_t^{(n+1)} f\big)(x) &= \mathrm{e}^{-\lambda(x)t} + \int_0^t \mathrm{e}^{-(t-s)\lambda(x)} \left(\int_E K(x, \mathrm{d}y)(P_t^{(n)} f)(x) \right) \mathrm{d}s \\ &\leq \mathrm{e}^{-\lambda(x)t} + \int_0^t \mathrm{e}^{-(t-s)\lambda(x)} \lambda(x) \mathrm{d}s \\ &= \mathrm{e}^{-\lambda(x)t} + \big(1 - \mathrm{e}^{-\lambda(x)t}\big) \\ &= 1.\end{aligned}$$

Similarly, $P_t^{n+1} f \geq 0$. Thus $0 \leq P_t^{(n+1)} f \leq 1$. This completes the inductive proof of the claim. By Lemma 27, $P_t^{(n)} f \to P_t f$ uniformly on E, therefore we also have $0 \leq P_t f \leq 1$. This proves (c).

For the continuity with respect to pointwise convergence of monotone increasing sequences, let $(f_k)_{k \in \mathbb{N}}$ be bounded non-negative functions with $f_k \nearrow f \in \mathscr{L}^\infty(E, \mathcal{E})$. An induction over n shows that $P_t^{(n)} f_k \nearrow P_t^{(n)} f$ as $k \to \infty$, for each $n \in \mathbb{N}_0$. The pointwise convergence $P_t f_k \to P_t f$ then follows from Lemma 27 and an $\varepsilon/3$-argument based on the bound

$$\begin{aligned}&\big|(P_t f_k)(x) - (P_t f)(x)\big| \\ &\qquad \leq ||P_t f_k - P_t^{(n)} f_k||_\infty + |(P_t^{(n)} f_k)(x) - (P_t^{(n)} f)(x)| + ||P_t f - P_t^{(n)} f||_\infty \\ &\qquad \leq 2\varepsilon_n(t)\, ||f||_\infty + |(P_t^{(n)} f_k)(x) - (P_t^{(n)} f)(x)|\end{aligned}$$

with $\varepsilon_n(t)$ as in Lemma 27. The pointwise convergence is monotone because $f_k \leq f_{k+1}$ implies $f_{k+1} - f_k \geq 0$ hence $P_t(f_{k+1} - f_k) \geq 0$ and $P_t f_k \leq P_t f_{k+1}$. This proves (d). $\square$

It follows from the Daniell-Stone theorem [Bau68, Chapter VII.39] that the semi-group $(P_t)_{t \geq 0}$ is associated with a sub-Markov transition function, i.e., there exists a family of kernels $P_t(x, \mathrm{d}y)$ such that $(P_t f)(x) = \int_E P_t(x, \mathrm{d}y) f(y)$ for all $t \geq 0$, $f \in \mathscr{L}^\infty(E, \mathcal{E})$, and $x \in E$, and the kernels form a sub-Markov transition function. Because of $P_t\mathbf{1} = \mathbf{1}$ and $P_t(x, E) = (P_t\mathbf{1})(x) = 1$, we are in fact dealing with a (Markov) transition function. In addition, the operator-valued differential equation $\frac{\mathrm{d}}{\mathrm{d}t} P_t = L P_t$ (differentiability in operator norm), which follows from the properties of the exponential of bounded operators, implies that the associated transition function satisfies the backward Kolmogorov equation.

Lemma 27 implies that the transition function associated with the exponential $(\exp(tL))_{t\geq 0}$ is in fact equal to the family $(P_t^*)_{t\geq 0}$ defined in Section 2.3. We have just checked that it is stochastic. The following proposition summarizes the relation between the functional-analytic approach for bounded jump kernels and the previous analytic approach to the backward Kolmogorov equation.

Proposition 29. *If* $\sup_{x\in E}\lambda(x) = \sup_{x\in E} K(x,E) < \infty$, *then the minimal solution* $(P_t^*)_{t\geq 0}$ *of the backward Kolmogorov equations is stochastic and it is the unique solution, moreover*

$$P_t^*(x,A) = \big(\mathrm{e}^{tL}\mathbb{1}_A\big)(x)$$

for all $t\geq 0$, $x\in E$, *and* $A\in\mathcal{E}$.

Remark (Dyson series and time-ordered exponential). We could also have written down a fixed point iteration starting from Eq. (16). The analogue of Lemma 27 for the operators $A(t)$ and $K(t)$, together with a representation similar to Eq. (5), yields

$$A(t) = \mathrm{id} + \sum_{n=1}^\infty \int_{0\leq t_1<\cdots<t_n\leq t} K(t_1)\cdots K(t_n)\mathrm{d}t_1\cdots\mathrm{d}t_n. \tag{19}$$

Physicists like to rewrite this series, with some abuse of notation, as follows. For $t_1,\ldots,t_n\in[0,t]$, let $\sigma\in\mathfrak{S}_n$ be a permutation such that $t_{\sigma(1)}\leq\cdots\leq t_{\sigma(n)}$. Write

$$\mathcal{T}\big[K(t_1)\cdots K(t_n)\big] := K(t_{\sigma(1)})\cdots K(t_{\sigma(n)})$$

and think of $\mathcal{T}$ as a time-ordering "operator". The notation is extended to integrals by

$$\mathcal{T}\Big[\int_{[0,t]^n} K(t_1)\cdots K(t_n)\mathrm{d}t_1\cdots\mathrm{d}t_n\Big] := \int_{[0,t]^n}\mathcal{T}\big[K(t_1)\cdots K(t_n)\big]\mathrm{d}t_1\cdots\mathrm{d}t_n.$$

Then

$$A(t) = \mathrm{id} + \sum_{n=1}^\infty \frac{1}{n!}\mathcal{T}\Big[\int_{[0,t]^n} K(t_1)\cdots K(t_n)\mathrm{d}t_1\cdots\mathrm{d}t_n\Big] \tag{20}$$

$$=: \mathcal{T}\Big[\exp\Big(\int_0^t K(s)\mathrm{d}s\Big)\Big]. \tag{21}$$

The last expression is sometimes referred to as *time-ordered exponential*, the series (20) and (19) as *Dyson series* or *Dyson expansion*, compare [RS75, Section X.12]. The time-ordered exponential connects to formulas for ordinary differential equations for real-valued functions. Indeed under mild regularity conditions on $k(t)$ the solution of the initial value problem for real-valued functions

$$a'(t) = k(t)a(t),\qquad a(0)=1$$

is given by

$$a(t) = \exp\Big(\int_0^t k(s)\mathrm{d}s\Big).$$

The considerations above tell us that a similar formula holds for the operator-valued equation $A'(t) = K(t)A(t)$ but the exponential has to be replaced by the time-ordered exponential (21), rigorously defined in terms of the Dyson expansion (19).

References

[Bau68] H. Bauer, *Wahrscheinlichkeitstheorie und Grundzüge der Maßtheorie*, Walter de Gruyter & Co., Berlin, 1968.

[BG68] R. M. Blumenthal and R. K. Getoor, *Markov processes and potential theory*, Pure and Applied Mathematics, Vol. 29, Academic Press, New York-London, 1968.

[Che86] M. Chen, *Coupling for jump processes*, Acta Math. Sinica (N.S.) **2** (1986), no. 2, 123–136.

[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.

[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.

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

[Kal97] O. Kallenberg, *Foundations of modern probability*, Probability and its Applications (New York), Springer-Verlag, New York, 1997.

[Kle08] A. Klenke, *Probability theory*, Universitext, Springer-Verlag London, Ltd., London, 2008, A comprehensive course, Translated from the 2006 German original.

[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.

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

[Nor98] J. R. Norris, *Markov chains*, Cambridge Series in Statistical and Probabilistic Mathematics, vol. 2, Cambridge University Press, Cambridge, 1998, Reprint of 1997 original.

[Pre75] C. Preston, *Spatial birth-and-death processes*, Bull. Inst. Internat. Statist. **46** (1975), no. 2, 371–391, 405–408 (1975), With discussion. MR 474532

[RS75] M. Reed and B. Simon, *Methods of modern mathematical physics. II. Fourier analysis, self-adjointness*, Academic Press [Harcourt Brace Jovanovich, Publishers], New York-London, 1975.

[Sch11] M. Scheutzow, *Stochastische Modelle*, Online lecture notes, TU Berlin. Available at `https://www3.math.tu-berlin.de/Vorlesungen/WS11/StoMo/stochmod.pdf`, 2011.

[Sch18] ———, *Stochastic Processes II / Wahrscheinlichkeitstheorie III*, Online lecture notes, TU Berlin. Available at `http://page.math.tu-berlin.de/~scheutzow/WT3main.pdf`, 2018.

Mathematisches Institut, Ludwig-Maximilians Universität, Theresienstr. 39, 80333 München, Germany