

# Markov Processes

• Tuesday and Thursday 12-14 ; Kleiner Hörsaal.

• Exercises : To be fixed. 1.008 Mo 12-14 or 1.007 Fr. 8-10.

• Assistant : Christian Kettner.

• Examination : Oral ; 50% exercise points. (groups up to 3 peoples).

## 1) Continuous time Markov chains

### 1.1) Markov chains, transition functions and infinitesimal description

• The processes we consider in this section are stochastic processes that live on a countable state space  $S$ .

For this reason we speak about chains. Moreover they satisfy the Markov property.

• Rem: The topology on  $S$  is the discrete topology, with respect to which all functions are continuous.

Def. 1.1) let us set:

•  $\mathcal{J} = \{ \text{right-continuous functions } w: \mathbb{R}_+ \rightarrow S, \text{ with finitely many jumps in any finite time interval} \}$

•  $\mathcal{F} = \text{the } \sigma\text{-algebra on } \mathcal{J} \text{ such that the mapping } w \rightarrow w(t) \text{ is measurable for each } t \geq 0$ .

• Denote  $X(t, \omega) := w(t)$  and the time shifted operator  $\theta_s : (\theta_s w)(t) := w(t+s)$ .

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

• Then, a  $\overset{\text{(continuous)}}{\text{stochastic process}} X$  with state space  $S$

| is a continuous-time Markov chain if:

(2)

③  $X$  is  $\mathcal{F}_{t \geq 0}$  adapted,

④  $\mathbb{P}^x(X(0)=x)=1$

⑤  $\mathbb{E}^x(Y_{t \geq 0} | \mathcal{F}_s) = \mathbb{E}^{X(s)}(Y), \mathbb{P}^x\text{-a.s.}, \forall s \in S, \text{ bounded measurable } Y \text{ on } \Omega.$

Revi: ③ is the Markov Property

•  $\mathbb{P}^x$  is the measure on the path when starting from  $x$ .

○  $\mathbb{P}^x(\cdot) = \mathbb{P}(\cdot | X(0)=x).$

• The next ingredient is the transition function.

Def. 12) A transition function is a function  $P_t(x,y)$

defined for  $t \geq 0, x, y \in S$  and satisfying:

⑥  $P_t(x,y) \geq 0$

•  $\sum_{y \in S} P_t(x,y) = 1$

•  $\lim_{t \downarrow 0} P_t(x,x) = P_0(x,x) = 1$

as well as the Chapman-Kolmogorov equations:

⑦  $P_{s+t}(x,y) = \sum_{z \in S} P_s(x,z) P_t(z,y), s, t \geq 0.$

Revi: Having the transition function, the finite-dimensional distributions of  $X$  starting from  $x$  are: for  $t_1 < t_2 < \dots < t_n$ ,

$$\mathbb{P}^x(X(t_1)=x_1, \dots, X(t_n)=x_n) = P_{t_1}(x, x_1) P_{t_2-t_1}(x_1, x_2) \cdots P_{t_n-t_{n-1}}(x_{n-1}, x_n)$$

and ⑦ ensures that they are consistent (and that we can apply Daniell-Kolmogorov Theorem).

,  $P_t(x,y)$  will be the probability that the Markov chain  $X$  starts at  $X(0)=x$  and is in  $y$  at time  $t$ .

- Now we come to the infinitesimal description.
- Since we would like to have only finitely many jumps in a finite time interval, one expects to have for small  $t$ ,

$$0 \leq P_t(x,y) = O(t) \text{ for } x \neq y$$

and  $0 \leq 1 - P_t(x,x) = O(t)$ ,

so that, for  $\overset{x \neq y}{q(x,y)} = \frac{d}{dt} P_t(x,y) \Big|_{t=0}$  would be the transition rate from  $x$  to  $y$ .

Motivated by this we define:

Def. 1.3) A  $Q$ -matrix is a collection  $q(x,y)$  of real numbers indexed by  $x, y \in S$  such that:

$$q(x,y) \geq 0 \text{ for } x \neq y$$

and

$$\sum_{y \in S} q(x,y) = 0$$

Notation:  $c(x) := -q(x,x) \geq 0$ .

This is the input when modelling!

Rec.: When  $S$  is finite, there is a one-to-one correspondence between:

- {
- . Markov chains,
- . Transition functions,
- .  $Q$ -Matrices

- What is the situation when  $S$  is countable but not finite?
- The answer is that it is not anymore the case.
- For Feller Processes (next chapter) it will be again the case.

Goal of this chapter: How to go from a  $Q$ -matrix to a transition function and then to the corresponding Markov chain.

## 1.2) Examples

1) Let  $P = (P(x,y))_{x,y \in S}$  be the transition matrix of a discrete-time Markov chain.

Let us consider an Poisson process of intensity 1 independent of the discrete chain.

Then, define a continuous-time M.C. to have jumps at the event of the Poisson process and the jumps are governed by  $P$ .

One can show (exercise):

$$P_t(x,y) = e^{-t} \sum_{n=0}^{\infty} \frac{t^n}{n!} (P^n)(x,y)$$

and that  $Q = P - I$ .

2)  $S = \{0, 1\} \Rightarrow$  The Q-matrix has the form

$$Q = \begin{pmatrix} -\beta & \beta \\ \gamma & -\gamma \end{pmatrix} \text{ for } \beta, \gamma > 0.$$

$$\Rightarrow P_t = \begin{pmatrix} \frac{\gamma}{\beta+\gamma} + \frac{\beta}{\beta+\gamma} e^{-t(\beta+\gamma)} & \frac{\beta}{\beta+\gamma} (1 - e^{-t(\beta+\gamma)}) \\ \frac{\gamma}{\beta+\gamma} (1 - e^{-t(\beta+\gamma)}) & \frac{\beta}{\beta+\gamma} + \frac{\gamma}{\beta+\gamma} e^{-t(\beta+\gamma)} \end{pmatrix}$$

3) If  $S$  is finite,  $Q$  and  $P_t = (P_t(x,y))_{x,y \in S}$  are finite matrices and it holds:

$$P_t = e^{tQ} = \sum_{n=0}^{\infty} \frac{t^n}{n!} Q^n$$

4) Birth-death chain on  $S = \{0, 1, 2, \dots\}$  with

$$q(k, k+1) = s_k, \quad q(k, k-1) = \gamma_k, \quad q(k, k) = -s_k - \gamma_k, \quad \gamma_0 = 0.$$

birth rate  
at  $k$       death rate  
at  $k$

and all other  $q(k, \ell)$  are zero.

- If  $X(t)$  = population size at time  $t$  and each individual gave birth at rate  $\gamma$  and dies at rate  $\lambda$ , then  $\beta_k = \gamma \cdot k$  and  $\lambda_k = \lambda \cdot k$ .  $\Rightarrow$  linear birth and death chain.  
 $\Rightarrow C(k) = -q(k, k) = (\lambda + \gamma) \cdot k$  is an unbounded function.

### 1.3) From a Markov chain to the infinitesimal description.

#### 1.3.1) Markov chain $\rightarrow$ transition function.

Thm 1.4) Given a Markov chain, let

$$P_t(x, y) := P^x(X(t)=y), \text{ for } t \geq 0, x, y \in S.$$

then:

- (a)  $P_t(x, y)$  is a transition function,
- (b)  $P_t(x, y)$  determines uniquely  $P^x$ .

Proof: (a)  $P_t(x, y) \geq 0$  and  $\sum_{y \in S} P_t(x, y) = 1$  are immediate.

let  $\tau = \inf \{t > 0 \mid X(t) \neq X(0)\}$ . Since the paths are right-continuous,

$\tau > 0$   $P^x$ -a.s. (for any of the  $x \in S$ ).

But  $P_t(x, x) \geq P^x(\tau > t) \Rightarrow \lim_{t \downarrow 0} P_t(x, x) \geq \lim_{t \downarrow 0} P^x(\tau > t) = 1$ .

• Chapman-Kolmogorov:

let  $Y = \prod_{\{X(t)=y\}}$  in the Markov Property

$$\begin{aligned} \Rightarrow E^x(Y_0 \otimes \dots \otimes Y_s) &= E^{X(s)}(Y) \\ &\stackrel{\parallel}{=} P^x(X(s+t)=y \mid F_s) \\ &\stackrel{\parallel}{=} P^x(X(s)=y) \\ &= P_t(x, y) \quad P^x\text{-a.s.} \end{aligned}$$

Taking  $E^x / \lim_{n \rightarrow \infty}$   $\overline{P}^x(X(n)=y) = P_t(x, y)$

b) By the Markov property, for  $0 < t_1 < \dots < t_n$ ,

$$\mathbb{P}^x(X(t_1)=x_1, \dots, X(t_n)=x_n) = P_{t_1}(x, x_1) P_{t_2-t_1}(x_1, x_2) \cdots P_{t_n-t_{n-1}}(x_{n-1}, x_n)$$

$\Rightarrow P_t$  determine the finite-dimensional distributions, which determine the full probability measures on  $(\mathcal{S}, \mathcal{F})$ . (# → theorem, see stochastic process lecture).

### 1.3.2) Transition functions $\rightarrow Q$ -matrices

- The first result is generic. We want to look at the (right)-derivative of  $P_t(x,y)$  in time.

Prop. 1-5) let  $P_t(x,y)$  be a transition function.

- (a) Then  $P_t(x,x) > 0$  for all  $t > 0$  and  $x \in S$ .
  - (b) If  $P_t(x,x) = 1$  for some  $t > 0$  and  $x \in S$ , then  $P_t(x,x) = 1$  for all  $t > 0$  and that given  $x$ .
  - (c)  $\forall x, y \in S$ ,  $P_t(y,x)$  is uniformly continuous in  $t$ . Indeed,
- $$|P_t(x,y) - P_s(x,y)| \leq 1 - P_{|t-s|}(x,x).$$

Rec: If (b) holds, then the M.C. stays for all times at  $x$ .

Proof: (a) From  $\lim_{t \downarrow 0} P_t(x,x) = 1$  it follows that

for  $t$  small enough,  $P_t(x,x) > 0$ .

But the Chapman-Kolmogorov equations implies  $P_{s+t}(x,x) \geq P_s(x,x) P_t(x,x)$

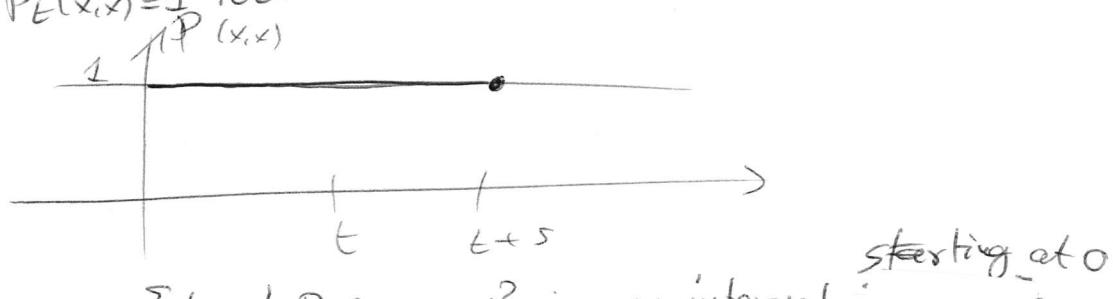
Thus  $P_t(x,x) > 0$  for all  $t > 0$ .

⑥ Chapman-Kolmogorov

(7)

$$\begin{aligned} \Rightarrow P_{S+t}(x, x) &= \sum_{z \in S} P_S(x, z) P_t(z, x) \\ &\leq P_S(x, x) P_t(x, x) + \underbrace{\sum_{z \in S \setminus \{x\}} P_S(x, z) \cdot 1}_{= 1 - P_S(x, x)} \\ &= 1 - P_S(x, x)(1 - P_t(x, x)). \end{aligned}$$

Thus, if  $P_{S+t}(x, x) = 1$ , since by ④  $P_S(x, x) > 0$ , it follows that  $P_t(x, x) = 1$  too.



$\Rightarrow \{t \geq 0 \mid P_t(x, x) = 1\}$  is an interval starting at 0

but by  $P_{S+t}(x, x) \geq P_S(x, x)P_t(x, x) \Rightarrow$  it is all  $\mathbb{R}_+$ .

⑦ By Chapman-Kolmogorov:

$$P_{t+s}(x, y) - P_t(x, y) = \overbrace{P_t(x, y)(P_S(x, y) - 1)}^{\leq 0} + \underbrace{\sum_{z \in S \setminus \{x\}} P_S(x, z) P_t(z, y)}$$

Further,  $|P_t(x, y)(P_S(x, y) - 1)| \leq |1 - P_S(x, y)| \geq 0$

and  $\left| \sum_{z \in S \setminus \{x\}} P_S(x, z) P_t(z, y) \right| \leq |1 - P_S(x, y)|$

$$\Rightarrow |P_{t+s}(x, y) - P_t(x, y)| \leq |1 - P_S(x, y)|$$

Replacing  $t+s \rightarrow t \Rightarrow |P_t(x, y) - P_S(x, y)| \leq |1 - P_{t-s}(x, y)|$ .

Finally, since  $\lim_{t \rightarrow s} |1 - P_{t-s}(x, y)| = 0$ , we have uniformly continuity.

(8)

We now use the results of Prop. 1.5 to show statements on the Q-matrix.

Theorem 1.6) let  $P_t(x,y)$  be a transition function.

a)  $\forall x \in S$ , the right derivative

$$c(x) = -q(x,x) = \frac{d}{dt} P_t(x,x) \Big|_{t=0} \in [0, \infty]$$

exists and satisfies

$$P_t(x,x) \geq e^{-c(x)t} \quad (\text{= Prob. of never leave } \{x\} \text{ during } [0,t]).$$

b) If  $c(x) < \infty$ , then for that  $x$  and all  $y \in S \setminus \{x\}$ ,  
the right derivative

$$q(x,y) = \frac{d}{dt} P_t(x,y) \Big|_{t=0} \in [0, \infty)$$

exists and

$$\sum_{y \in S} q(x,y) \leq 0.$$

c) If for some  $x \in S$ ,  $c(x) < \infty$  and  $\sum_{y \in S} q(x,y) = 0$ ,

then  $P_t(x,y)$  is  $C^1$  int for that  $x$  and every  $y \in S$ .

Further, it satisfies the Kolmogorov backward equations

$$\frac{d}{dt} P_t(x,y) = \sum_{z \in S} q(x,z) P_t(z,y)$$

- Remarks:
- If  $c(x)=\infty \Rightarrow P_t(x,x)=1$  for all  $t \geq 0$  : absorbing states or traps
  - If  $c(x)=0 \Rightarrow$  The M.C. leaves immediately  $x$  : instantaneous states
  - States with  $c(x) < \infty$  are called stable.

(3)

Proof: @ Consider the function

$$f(t) = -\ln P_t(x,x)(\gamma, \alpha)$$

By Prop 1.5 @ it is well-defined

By Prop 1.5 (C) :  $f$  is continuous ( $f(\epsilon+\delta) - f(\epsilon) =$

$$= \ln \left( \frac{P_{\epsilon+\delta}(x,x)}{P_\epsilon(x,x)} \right)$$

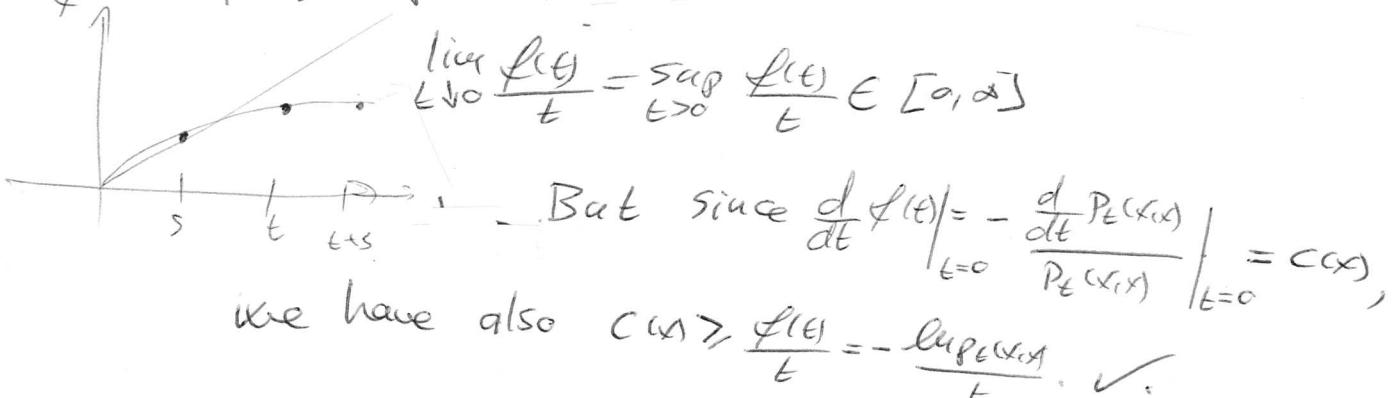
$$= \ln \left( 1 + \frac{P_{\epsilon+\delta}(x,x) - P_\epsilon(x,x)}{P_\epsilon(x,x)} \right)$$

$$\underset{\epsilon \rightarrow 0}{\equiv} \frac{1 - P_\epsilon(x,x)}{P_\epsilon(x,x)} \underset{\epsilon \rightarrow 0}{\rightarrow} 0.$$

By  $P_{\epsilon+t}(x,x) \geq P_\epsilon(x,x) P_t(x,x)$

$\Rightarrow f(\epsilon+s) \leq f(s) + f(\epsilon)$ , i.e.,  $f$  is subadditive.

This implies that



(b) Assume now  $c(x) < \infty$  (i.e., we can leave state  $x$ ).

By @  $\Rightarrow 1 - P_t(x,x) \leq 1 - e^{-c(x)t} \leq c(x)t$ .

$$\Rightarrow \sum_{y \in S \setminus \{x\}} \frac{P_t(x,y)}{t} = \frac{1 - P_t(x,x)}{t} \leq c(x).$$

$$\Rightarrow \limsup_{t \rightarrow 0} \frac{P_t(x,y)}{t} < \infty \text{ for } y \neq x.$$

Denote  $q(x,y)$  this limsup. We need to show that the limit exists.

Idea: discretize.

Take  $\delta > 0$  and  $n \in \mathbb{N}$ . (We look at the chain at discrete times  $n \cdot \delta$  with transition matrix  $P_{n\delta}(x,y)$ ).

Chap-Kolmogorov  $\Rightarrow P_{n\delta}(x,y) \geq \sum_{k=0}^{n-1} P_{\delta}(x,x) \cdot P_{\delta}(x,y) \cdot P_{n-k-1}\delta(y,y)$

$\overset{k \text{ steps}}{\circlearrowright} \xrightarrow{k+1 \text{ step}} \overset{n-k-1 \text{ steps}}{\circlearrowright}$

Since  $P_E(x,y) \geq e^{-C(x)t}$ , we have

$$\frac{P_{E\delta}(x,y)}{\delta} \geq \frac{P_S(x,y)}{\delta} \cdot e^{-C(x)\cdot \frac{\delta}{\delta}} \inf_{0 \leq t \leq \delta} P_S(y,y)$$

Now we choose a sequence of  $S, \delta > 0$  s.t.  $\frac{P_S(x,y)}{\delta} \rightarrow q(x,y)$ , and  $\delta \rightarrow 0$  s.t.  $\delta \rightarrow 0$ .

Then,  $\frac{P_E(x,y)}{\delta} \geq q(x,y) \cdot e^{-C(x)t} \inf_{0 \leq t \leq \delta} P_S(y,y)$  for  $t > 0$ .

$$\Rightarrow \liminf_{t \downarrow 0} \frac{P_E(x,y)}{t} \geq q(x,y) \stackrel{\text{def}}{=} \limsup_{t \downarrow 0} \frac{P_E(x,y)}{t} \Rightarrow \text{limit exists.}$$

Finally,  $\liminf_{t \downarrow 0} \sum_{y \in S \setminus \{x\}} \frac{P_E(x,y)}{t} \stackrel{(C(x))}{\geq} \sum_{y \in S \setminus \{x\}} \liminf_{t \downarrow 0} \frac{P_E(x,y)}{t} = \sum_{y \in S \setminus \{x\}} q(x,y)$

$$\Rightarrow \sum_{y \in S^c} q(x,y) = \sum_{y \in S \setminus \{x\}} q(x,y) - C(x) \leq 0.$$

### (c) Differentiability:

$$\text{Chap-Kolmogorov} \Rightarrow \frac{P_{E+\varepsilon}(x,y) - P_E(x,y)}{\varepsilon} - \sum_{z \in S} q(x,z) P_E(z,y) =$$

$$= \sum_{z \in S} \left( \underbrace{\frac{P_E(x,z) - P_0(x,z)}{\varepsilon}}_{\substack{\rightarrow 0 \text{ by } @ \\ \varepsilon \rightarrow 0}} - q(x,z) \right) P_E(z,y)$$

for  $z=x$ ,  
by (b) for  $z \neq x$ .

$\Rightarrow$  Need to control the tail of the sum: take  $T \subset S$  finite with  $x \in T$ . Then, by:

$$\sum_{z \notin T} \left| \frac{P_E(x,z)}{\varepsilon} - q(x,z) \right| P_E(z,y) \leq \sum_{z \notin T} \frac{P_E(x,z)}{\varepsilon} + \sum_{z \notin T} q(x,z)$$

$$= \frac{1}{\varepsilon} \cdot \left( 1 - \sum_{z \in T} P_E(x,z) \right) - \sum_{z \in T} q(x,z)$$

$\xrightarrow{\varepsilon \rightarrow 0} -2 \cdot \sum_{z \in T} q(x,z).$

(ii)

But since  $\sum_{z \in S} q(x, z) = 0$  and  $x \in T$ ,

by making  $T$  large we can make  $\sum_{z \in T} q(x, z)$  as small as desired  $\Rightarrow$  The tail contribution vanishes and indeed the right derivative exists and satisfies the backward equations.

Finally, by Prop. 1.50, the r.h.s. of the backwards eq. is continuous in  $t$ . But a continuous function with a continuous right-derivative is differentiable  $\#$

Remark: To prove the strong Markov property one needs to use:

- 1) The Markov property,
- 2) The right-continuity of the paths,
- 3) The continuity of  $\mathbb{E}^x(Y)$  for  $Y$  of the form  $\prod_{n=1}^N f_n(\omega(t_n))$ .

Theorem 1.7 (Strong MP) For any Markov chain, if  $Y_s(\omega)$  is bounded and jointly measurable on  $[0, \infty) \times \mathbb{R}$ , and  $\tau$  is a stopping time, then  $\forall x \in S$ ,

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

We will come back to this when studying Feller processes.

# 1.4) From infinitesimal description to Markov chain

(12)

- Goal: determine when  $Q$  determines a unique transition function  $P_t(x,y)$  that satisfies the Kolmogorov backward equation,

$$\frac{d}{dt} P_t(x,y) = \sum_{z \in S} q(x,z) P_t(z,y). \quad (\text{KBE})$$

For finite  $S$ ,  $(P_t(x,y))_{x,y \in S} = e^{tQ} \equiv \sum_{n \geq 0} \frac{t^n}{n!} Q^n$ .

## 1.4.1) The backward equation

Prop. 1.8) Let  $P_t(x,y)$  be a uniformly bounded function of  $t, x, y$ . Then the following statements are equivalents:

- (a)  $P_t(x,y)$  is continuously differentiable in  $t$ , satisfies (KBE) and  $P_0(x,y) = \delta_{x,y}$  for each  $x,y \in S$

- (b)  $P_t(x,y)$  is continuous in  $t$  and satisfies:

$$P_t(x,y) = \delta_{x,y} e^{-Cx t} + \int_0^t ds e^{-Cx(t-s)} \sum_{z \in S \setminus \{x\}} q(x,z) P_s(z,y) \quad \text{(*)}$$

for  $x,y \in S$  and  $t \geq 0$ .

Proof:  $\text{(a)} \Rightarrow \text{(b)}$ : (KBE)  $\Leftrightarrow \frac{d}{dt} P_t(x,y) + C(x) P_t(x,y) = \sum_{z \in S \setminus \{x\}} q(x,z) P_t(z,y)$  (12)

$$\Rightarrow \frac{d}{dt} (e^{Cx t} P_t(x,y)) = e^{Cx t} \cdot ( ) = e^{Cx t} \sum_{z \in S \setminus \{x\}} q(x,z) P_t(z,y)$$

$$\Rightarrow \int_0^t \frac{d}{ds} (e^{Cxs} P_s(x,y)) ds = \int_0^t ds e^{Cxs} \sum_{z \in S \setminus \{x\}} q(x,z) P_s(z,y)$$

$$\leq e^{Cx t} P_t(x,y) - \delta_{x,y} \quad \text{(*)}$$

$$\Rightarrow P_t(x,y) = \delta_{x,y} e^{-Cx t} + \int_0^t ds e^{-Cx(t-s)} \sum_{z \in S \setminus \{x\}} q(x,z) P_s(z,y).$$

(b)  $\Rightarrow$  (a): The r.h.s. of (\*) is  $C^1 \Rightarrow$  also the l.h.s. is, i.e.,  $P_t(x,y)$  is  $C^1$  (in time). Setting  $t=0$ ,  $P_0(x,y) = \delta_{x,y}$  ✓.

Comparing  $\frac{d}{dt} P_t(x,y) \Rightarrow (*) \Leftrightarrow (\text{KBE})$  ✓

- The reason we stated Prop. 1.8 is that the equation in (b) contains only positive coefficients and we can try to use it to compute its fix point(s).

- Define  $P_t^{(0)}(x,y) = 0$  for all  $x,y \in S$ ,  $t \geq 0$ .

Set 
$$P_t^{(n+1)}(x,y) := \delta_{xy} e^{-cx t} + \int_0^t ds e^{-c(x)(t-s)} \sum_{z \in S \setminus \{x\}} q(x,z) P_s^{(n)}(z,y)$$
 for  $n \geq 0$ .

- $P_t^{(n)}$  turns out to be monotone and bounded:

Lemma 1.9) It holds:

- (a)  $P_t^{(n)}(x,y) \geq 0$ , for  $t \geq 0$ ,  $x,y \in S$ ,
- (b)  $\sum_{y \in S} P_t^{(n)}(x,y) \leq 1$ , for  $t \geq 0$ ,  $x \in S$ ,
- (c)  $P_t^{(n+1)}(x,y) \geq P_t^{(n)}(x,y)$ , for  $t \geq 0$ ,  $y \neq 0$ ,  $x,y \in S$ .

Proof: It is trivial (use induction). #

Because of (c),  $\lim_{n \rightarrow \infty} P_t^{(n)}(x,y)$  exists.

Def. 1.10) We denote by  $P_t^*(x,y) := \lim_{n \rightarrow \infty} P_t^{(n)}(x,y)$  and we call it the minimal solution to (KBE).

The reason of calling  $P_t^*$  the minimal solution will become clear very soon.

Rem.:  $P_t^*(x,y)$  is a function entirely determined by the Q-matrix.

First let us state some properties of  $P_t^*(x,y)$ :

↑ 17.10.2013

Thm 1.11)  $P_t^*(x,y)$  satisfies:

- (a)  $P_t^*(x,y) \geq 0$ , for  $t \geq 0, x,y \in S'$ ,
- (b)  $\sum_{y \in S'} P_t^*(x,y) \leq 1$ , for  $t \geq 0, x \in S'$ ,
- (c) the Chapman-Kolmogorov equations,
- (d) the Kolmogorov backward equations.

Proof: (a) follows from Lemma 1.9(a)

(b) (i) (b).

(d) take the limit in  $P_t^{(n+1)}(x,y) = \dots$

and use monotone convergence (possible by Lemma 1.9(c))

The limit is continuous so by Prop 1.8 (b)  $\Rightarrow$  Prop 1.8(a), i.e;  $P_t^*(x,y)$  satisfies the (KBE).

(c) To prove this we need to work a bit more:

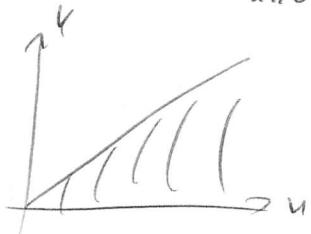
Define  $D_t^{(n)}(x,y) = P_t^{(n+1)}(x,y) - P_t^{(n)}(x,y)$ . By Lemma 1.9(c),

$D_t^{(n)}(x,y) \geq 0$ . In Lemma 1.12 below we will prove

that  $D_{t+s}^{(n)}(x,y) = \sum_{z \in S} \sum_{k=0}^n D_s^{(k)}(x,z) D_t^{(n-k)}(z,y)$ .

Further,  $P_t^*(x,y) = \sum_{u \geq 0} D_t^{(u)}(x,y)$ . Thus

$$P_{t+s}^*(x,y) = \sum_{u \geq 0} D_{t+s}^{(u)}(x,y) = \sum_{z \in S} \sum_{u \geq 0} \sum_{k=0}^n D_s^{(k)}(x,z) D_t^{(u-k)}(z,y)$$



$$= \sum_{z \in S} \underbrace{\sum_{k \geq 0} D_s^{(k)}(x,z)}_{\equiv P_s^*(x,z)} \underbrace{\sum_{u \geq k} D_t^{(u-k)}(z,y)}_{\equiv P_t^*(z,y)}$$

$$= \sum_{z \in S} P_s^*(x,z) P_t^*(z,y). \quad \#$$

It remains to prove:

Lemma 1.12) It holds

$$\Delta_{t+s}^{(n)}(x,y) = \sum_{z \in S} \sum_{k=0}^n \Delta_s^{(k)}(x,z) \Delta_t^{(n-k)}(z,y). \quad (*)$$

Proof: To prove  $(*)$  we prove that (for  $\lambda, \mu > 0$ )

$$(*) \quad \int_0^\infty ds \int_0^\infty dt e^{-\lambda s} e^{-\mu t} \Delta_{t+s}^{(n)}(x,y) = \int_0^\infty ds \int_0^\infty dt e^{-\lambda s} e^{-\mu t} \quad (\text{r.h.s. of } (*)).$$

Denote by  $\Psi_{n,\lambda}(x,y) = \int_0^\infty e^{-\lambda t} \Delta_t^{(n)}(x,y) dt$  the Laplace transform of  $\Delta_t^{(n)}(x,y)$ .

Then,  $(*)$  becomes:

$$(*) \quad \frac{\Psi_{n,\lambda}(x,y) - \Psi_{n,0}(x,y)}{\lambda - \mu} = \sum_{z \in S} \sum_{k=0}^n \Psi_{k,0}(x,z) \Psi_{n-k,\lambda}(z,y).$$

Another identity is  $\Psi_{n+1,\lambda}(x,y) = \sum_{z \in S \setminus \{x\}} \frac{q(x,z)}{\lambda + c(x)} \Psi_{n,0}(z,y)$   $(***)$

which is proven easily by introducing the definition of

$\Delta_t^{(n)}(x,y) = P_t^{(n+1)}(x,y) - P_t^{(n)}(x,y) = \int_0^t ds e^{-c(x)(t-s)} \sum_{z \neq x} q(x,z) \cdot \Delta_s^{(n-1)}(z,y)$  in the Laplace transform (and performing first the integral over  $t$ ).

In matrix-form,  $(**)$  becomes

$$\Psi_{n,\lambda}(x,y) = (A_\lambda)^n \Psi_{0,\lambda}(x,y) \quad \text{where } A_\lambda(x,y) = \frac{q(x,y)}{\lambda + c(x)} \Delta_{\{x\}}(x,y).$$

Then, l.h.s. of  $(*)$  writes

$$\frac{(A_\lambda)^n \Psi_{0,\lambda}(x,y) - (A_\lambda)^n \Psi_{0,0}(x,y)}{\lambda - \mu} = \frac{(A_\lambda)^n (\Psi_{0,\lambda}(x,y) - \Psi_{0,0}(x,y))}{\lambda - \mu}$$

(Telescopic)  $+ \sum_{k=0}^{n-1} (A_\lambda)^k \frac{A_\lambda - A_\lambda}{\lambda - \mu} A_\lambda^{n-k-1} \Psi_{0,\lambda}(x,y)$

$$\text{Now, } \Psi_{0,\lambda}(x,y) = \frac{\delta_{x,y}}{\lambda + c(x)} \Rightarrow \frac{\Psi_{0,\lambda}(x,y) - \Psi_{0,0}(x,y)}{\lambda - \mu} = \Psi_{0,0}(x,y) \cdot \Psi_{0,\lambda}(x,y)$$

(16)

and  $\frac{A_\mu - A_\lambda}{\lambda - \mu} (x, y) = \frac{\mathbb{I}_{x \neq y}}{\lambda + c(x)} \frac{q(x, y)}{(\lambda + c(x))(\lambda + c(y))} = \sum_z \delta_{x,z} \frac{1}{\lambda + c(x)} \cdot \frac{\mathbb{I}_{(z \neq y)}}{\lambda + c(z)} \frac{q(z, y)}{\lambda + c(z)}$   
 $= (\varphi_{0,x} \cdot A_\mu)(x, y).$

$$\Rightarrow \text{l.h.s. of } \textcircled{*} = (A_\lambda)^n \varphi_{0,x} \varphi_{0,y} + \sum_{k=0}^{n-1} (A_\lambda)^k \varphi_{0,x} (A_\mu)^{n-k} \varphi_{0,y}$$

$$= \sum_{k=0}^n \underbrace{(A_\lambda)^k \varphi_{0,x}}_{=\varphi_{k,x}} \underbrace{(A_\mu)^{n-k} \varphi_{0,y}}_{=\varphi_{n-k,y}} = \text{r.h.s. of } \textcircled{*}. \quad \#$$

Rem.: The issue that it is not yet solved by Thm 1.11 is whether  $\sum_{y \in S} P_t^*(x, y) = 1$  or not.

There are indeed situations where  $\sum_{y \in S} P_t^*(x, y) < 1$ , i.e., the solution  $P_t^*(x, y)$  is substochastic.

In this situation one loses mass "at infinity" in finite times and there will not be a unique transition function for a given Q-matrix.

Thm 1.13 (a) If  $P_t(x, y)$  is a non-negative solution to the (KBE) satisfying  $P_0(x, y) = \delta_{x,y}$ , then

$$P_t(x, y) \geq P_t^*(x, y) \quad \text{for all } t \geq 0, x, y \in S.$$

(b) If  $\sum_{y \in S} P_t^*(x, y) = 1$  for all  $t \geq 0$ ,  $x \in S$ ,

then  $P_t^*(x, y)$  is the unique transition function satisfying the (KBE).

Rem.: Part (a) of the theorem explains why  $P_t(x, y)$  is called minimal solution.

Proof: @ let us show that  $P_t(x,y) \geq P_t^{(n)}(x,y)$  (by induction).

For  $n=0$  it is true, since  $P_t^{(0)}(x,y)=0$ .

Assume the inequality to hold for some  $n$ .

Then:

$$P_t^{(n+1)}(x,y) \leq S_{x,y} e^{-\alpha x t} + \int_0^t ds e^{-\alpha x(t-s)} \sum_{z \neq x} q(x,z) P_s(z,y)$$

$\stackrel{P_t \text{ satisfies KBE}}{=} P_t(x,y)$

Prop. 1.8

(b) By Thm 1.11,  $P_t^*(x,y)$  satisfies the (KBE).

By part @, if  $P_t(x,y)$  is another solution, then

$$0 \leq \sum_{y \in S} (P_t(x,y) - P_t^*(x,y)) = 1 - 1 = 0 \Rightarrow P_t(x,y) = P_t^*(x,y).$$

for all  $t \geq 0, x, y \in S$ . #

The question is now what can "go wrong", what happens in the case that  $P_t^*(x,y)$  is only substochastic but not stochastic.

Example: Consider a pure birth chain on  $S = \{0, 1, 2, \dots\}$

with  $q(i,j) = \begin{cases} -\beta_i, & \text{if } j=i, \\ \beta_i, & \text{if } j=i+1, \\ 0, & \text{otherwise,} \end{cases}$

where  $\beta_i > 0$  for each  $i \geq 0$ .

One can show (see exercises) that

$$\sum_{y \in S} P_t^*(x,y) = 1 \text{ for all } t \geq 0, x \in S \quad \textcircled{D}$$

if and only if  $\sum_{k=0}^{\infty} \frac{1}{\beta_k} = \infty$ .

What happens if  $\sum_{k=0}^{\infty} \frac{1}{\beta_k} < \infty$  is that a chain starting at 0 goes to infinity in a finite time (one says that the chain "explodes").

Def: If  $\textcircled{D}$  holds  $\Rightarrow$  the chain is called non-explosive

If  $\textcircled{D}$  does not hold  $\Rightarrow$  there are  $\infty$ -many transition functions

## 1.4.2) The probabilistic construction

- . In the previous section (1.4.1) we derived a transition function from the Q-matrix (in the case that  $P_t^*(x, \cdot)$  is stochastic). For that purpose we have introduced a sequence  $P_t^{(n)}(x, y)$ , satisfying the relations

$$\begin{cases} P_t^{(0)}(x, y) = 0 \\ P_t^{(n+1)}(x, y) = \delta_{x,y} e^{-c(x)t} + \int_0^t ds e^{-c(x)(t-s)} \sum_{z \in S^1 \setminus \{x\}} q(x, z) P_s^{(n)}(z, y) \end{cases}$$

- . In this chapter, between other results, we will see that these  $P_t^{(n)}(x, y)$  have also a probabilistic meaning.

- . Lemma 1.14) Let  $X(\epsilon)$  be a continuous time Markov Chain and let us define the stopping time

$$\tau = \inf \{ t \geq 0 : X(t) \neq X(0) \},$$

that is, the time of the first jump.

Then,  $\mathbb{P}^X(\tau > t) = e^{-c(x)t}$  for some  $0 \leq c(x) \leq \infty$ .

Proof: The M.P.  $\mathbb{E}^X(Y_{0 \leq s \leq t} | \mathcal{F}_s) = \mathbb{E}^{X(s)}(Y)$ , dP-a.s.

for  $Y = \prod_{\{w: w(r)=x, 0 \leq r \leq t-s\}} 1$ .

This gives, for fixed  $s \leq t$  and  $x \in S^1$ :

$$\Rightarrow \mathbb{P}^X(X(r)=x, s \leq r \leq t | \mathcal{F}_s) = \mathbb{P}^{X(s)}(\tau > t-s), \text{P-a.s.}$$

$$\text{Then, } \mathbb{E}^X(\prod_{\{\tau > s\}} \mathbb{P}^X(X(r)=x, s \leq r \leq t | \mathcal{F}_s)) = \mathbb{E}^X(\mathbb{P}^{X(s)}(\tau > t-s) \cdot \prod_{\{\tau > s\}} 1)$$

$$\mathbb{P}^X(\{\tau > t\})$$

$$\mathbb{P}^X(\tau > t)$$

$$\mathbb{P}^X(\tau > t-s) \mathbb{P}^X(\tau > s)$$

$$\mathbb{P}^X(\tau > t-s) \mathbb{P}^X(\tau > s)$$

We have found:

$$\mathbb{P}^X(\tau > t) = \mathbb{P}^X(\tau > t-s) \mathbb{P}^X(s > s).$$

This means that  $\mathbb{P}^X(\tau > t) = 0$  for all  $t > 0$  (i.e.  $c(x) = \infty$ ), or  $\mathbb{P}^X(\tau > t) > 0$  for all  $t > 0$ .

In this second case,

$$f(t) := \ln \mathbb{P}^X(\tau > t) = \ln \mathbb{P}^X(\tau > t-s) + \mathbb{P}^X(s > s).$$

As  $f(t)$  is monotone (decreasing)  $\Rightarrow f(t) = ct$ , i.e.,  $\exists c(x) < \infty$  s.t.  $\mathbb{P}^X(\tau > t) = e^{-c(x)t}$ .  $\#$

Now, let us go to the construction:

let, for  $x \in S$ ,  $i \geq 1$ ,  $T_{x,i}$  be independent exponential distributed random variables with parameter  $c(x)$ .

• Define a discrete time Markov chains follows:

For a given Q-matrix:

$$\begin{cases} \text{If } c(x)=0 \Rightarrow p(x,y)=1 \text{ and } p(x,y)=0 \text{ for } y \neq x, \\ \text{If } c(x)>0 \Rightarrow p(x,y)=\begin{cases} \frac{q(x,y)}{c(x)}, & \text{if } y \neq x, \\ 0, & \text{if } y=x \end{cases} \end{cases}$$

$$\Rightarrow \sum_{y \in S} p(x,y) = 1. \quad \text{let } \pi(x) \text{ be a probability measure on } S. \quad (\text{the initial distribution}), \text{ indep. of all other}$$

• let  $Z_n$  be a discrete M.C. on  $S$  starting from  $x$ , i.e., R.V.  $Z_0 = x$ ,  $Z_1, Z_2, \dots, Z_n$   
 $\mathbb{P}(Z_0=x_0, \dots, Z_n=x_n) = \prod_{i=1}^n p(x_i, x_{i+1}) = p(x_0, x_1) p(x_1, x_2) \dots p(x_{n-1}, x_n).$

This is the chain of the visited sites.

• Then, the continuous time Markov chain is

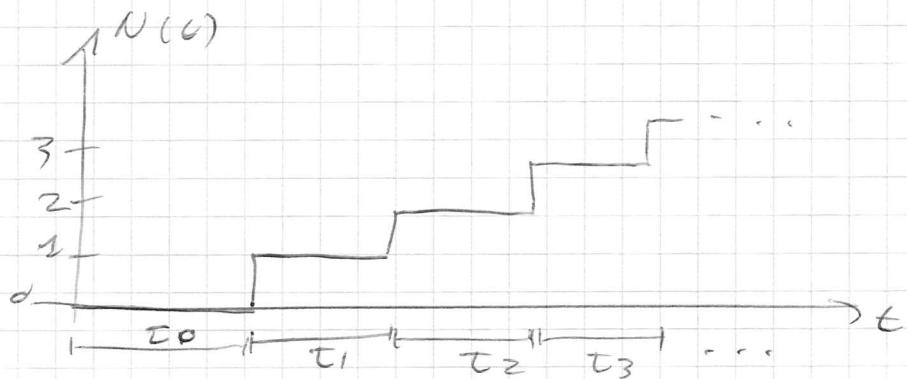
obtained by setting the waiting times  $\tau_0, \tau_1, \dots$

at sites  $x_0, x_1, \dots$  to be exponentially distributed with parameter  $c(x)$ .

The  $\tau_u$ 's are some of the  $\tau_{x,i}$ 's, namely,  $\tau_{x,i}$  is the waiting time of the  $i$ th visit at  $x$ .

$$\text{Define } N(\epsilon) = \begin{cases} \min\{u \geq 0 : \tau_0 + \dots + \tau_u > \epsilon\}, & \text{if } \sum_{u \geq 0} \tau_u > \epsilon, \\ \infty & \text{otherwise.} \end{cases}$$

$N(\epsilon)$  is the number of jumps in  $[0, \epsilon]$ .



We are ready to define the continuous-time Markov chain by:

Def. 1.15)

$$X(t) = Z_{N(t)} \text{ on } \{N(t) < \infty\}.$$

Rem.: If  $N(t) = \infty \Rightarrow$  there is an "explosion" before time  $t$  and the chain is not uniquely defined (see later).

Now we give the probabilistic interpretation of  $P_t^{(n)}(x,y)$ .

Prop 1.16) It holds:

$$(a) P_t^{(n)}(x,y) = \mathbb{P}(X(t)=y, N(t) \leq n | X(0)=x)$$

$$(b) P_t^*(x,y) = \mathbb{P}(X(t)=y, N(t) < \infty | X(0)=x)$$

$$(c) \sum_{y \in S} P_t^*(x,y) = \mathbb{P}(N(t) < \infty | X(0)=x).$$

Proof: Clearly  $\textcircled{b} = \lim_{n \rightarrow \infty} \textcircled{a}$  and  $\textcircled{c} = \sum_{y \in S} \textcircled{b}$ .

So, we need to show  $\textcircled{a}$ .

Consider:

$$\mathbb{P}(X(t)=y, N(t) \leq n+1 \mid \tau_0=s, Z_1=z, Z_0=x)$$

$$= \begin{cases} \delta_{x,y}, & \text{if } s > t, \\ \mathbb{P}(X(t-s)=y, N(t-s) \leq n \mid X(0)=z), & \text{if } s \leq t. \end{cases}$$

(from the Markov property of the discrete chain  
+ "forgetfulness" of the Poisson process giving  
the holding times.)

• By the construction of the chain,

$$\begin{aligned} & \mathbb{P}(Z_0=x_0, \dots, Z_n=x_n, \tau_0 > t_0, \tau_1 > t_1, \dots, \tau_n > t_n) \\ &= \pi(x_0) P(X_1, x_1) \dots P(X_{n-1}, x_n) \cdot e^{-\frac{(x_0-t_0)}{c(x_0)} - \frac{(x_1-t_1)}{c(x_1)} - \dots - \frac{(x_n-t_n)}{c(x_n)}}. \end{aligned}$$

Thus, taking expectations of  $\textcircled{a}$  we get:

$$\mathbb{P}(X(t)=y, N(t) \leq n+1 \mid X(0)=x) = \delta_{x,y} e^{-\frac{(x)}{c(x)}t}$$

$$+ \int_0^t ds \underbrace{\frac{-c(x)s}{e}}_{\text{density of exponential v.v.}} \sum_{z \neq x} p(x,z).$$

since  $P(X_0=x)$   
per def.

$$\cdot \mathbb{P}(X(t-s)=y, N(t-s) \leq n \mid X(0)=z)$$

Finally, since  $c(x)p(x,z) = q(x,z)$  and changing the variable  $s \mapsto t-s$  leads to the same recursion as for  $P_t^{(n)}(x,y)$ . As at  $n=0$  they agree w.

We compute  $\sum_{z \in S} \int_0^\infty \mathbb{P}(\tau_0=s, Z_1=z) \cdot \textcircled{a}$  on both left and right parts.

We see from Prop. 1.16 the following:

Theorem 1.17) The following statements are equivalent:

- (a) The minimal solution  $P_t^*(x,y)$  to the (KBE) is stochastic.
- (b)  $\mathbb{P}(N(t) < \infty) = 1, \forall t \geq 0,$
- (c)  $\sum_{n \geq 0} \tau_n = \infty \text{ a.s.}$
- (d)  $\sum_{n \geq 0} \frac{1}{C(z_n)} = \infty \text{ a.s.}$

Proof: (a)  $\Rightarrow$  (b): by Prop 1.16 (c).

(b)  $\Rightarrow$  (c): by  $\{N(t) < \infty\} = \left\{ \sum_{k \geq 0} \tau_k > t \right\}$

(c)  $\Leftrightarrow$  (d): let  $\lambda > 0$ . Then,

$$\begin{aligned} \mathbb{E}\left(e^{-\lambda \sum_{k=0}^n \tau_k} \mid z_0, z_1, \dots\right) &= \prod_{k=0}^n \underbrace{\int_0^\infty ds_k e^{-\lambda s_k} \cdot C(z_k) e^{-C(z_k)s_k}}_C \\ &= e^{\sum_{k=0}^n \ln\left(1 - \frac{\lambda}{C(z_k)+\lambda}\right)} \end{aligned}$$

$\boxed{=} \frac{C(z_n)}{C(z_0)+\lambda}.$

- Take  $\mathbb{E}(\dots)$  and  $n \rightarrow \infty$ :

$$\mathbb{E}\left(e^{-\lambda \sum_{k=0}^n \tau_k}\right) = \mathbb{E}\left(e^{\sum_{k=0}^\infty \ln\left(1 - \frac{\lambda}{C(z_k)+\lambda}\right)}\right) \underset{\lambda \text{ small}}{\approx} \mathbb{E}\left(e^{\sum_{k=0}^\infty \frac{-\lambda}{C(z_k)}}\right)$$

Take now  $\lambda \downarrow 0$

for  $\lambda$  small.

$$\Rightarrow \mathbb{E}\left(-\frac{1}{\lambda} \left(\sum_{k \geq 0} \tau_k < \infty\right)\right) = \mathbb{E}\left(-\frac{1}{\lambda} \left\{ \sum_{k \geq 0} \frac{1}{C(z_k)} \right\}\right), \text{ i.e.,}$$

we get (c)  $\Leftrightarrow$  (d).

#

The statements of Theorem 1.17 are not always simple to check. There are two cases where it is the case:

Cor. 1.18) The minimal solution  $P_t^*(x,y)$  to (KBE) is stochastic if either

$$\textcircled{a} \quad \sup_{x \in S} C(x) < \infty$$

or

\textcircled{b} the discrete time M.C.  $Z_n$  is irreducible and recurrent.

(i.e., every state is reachable by any state and eventually, it comes back to any starting point with proba. 1).

Proof.: \textcircled{a}:  $\frac{1}{C(x)} \geq \varepsilon$  for some  $\varepsilon \Rightarrow \sum_{n=1} \frac{1}{C(Z_n)} \geq \varepsilon \cdot \sum_{n=1} 1 = \infty$ .

\textcircled{b}:  $Z_n = x$  infinitely often a.s., fix. Thus,

$$\frac{1}{C(x)} \text{ occurs } \infty \text{ often in } \sum_{n=1} \frac{1}{C(Z_n)}. \quad \#$$

We are finally ready for the main statement of this section.

Their 1.19) Suppose that the minimal solution  $P_t^*(x,y)$

to (KBE) is stochastic.

Then,  $P_t^*(x,y)$  is a transition function and there is a unique Markov chain satisfying  $\mathbb{P}^x(X(t)=y) = P_t^*(x,y)$ .

Proof: By  $P_t^*(x,y) = \mathbb{P}(X(t)=y, N(e) \leq t | X(0)=x)$

we have  $P_t^*(x,x) \geq \mathbb{P}(N(e)=0 | X(0)=x) = e^{-\text{cost}}$ .

Thus,  $\lim_{t \downarrow 0} P_t^*(x,x) = 1$ , which is the missing property for instance in Thm 1.11.

The M.C. is then defined by setting the family  $\{P_t^*\}$  to be the distributions of  $(X(\cdot) | X(0)=x)$ , where  $X(\cdot)$  is defined in Def 1.15.

Uniqueness comes from Thm 1.4(b).

It remains to verify the Markov property.

Without entering in the details (if you are interested, look at the book, Remark 1.48 and Thm 1.46), one needs to have:

right-continuity of the profiles: ok by construction,

for  $\phi(y,u) = \mathbb{E}\left[\phi_i(X(t_{i-1})) \dots \phi_u(X(t_{u-1}))\right]$ : (4a)

joint continuity of  $\phi(y,u)$ : ok by Prop 1.5

$\phi(y,0) = \mathbb{E}\left(\prod_{u=1}^U \phi_u(X(t_u))\right)$  + choice of the discrete topology over  $S^U$ .

the Chapman-Kolmogorov eqn.: ok.

#

Remark: What happens if  $\sum_{k \geq 0} \tau_k = \infty$ ?

- The construction in Def. 1.15 defines a process  $X(t)$  for  $t < \sum_{k=0}^{\infty} \tau_k$ .
- After this time, one has to define how it continues. For instance we can choose an  $u \in S$  and define  $X(\sum_{k=0}^{\infty} \tau_k) = u$  and after time  $u$  run the chain until the next explosion and then return to  $u$  (or to another site).
- As the construction depends on  $u$ , and there are  $\infty$ -many  $u$  (otherwise, for  $S$  finite, there are no explosions)  $\Rightarrow \infty$ -many "versions" of the process.

Rem.: Although generically not enough to determine the process, it is sometimes useful to have the Kolmogorov forward equations.

Thm 1.20) Assume  $\sum_{y \in S} q(z,y) < \infty$  and

$\sum_{y \in S} p_t^*(x,y) < \infty$  for all  $t$ . Then

$$\frac{d}{dt} p_t^*(x,y) = \sum_{z \in S} p_t^*(x,z) q(z,y)$$

(without proof).

L 24.10.2013