On Markov chains and filtrations

Peter Spreij
Department of Econometrics
Vrije Universiteit
De Boelelaan 1105
1081 HV Amsterdam
pspreij@econ.vu.nl

February 14, 1997

Abstract

In this paper we rederive some well known results for continuous time Markov processes that live on a finite state space. Martingale techniques are used throughout the paper. Special attention is paid to the construction of a continuous time Markov process, when we start from a discrete time Markov chain. The Markov property here holds with respect to filtrations that need not be minimal.

1 Introduction

In many classical treatments of the theory of Markov processes it is not always easy to distinguish in lines of thought in a proof of a theorem hard analysis from heuristic reasoning. The purpose of the present paper is to give rigorous proofs –of well known results– based on martingale methods and by paying more attention to the essentials in the structure of the various possible underlying filtrations. Thus we apply results from the theory of multivariate point processes, which serves as a convenient vehicle, and leads to an elegant formulation of issues about Markov chains, which ought to be considered as an alternative to the classical approaches.
The rest of the paper is organised as follows. In section 2 we start with a continuous time process, which is assumed to be Markov with respect to a certain filtration. By application of stochastic calculus one can straightforwardly derive properties of the embedded chain (sampled at the jump times). In section 3–that constitutes the bulk of the paper–we start with discrete time filtrations and stochastic processes and study the construction of a related continuous time process. We allow for a generalization of similar work in Doob [6] in that we consider arbitrary filtrations (that make all the processes involved adapted, but need not be generated by these processes) and that the sampled continuous time process is not necessarily the same as the discrete time process from which it was constructed. We also discuss in detail

the construction of (various) filtrations in continuous time. This is a relative novelty in the present approach. Usually one encounters in the literature (see e.g. Brémaud [3]) only properties, like a characterization of a pre-stopping time σ- algebra, of the filtration generated by a point process, whereas here we put the emphasis on on discrete time objects as building blocks for a continuous time theory.
An application to marked point processes is also presented, which slightly differs from certain well known results. Finally we prove the Markov property in continuous time by means of a measure transformation.
The paper is essentially self-contained and readers are assumed to be familiar with some basic notions from the general theory of stochastic processes only.

2 Markov chains in continuous time

Let $(\Omega, \mathcal{F}, \mathbb{F}, P)$ be a filtered probability space. Assume that the *filtration* $\mathbb{F}$ satisfies the usual conditions in the sense of Dellacherie & Meyer [5]. Let X be a cadlag $\mathbb{F}$-Markov process with a finite state space. Without loss of generality we can assume that the state space is the *standard orthogonal basis* of the Euclidian space $\mathbb{R}^m$. Call this set $B^m = \{b_1, \ldots, b_m\}$. (Indeed, if ξ is a stochastic process with values in a set $\{z_1, \ldots, z_m\}$, where all the z_i are different, then we can define the process X with values in B^m by $X_t = b_i$ iff $\xi_t = z_i$. Hence the probabilistic structure of ξ determines that of X and vice versa). By X is $\mathbb{F}$-Markov it is meant that for all $t \geq s$ and for all $b \in B^m$ one has $P(X_t = b|\mathcal{F}_s) = P(X_t = b|\sigma(X_s))$. Denote by $\Phi(t, s)$ the $m \times m$ matrix with elements $\Phi_{ij}(t, s) = P(X_t = b_i|X_s = b_j)$ and let (the limit is assumed to exist) $A(t) = \lim_{h\downarrow 0} \frac{1}{h}[\Phi(t + h, t) - I]$. In this paper we assume that actually $A(t)$ is independent of t, so we write A instead. We call A the generating matrix of X. Notice that the *column sums* of $\Phi(t, s)$ are all equal to one and that all *the column sums* of A equal zero.
The advantage of working with the state space B^m can be illustrated with the following observation (and also the formulation of all the results below): X is $\mathbb{F}$-Markov if and only if $E[X_t|\mathcal{F}_s] = \Phi(t, s)X_s$ for all $t \geq s$.
Necessity of this equality can be shown as follows. First we note that $P(X_t = b_i|\mathcal{F}_s) = \sum_{j=1}^m P(X_t = b_i|X_s = b_j)1_{\{X_s=b_j\}} = b_i^T\Phi(t, s)X_s$. Hence $E[X_t|\mathcal{F}_s] = \sum_{i=1}^m b_iP(X_t = b_i|\mathcal{F}_s) = \sum_{i=1}^m b_ib_i^T\Phi(t, s)X_s = \Phi(t, s)X_s$, since $\sum_{i=1}^m b_ib_i^T = I$.
In order to show sufficiency we use that $X_t = b_i$ iff $b_i^TX_t = 1$. So we obtain $P(X_t = b_i|\mathcal{F}_s) = E[1_{\{X_t=b_i\}}|\mathcal{F}_s] = b_i^TE[X_t|\mathcal{F}_s] = b_i^T\Phi(t, s)X_s$.

The main result of this section is theorem 2.3 below. Its content can be found in many textbooks on Markov chains for the situation where the filtration $\mathbb{F}$ is generated by X itself (see e.g. Çinlar [4], section 8.3. We provide alternative proofs based on stochastic calculus.
The following result (which can be generalized to the time varying case without the differentiability assumption, see the appendix, proposition A.3) is fundamental for this section and can be found in e.g. [17]. It gives a convenient equivalent representation of an $\mathbb{F}$-Markov process with values in B^m as the solution of a stochastic differential equation.

Proposition 2.1 *A stochastic process $X : \Omega \times [0,\infty) \to B^m$ is $\mathbb{F}$-Markov with generating matrix A iff X satisfies the stochastic differential equation*

$$dX_t = AX_t dt + dM_t,\ X_0 \tag{2.1}$$

with $M : \Omega \times [0,\infty) \to \mathbb{R}^n$ a n-dimensional $\mathbb{F}$-martingale.

The rest of this section is mainly devoted to proving some basic results of the process that is obtained by sampling X at its jump times. The main tool used in our proofs is the stochastic calculus rule.
We need some notation. Define N as the counting process that counts the number of transitions of X. So $N_t = \sum_{s\leq t} 1_{\{X_s \neq X_{s-}\}}$. Let λ be the row vector whose i-th element $\lambda_i = -A_{ii}$. Then we have the following

Proposition 2.2 *The intensity process of N is λX.*

PROOF: Notice first that $X_t^T X_t \equiv 1$. Since X is a process of bounded variation, the optional quadratic variation process $[X^T, X]$ is such that $[X^T, X]_t = \sum_{s\leq t} \Delta X_s^T \Delta X_s$. But $\Delta X_s^T \Delta X_s = 2.1_{\{X_s \neq X_{s-}\}} = \Delta N_s$, so $[X^T, X] = 2N$.
Apply now the product rule for differentials and equation (2.1) to get

$$\begin{aligned} d(X_t^T X_t) &= X_{t-}^T dX_t + dX_t^T X_{t-} + d[X^T, X]_t \\ &= X_t^T A X_t dt + X_{t-}^T dM_t + X_t^T A^T X_t dt + dM_t^T X_{t-} + 2dN_t. \end{aligned}$$

Use now that $d(X_t^T X_t) = 0, X_t^T A X_t = X_t^T A^T X_t = -\lambda X_t$ and write dm_t for the martingale terms in this equation to get $dN_t = \lambda X_t dt + dm_t$. □

Introduce the following notation. For $k \in \{0, 1, 2, \ldots\}$ let $T_k = \inf\{t > 0 : N_t = k\}$, the time of the k-th transition of X and $S_{k+1} = T_{k+1} - T_k$. Let furthermore Λ be the diagonal matrix with elements $\Lambda_{ii} = \lambda_i = -A_{ii}$. Assume that the $\lambda_i > 0$, then Λ is invertible, and the T_k are finite a.s. We have the following well known result.

Theorem 2.3 *(i) For all $k \geq 0$ we have that S_{k+1} has, conditionally on $\mathcal{F}_{T_k}$, an exponential distribution with mean $\frac{1}{\lambda X_{T_k}}$.*
(ii) For all $k \geq 0$ it holds that $E[X_{T_{k+1}} | \mathcal{F}_{T_k}] = [\frac{1}{\lambda X_{T_k}} A + I] X_{T_k} = [A\Lambda^{-1} + I] X_{T_k}$.
(iii) $X_{T_{k+1}}$ and $S_{T_{k+1}}$ are conditionally independent given $\mathcal{F}_{T_k}$.

PROOF: (i) One possibility is to use formula (2.12) on page 63 in Brémaud [3], which actually only refers to the case where $\mathbb{F}$ is generated by X. We give an alternative proof. Let $g_t = \exp(iuN_t + ivt)$. An application of the stochastic calculus rule gives

$$dg_t = g_{t-}((iu\lambda_t + iv)dt + iudm_t) + g_{t-}(e^{iu} - 1 - iu)\Delta N_t$$

So

$$\exp(iuN_{T_{k+1}} + ivT_{k+1}) - \exp(iuN_{T_k} + ivT_k) =$$

$$\int_{(T_k, T_{k+1}]} g_{t-}((iu\lambda_t + iv)dt + iudm_t) + \int_{(T_k, T_{k+1}]} g_{t-}(e^{iu} - 1 - iu)dN_t =$$

$$\int_{(T_k,T_{k+1}]} \exp(iuN_{T_k}+ivt)((iu\lambda X_{T_k}+iv)dt+iudm_t)+$$

$$\exp(iuN_{T_k}+ivT_{k+1})(e^{iu}-1-iu)$$

Since $N_{T_k}=k$, this formula reduces to

$$\exp(iu(k+1)+ivT_{k+1})-\exp(iuk+ivT_k)=$$

$$\int_{(T_k,T_{k+1}]} \exp(iuk+ivt)((iu\lambda X_{T_k}+iv)dt+iudm_t)+$$

$$\exp(iuk+ivT_{k+1})(e^{iu}-1-iu)$$

Take now conditional expectations given $\mathcal{F}_{T_k}$ to get

$$E[\exp(iu(k+1)+ivT_{k+1})-\exp(iuk+ivT_k)|\mathcal{F}_{T_k}]=$$

$$E[\int_{(T_k,T_{k+1}]} \exp(iuk+ivt)(iu\lambda X_{T_k}+iv)dt|\mathcal{F}_{T_k}]+$$

$$E[\exp(iuk+ivT_{k+1})|\mathcal{F}_{T_k}](e^{iu}-1-iu)$$

Since T_k is $\mathcal{F}_{T_k}$-measurable, multiplication by $\exp(-iuk-ivT_k)$ yields

$$E[\exp(iu+ivS_{k+1})|\mathcal{F}_{T_k}]=$$

$$1+(iu\lambda X_{T_k}+iv)E[\int_{(T_k,T_{k+1}]} \exp(iv(t-T_k))|\mathcal{F}_{T_k}]+$$

$$E[\exp(ivS_{k+1})|\mathcal{F}_{T_k}](e^{iu}-1-iu)$$

$$=1+\frac{iu\lambda X_{T_k}+iv}{iv}E[\exp(ivS_{k+1})-1|\mathcal{F}_{T_k}]+E[\exp(ivS_{k+1})|\mathcal{F}_{T_k}](e^{iu}-1-iu)$$

Write now $h(v)=E[\exp(ivS_{k+1})|\mathcal{F}_{T_k}]$. Then we obtain that $h(v)$ satisfies for all u the equation

$$e^{iu}h(v)=1+\frac{u\lambda X_{T_k}+v}{v}[h(v)-1]+h(v)(e^{iu}-1-iu).$$

Solving this equation, we obtain that $h(v)=(1-\frac{iv}{\lambda X_{T_k}})^{-1}$. So $h(v)$ is the characteristic function of the exponential distribution with mean $(\lambda X_{T_k})^{-1}$. This proves the first assertion of the theorem.
(ii) From equation (2.1) we get

$$X_{T_{k+1}}-X_{T_k}=\int_{(T_k,T_{k+1}]} AX_t dt+M_{T_{k+1}}-M_{T_k}$$

$$AX_{T_k}(T_{k+1}-T_k)+M_{T_{k+1}}-M_{T_k}.$$

Take now conditional expectations, use that X_{T_k} is $\mathcal{F}_{T_k}$- measurable, that (see assertion (i) of this theorem) $E[S_{k+1}|\mathcal{F}_{T_k}] = (\lambda X_{T_k})^{-1}$ and that on the events $\{X_{T_k} = b_i\}$ it holds that $\frac{1}{\lambda X_{T_k}} AX_{T_k} = A\Lambda^{-1}X_{T_k}$. Then the result follows.

(iii) We have to prove that for all $u^T = [u_1, \ldots, u_n] \in \mathbb{R}^{1\times n}$ and for all $v \in \mathbb{R}$ the following factorization holds: $E[e^{iu^T X_{T_{k+1}} + ivS_{k+1}}|\mathcal{F}_{T_k}] = E[e^{iu^T X_{T_{k+1}}}|\mathcal{F}_{T_k}]E[e^{ivS_{k+1}}|\mathcal{F}_{T_k}]$. But since $e^{iuX_{T_{k+1}}} = [e^{iu_1}, \ldots, e^{iu_n}]X_{T_{k+1}}$ it is sufficient to prove that the following holds:

$$E[X_{T_{k+1}}e^{ivS_{k+1}}|\mathcal{F}_{T_k}] = E[X_{T_{k+1}}|\mathcal{F}_{T_k}]E[e^{ivS_{k+1}}|\mathcal{F}_{T_k}]. \tag{2.2}$$

Use the product rule and equation (2.1) to get

$$d(X_t e^{ivt}) = (AX_t dt + dM_t)e^{ivt} + X_t iv e^{ivt} dt.$$

From this equation we get by integrating over $(T_k, T_{k+1}]$ and by taking conditional expectations

$$E[X_{T_{k+1}}e^{ivT_{k+1}}|\mathcal{F}_{T_k}] - X_{T_k}e^{ivT_k} =$$

$$(A + ivI)X_{T_k}E[\frac{e^{ivT_{k+1}} - e^{ivT_k}}{iv}|\mathcal{F}_{T_k}]$$

Multiplication by e^{-ivT_k} yields

$$E[X_{T_{k+1}}e^{ivS_{k+1}}|\mathcal{F}_{T_k}] - X_{T_k} = (A + ivI)X_{T_k}E[\frac{e^{ivS_{k+1}} - 1}{iv}|\mathcal{F}_{T_k}] =$$

$$(A + ivI)X_{T_k}\frac{(1 - \frac{iv}{\lambda X_{T_k}})^{-1} - 1}{iv} = (A + ivI)X_{T_k}\frac{1}{\lambda X_{T_k} - iv}.$$

Here we used the expression for $h(v)$ used in the proof of part (i). From this equation we obtain $E[X_{T_{k+1}}e^{ivS_{k+1}}|\mathcal{F}_{T_k}] = \frac{1}{\lambda X_{T_k} - iv}(A + \lambda X_{T_k}I)X_{T_k}$. The factorization (2.2) now follows from assertions (i) and (ii). $\square$

REMARK: The assertions of theorem 2.3 can be modified to take into account the situation where X has absorbing states, in which case some of the $A_{ii} = -\lambda_i$ are zero. This leads to degenerate exponential distributions for the S_{k+1}. Incorporation of such a situation leads to some subtle changes in the proofs, since it may happen that the T_k become infinite. For instance assertion (ii) of theorem 2.3 should be replaced by $E[X_{T_{k+1}}1_{\{T_{k+1}<\infty\}}|\mathcal{F}_{T_k}] = (A + \Lambda)\Lambda^+ X_{T_k}1_{\{T_k<\infty\}}$, where Λ^+ stands for the Moore-Penrose inverse of Λ. We omit a detailed treatment.

Corollary 2.4 *The embedded process $x : \Omega \times \{0, 1, 2, \ldots\} \to B^m$, defined by $x_k = X_{T_k}$, is a Markov chain w.r.t. the discrete time filtration $\{\mathcal{G}_n\}_{n\geq 0}$, defined by $\mathcal{G}_n = \mathcal{F}_{T_n}$, and has transition matrix $\tilde{A}$ given by $\tilde{A} = A\Lambda^{-1} + I$.*

PROOF: We show that the conditional probability $P(x_{k+1} = b_i|\mathcal{G}_k)$ only depends on the conditioning σ-algebra $\mathcal{G}_k$ through x_k: $P(x_{k+1} = b_i|\mathcal{G}_k) = E[1_{\{x_{k+1}=b_i\}}|\mathcal{G}_k] = E[b_i^T x_{k+1}|\mathcal{G}_k] = b_i^T E[X_{T_{k+1}}|\mathcal{F}_{T_k}] = b_i(A\Lambda^{-1}+I)X_{T_k} = b_i(A\Lambda^{-1}+I)x_k$, where we used (i) of theorem 2.3. From this relation it also follows that $\tilde{A}$ is the transition matrix of x. □

REMARK 1: It follows from theorem 2.3 (i) that the process Δm, defined by $\Delta m_k = x_k - \tilde{A}x_{k-1}$ is a martingale difference sequence w.r.t. the filtration $\{\mathcal{G}_n\}_{n\geq 0}$. The proof of corollary 2.4 shows that this property is sufficient for x with state space B^m to be Markov. But this property is also sufficient. If x is Markov with transition matrix $\tilde{A}$, then $E[\Delta m_{k+1}|\mathcal{G}_k] = E[x_{k+1}|\mathcal{G}_k] - \tilde{A}x_k = E[x_{k+1}|\sigma(x_k)] - \tilde{A}x_k = \sum_{i=1}^m E[b_i 1_{\{x_k=b_i\}}|\sigma(x_k)] - \tilde{A}x_k = \sum_{i=1}^m b_i P(x_k = b_i|\sigma(x_k)) - \tilde{A}x_k = \sum_{i=1}^m b_i b_i^T \tilde{A}x_k - \tilde{A}x_k = 0$, because $\sum_{i=1}^m b_i b_i^T = I$.
In this remark we actually proved the discrete time analog of proposition 2.1. Again we refer to the appendix (proposition A.3) for a more general result.

REMARK 2: Observe that for an embedded Markov chain x necessarily the $\tilde{A}_{ii} = 0$, and hence $P(x_{k+1} = x_k) = 0$ for all k.

REMARK 3: Trivially we have

$$\forall n,t : \mathcal{F}_{T_n} \cap \{T_n \leq t < T_{n+1}\} \subset \mathcal{F}_t \cap \{T_n \leq t < T_{n+1}\} \tag{2.3}$$

Consider the following problem. Find the smallest filtration, $\{\mathcal{F}_t^0\}$ say, such that for given $\{\mathcal{F}_{T_n}\}$ the following inclusion holds

$$\forall n,t : \mathcal{F}_{T_n} \cap \{T_n \leq t < T_{n+1}\} \subset \mathcal{F}_t^0 \cap \{T_n \leq t < T_{n+1}\}. \tag{2.4}$$

Then for instance also the following question arises: Is X adapted to $\{\mathcal{F}_t^0 : t \geq 0\}$? This problem can be answered by the results of the next section from which it also follows that for the resulting filtration actually equality holds and that X is an adapted process w.r.t. this filtration.
In general the inclusion (2.3) is strict, which follows e.g. from a result in a recent paper by Jacod and Skorohod [14] where it is shown that equality in (2.3) is equivalent to the assertion that all (adapted) martingales are a.s. of finite variation. For the filtration with which we started this section this does not necessarily hold.

3 From discrete to continuous time

3.1 Introduction

In the previous section we obtained in theorem 2.3 and in corollary 2.4 the distribution of the embedded chain and the distribution of the jump times of the Markov chain. The purpose of the present section is to follow the road in the opposite direction. That is, starting from a Markov chain in discrete time (w.r.t. to some filtration) and a sequence of conditionally exponentially distributed random variables, we construct

a Markov chain in continuous time. One of the questions we address is to what filtration the Markov property here refers.
A similar construction based on another approach can be found in Doob [6], section VI.1 and Gihman & Skorohod [7] sections III.1 and III.3, with the restrictions that the filtrations are generated by the processes involved. Also an elementary treatment can be found in chapter 8 of Çinlar [4]. Here we allow more general filtrations. In Jacod [12], section III.2b properties of filtrations like the one that is introduced below are described, for the case where these are generated by a multivariate point process. We also mention the paper [1] by Boel, Varaiya and Wong for results on jump processes, [9], [10], [15] for some results on filtrations similar to the one that will be discussed in the next subsection and Jacobsen [11] for some results on Markov chains. As a consequence there is some duplication with the existing literature (although we treat in some sense a more general situation), but we prefer to give full proofs to make this paper self contained and also because the methods we use are different.
The basic assumptions are the following. Let $(\Omega, \mathcal{F}, \mathbb{G}, P)$ be a filtered probability space. $\mathbb{G}$ is a filtration in discrete time, $\mathbb{G} = \{\mathcal{G}_k\}_{k\in\{0,1,2,\ldots\}}$. Furthermore we assume that we have a stochastic process $x : \Omega \times \{0, 1, 2, \ldots\} \to B^m$, which is assumed to be $\mathbb{G}$-Markov with generating matrix A', that does not necessarily have the property that all its diagonal elements are zero. (see also at the end of this section). As observed in remark 1 of the previous section, the Markov property in this case is equivalent with saying that the process m, defined by $\Delta m_k = x_k - A'x_{k-1}$, is a $\mathbb{G}$-martingale.
Let Λ be a $m \times m$ diagonal matrix with all its entries on the diagonal $\Lambda_{ii} = \lambda_i$ positive. Define $\lambda = [\lambda_1, \ldots, \lambda_m]$. Assume that we also have a sequence of random variables S_k, that are $\mathcal{G}_k$-measurable, S_{k+1} has, given $\mathcal{G}_k$, an exponential distribution with mean $(\lambda x_k)^{-1}$. This entails that the conditional density (on $(0, \infty)$) of S_{k+1} given $\mathcal{G}_k$ takes the form

$$\mathbf{1}^T \Lambda \exp(-\Lambda s) x_k \tag{3.1}$$

(Here $\mathbf{1}^T$ is the vector $[1, \ldots, 1] \in \mathbb{R}^{1\times m}$).
Furthermore x_{k+1} and S_{k+1} are assumed to be conditionally independent given $\mathcal{G}_k$. (Cf. the assertions of theorem 2.3). Write $T_k = \sum_{l=0}^{k} S_l$.
The main purpose of this section is to show that the continuous time process $X : \Omega \times [0, \infty) \to B^m$, defined by

$$X_t = \sum_{k=0}^{\infty} x_k 1_{\{T_k \le t < T_{k+1}\}} \tag{3.2}$$

has the Markov property w.r.t. the filtration $\mathbb{H}$, that will be defined in the next subsection. The basic assumptions will be introduced step by step, at the stages where they are needed, while all the assumptions made before remain valid.

3.2 The filtration

Consider the filtered probability space $(\Omega, \mathcal{F}, \mathbb{G}, P)$, where $\mathbb{G}$ is a filtration in discrete time, $\mathbb{G} = \{\mathcal{G}_k\}_{k\in\{0,1,2,\ldots\}}$. Denote by $\mathcal{G}_\infty$ the σ-algebra $\bigvee_{n\ge 0} \mathcal{G}_n$. Let $T_n : \Omega \to [0, \infty]$ for each $n \in \{0, 1, 2, \ldots\}$ be a random variable. Assume moreover that for all n

$T_{n+1} \geq T_n$ and that strict inequality holds if $T_n < \infty$ and that the T_n are $\mathcal{G}_n$-measurable. Then a filtration in continuous time is defined in

Definition 3.1 *Let for each $t \in [0, \infty)$ the set $\mathcal{H}_t$ be defined as follows: $\mathcal{H}_t = \{F \in \mathcal{F} : \forall k : \exists G_k \in \mathcal{G}_k$ such that $F \cap \{T_{k+1} > t\} = G_k \cap \{T_{k+1} > t\}\}$.*

Then we have the following (like in Jacod [12], proposition (3.39))

Proposition 3.2 *The collection $\mathbb{H} = \{\mathcal{H}_t\}_{t \in [0,\infty)}$ is a right continuous filtration on Ω.*

PROOF: First we will show that each $\mathcal{H}_t$ is a σ-algebra. Clearly, both $\emptyset$ and Ω belong to $\mathcal{H}_t$. Let $A_1, A_2, \ldots$ belong to $\mathcal{H}_t$. Fix k, then there are sets $A_{ik} \in \mathcal{G}_k$ such that $A_i \cap \{T_{k+1} > t\} = A_{ik} \cap \{T_{k+1} > t\}$. But then $\bigcup_i A_i \cap \{T_{k+1} > t\} = \bigcup_i (A_i \cap \{T_{k+1} > t\}) = \bigcup_i (A_{ik} \cap \{T_{k+1} > t\}) = \bigcup_i A_{ik} \cap \{T_{k+1} > t\}$, and $\bigcup_i A_{ik} \in \mathcal{G}_k$.
The same way of reasoning gives that also countable intersections of elements of $\mathcal{H}_t$ again belong to $\mathcal{H}_t$.
Let now $A \in \mathcal{H}_t$ and $A_k \in \mathcal{G}_k$ such that $A \cap \{T_{k+1} > t\} = A_k \cap \{T_{k+1} > t\}$. Then $A^c \cap \{T_{k+1} > t\} = \{T_{k+1} > t\} \setminus (A \cap \{T_{k+1} > t\}) = \{T_{k+1} > t\} \setminus (A_k \cap \{T_{k+1} > t\}) = A_k^c \cap \{T_{k+1} > t\}$. Hence $A^c \in \mathcal{H}_t$, because $A_k^c \in \mathcal{G}_k$. So we have proved that $\mathcal{H}_t$ is a σ-algebra.
The proof that $\mathbb{H}$ is an increasing family follows from the fact that for $s < t$ it holds that $\{T_{k+1} > t\} \cap \{T_{k+1} > s\} = \{T_{k+1} > t\}$. Finally we have to show that $\mathbb{H}$ is right continuous. This follows from lemma L29 on page 306 in Brémaud [3]. We give this result in the appendix (lemma A.1).
Let $H \in \mathcal{H}_{t+}$. Then $H \in \mathcal{H}_{t+\varepsilon}$ for all $\varepsilon > 0$. Hence for all k we have $H \cap \{T_{k+1} > t + \varepsilon\} \in \{G \cap \{T_{k+1} > t + \varepsilon\} : G \in \mathcal{G}_k\}$. Apply now lemma A.1 to obtain that $H \cap \{T_{k+1} > t\} \in \{G \cap \{T_{k+1} > t\} : G \in \mathcal{G}_k\}$, which means that $H \in \mathcal{H}_t$. □

REMARK: Notice that in the proof of the previous proposition we have not used the fact that the T_n are $\mathcal{G}_n$- measurable.

Another characterization of the filtration $\mathbb{H}$ is given in the next

Proposition 3.3 *Let a family $\{\mathcal{K}_t : t \geq 0\}$ of subsets of Ω be defined by $\mathcal{K}_t = \{F \in \mathcal{F} : \forall k : \exists G_k \in \mathcal{G}_k$ such that $K \cap \{T_{k+1} > t \geq T_k\} = G_k \cap \{T_{k+1} > t \geq T_k\}\}$. Then each $\mathcal{K}_t$ is a σ-algebra. Moreover, $\mathcal{K}_t = \mathcal{H}_t$.*

PROOF: The fact that the $\mathcal{K}_t$ are σ-algebras can be proved in the same way as in the proof of proposition 3.2. So we only need to prove that $\mathcal{K}_t = \mathcal{H}_t$ for each t, from which it also follows that the $\mathcal{K}_t$ are σ-algebras, of course.
Let $H \in \mathcal{H}_t$. Then it follows immediately from the definition of $\mathcal{H}_t$ that $H \cap \{T_{k+1} > t \geq T_k\} = G_k \cap \{T_{k+1} > t \geq T_k\}$ by taking intersections with $\{T_k \leq t\}$. This shows that $\mathcal{H}_t \subset \mathcal{K}_t$. In order to prove the converse, we use an induction argument. Let $K \in \mathcal{K}_t$ and let $k = 0$. Then $K \cap \{T_1 > t\} = K \cap \{T_0 \leq t < T_1\} = G_0 \cap \{T_0 \leq t < T_1\} = G_0 \cap \{\leq T_1 > t\}$ by definition of $\mathcal{K}_t$. Suppose now that for a certain $k \geq 1$ we know that there is $G'_{k-1} \in \mathcal{G}_{k-1}$ such that $K \cap \{t < T_k\} = G'_{k-1} \cap \{t < T_k\}$. By

definition of $\mathcal{K}_t$ there also exists $G'_k \in \mathcal{G}_k$ such that $K \cap \{T_k \leq t < T_{k+1}\} = G'_k \cap \{T_k \leq t < T_{k+1}\}$. But then

$$K \cap \{T_{k+1} > t\} = K \cap (\{T_k > t\} \cup \{T_k \leq t < T_{k+1}\}) =$$

$$(K \cap (\{T_k > t\}) \cup (K \cap \{T_k \leq t < T_{k+1}\}) =$$

$$(G'_{k-1} \cap \{T_k > t\}) \cup (G'_k \cap \{T_k \leq t < T_{k+1}\}) =$$

$$(G'_{k-1} \cap \{T_k > t\} \cap \{T_{k+1} > t\}) \cup (G'_k \cap \{T_k \leq t < T_{k+1}\}) =$$

$$\{T_{k+1} > t\} \cap ((G'_{k-1} \cap \{T_k > t\}) \cup (G'_k \cap \{T_k \leq t\})) =$$

$$\{T_{k+1} > t\} \cap G_k,$$

with $G_k = (G'_{k-1} \cap \{T_k > t\}) \cup (G'_k \cap \{T_k \leq t\}) \in \mathcal{G}_k$. Hence $K \in \mathcal{H}_t$. □

REMARK: One can also define filtrations by $\mathcal{H}'_t = \mathcal{H}_t \cap \mathcal{G}_\infty$ and $\mathcal{K}'_t = \mathcal{K}_t \cap \mathcal{G}_\infty$. Then it follows that $\mathcal{K}'_t = \mathcal{H}'_t$ and moreover $\mathcal{K}'_t = \{K \in \mathcal{K}_t : K \cap \{t \geq T_\infty\} \in \mathcal{G}_\infty \cap \{t \geq T_\infty\}\}$. This corresponds to the filtration defined in Jacod [12] on page 84. Of course the $\mathcal{H}_t$ and $\mathcal{H}'_t$ coincide if $\mathcal{F} = \mathcal{G}_\infty$.

Proposition 3.4 *The T_n are $\mathbb{H}$-stopping times.*

PROOF: We have to show that for all n and for all t the set $\{T_n \leq t\} \in \mathcal{H}_t$. For $k \leq n-1$ we have $\{T_n \leq t\} \cap \{T_{k+1} > t\} = \emptyset \cap \{T_{k+1} > t\}$, whereas for $k \geq n$ the set $\{T_n \leq t\} \in \mathcal{G}_n \subset \mathcal{G}_k$. Hence for all k there is $A_k \in \mathcal{G}_k$ such that $\{T_n \leq t\} \cap \{T_{k+1} > t\} = A_k \cap \{T_{k+1} > t\}$. □

Proposition 3.5 *For all n one has $\mathcal{H}_{T_n} = \mathcal{G}_n \cap \mathcal{H}_\infty$, where $\mathcal{H}_\infty = \vee_{t \geq 0} \mathcal{H}_t$. and $\mathcal{H}_{T_{n+1}-} = \mathcal{H}_{T_n} \vee \sigma(T_{n+1})$.*

PROOF: Let $G \in \mathcal{G}_n \cap \mathcal{H}_\infty$. In order to prove that $G \in \mathcal{H}_{T_n}$, one has to show that for all t the set $G \cap \{T_n \leq t\}$ belongs to $\mathcal{H}_t$. Consider thereto $G \cap \{T_n \leq t\} \cap \{T_{k+1} > t\}$. For $k \leq n-1$ this set is empty. For $k \geq n$ one has $\{T_n \leq t\} \in \mathcal{G}_n \subset \mathcal{G}_k$. Hence $G \cap \{T_n \leq t\} \in \mathcal{G}_k$. It follows that $G \in \mathcal{H}_{T_n}$.
We proceed to prove the converse. Let $H \in \mathcal{H}_{T_n}$. First we notice that $H \cap \{T_n = \infty\} \in \mathcal{H}_{T_n-}$, by T11 in [3], page 299. It is immediately clear from the definition of $\mathcal{H}_{T_n-}$ that this σ-algebra is a subset of $\mathcal{G}_n$, since $\mathcal{H}_{T_n-} = \sigma\{H \cap \{T_n > t\} : H \in \mathcal{H}_t, t \geq 0\} \subset \sigma\{G \cap \{T_n > t\} : G \in \mathcal{G}_{n-1}, t \geq 0\} = \mathcal{G}_{n-1} \vee \sigma(T_n) \subset \mathcal{G}_n$, by $\mathcal{G}_n$-measurability of T_n.
So it is sufficient to prove that $H \cap \{T_n < \infty\} \in \mathcal{G}_n$. First we observe that this set also belongs to $\mathcal{H}_{T_n}$. Hence for all t there is a set G_t in $\mathcal{G}_n$ such that $H \cap \{T_n < \infty\} \cap \{T_n \leq t < T_{n+1}\} = G_t \cap \{T_{n+1} > t\}$. Denote by Z the optional process defined by $Z_t = 1_{\{T_n \leq t\}}$. Then $Z_t 1_{\{T_n \leq t < T_{n+1}\}} = Y^n_t 1_{\{T_n \leq t < T_{n+1}\}}$, where the process Y^n is $\mathcal{G}_n$- measurable and rightcontinuous (Take $Y^n_t = \liminf_{h \downarrow 0} 1_{G^n_{t+h}}$). Similarly, we can consider the optional process Z^H given by $Z^H_t = 1_{\{T^H_n \leq t\}}$ (Recall that for an $\mathbb{H}$-stopping time T and a set $H \in \mathcal{H}_T$ the random variable $T^H := T.1_H + \infty.1_{H^c}$ is

an $\mathbb{H}$- stopping time too). And then there is a $\mathcal{G}_n$-measurable process $Y^{n,H}$ such that $Z^H_t 1_{\{T_n \leq t < T_{n+1}\}} = Y^{n,H}_t 1_{\{T_n \leq t < T_{n+1}\}}$. But $Z^H_t = 1_H Z_t$. So, on $\{T_n \leq t < T_{n+1}\}$ we have $Y_{n,H} = 1_H Y^n$. Next $1_{\{T_n<\infty\}} = 1_{\{T_n<\infty\}} Z_{T_n} = 1_{\{T_n<\infty\}} \sum_{k\geq 0} Y^k_{T_n} 1_{\{T_k \leq T_n < T_{k+1}\}} = 1_{\{T_n<\infty\}} Y^n_{T_n}$. Hence $1_H 1_{\{T_n<\infty\}} = 1_{\{T_n<\infty\}} Z^H_{T_n} = 1_{\{T_n<\infty\}} Y^{n,H}_{T_n}$, which is $\mathcal{G}_n$- measurable. So we proved that $H \cap \{T_n < \infty\} \in \mathcal{G}_n$, which is what we needed.
Consider now $\mathcal{H}_{T_{n+1}-}$. It is generated by the sets of the form $H_t \cap \{T_{n+1} > t\}$, with $H_t \in \mathcal{H}_t$. By definition of $\mathcal{H}_t$ such a set belongs to $\mathcal{G}_n \cap \{T_{n+1} > t\} \cap \mathcal{H}_\infty$. Hence it follows that $\mathcal{H}_{T_{n+1}-} \subset \mathcal{H}_{T_n} \vee \sigma(T_{(n+1)})$. Since T_{n+1} is $\mathcal{H}_{T_{n+1}-}$-measurable and $\mathcal{G}_n \cap \mathcal{H}_\infty = \mathcal{H}_{T_n}$ we conclude from the first part of the proof that actually equality holds here: $\mathcal{H}_{T_{n+1}-} \subset \mathcal{G}_n \cap \mathcal{H}_\infty \vee \sigma(T_{n+1}) = \mathcal{H}_{T_n} \vee \sigma(T_{n+1}) \subset \mathcal{H}_{T_{n+1}-} \vee \sigma(T_{n+1}) = \mathcal{H}_{T_{n+1}-}$. □

REMARK: It is in general not true that $\mathcal{H}_{T_n} = \mathcal{G}_n$. Consider for instance the (somewhat degenerate) case where $T_0 = 0, T_n = \infty$ for $n \geq 1$. Then it is easy to check that $\mathcal{H}_t = \mathcal{G}_0$ for all t and hence $\mathcal{H}_\infty = \mathcal{G}_0$. It also follows that $\mathcal{H}_{T_n} = \mathcal{H}_\infty = \mathcal{G}_0$.
However, if all the T_n are finite, then $\mathcal{G}_\infty \subset \mathcal{H}_\infty$ (and consequently $\mathcal{H}_{T_n} = \mathcal{G}_n$). Indeed, let $G \in \mathcal{G}_n$. Then $G = \cup_{t\geq 0} H_t$, with $H_t = G \cap \{T_n \leq t\}$. It follows from the first part of the proof of proposition 3.5 that $H_t \in \mathcal{H}_t$, hence $G \in \mathcal{H}_\infty$ and consequently $\mathcal{G}_n \subset \mathcal{H}_\infty$ and $\mathcal{G}_\infty \subset \mathcal{H}_\infty$.
The opposite inclusion $\mathcal{H}_\infty \subset \mathcal{G}_\infty$ is (even in the case of finite T_n) in general not true. Consider as an example the deterministic $T_n = \frac{n}{n+1}$. (So $T_\infty = 1 < \infty$). Then it is not difficult to check that $\mathcal{H}_t = \mathcal{G}_{[\frac{t}{1-t}]}$ for $t < 1$, and $\mathcal{H}_t = \mathcal{F}$ for $t \geq 1$. Consequently $\mathcal{H}_\infty = \mathcal{F}$, which is in general bigger than $\mathcal{G}_\infty$. However, under the assumption that $T_\infty = \infty$ it follows that $\mathcal{H}_\infty \subset \mathcal{G}_\infty$, because $\mathcal{H}_t \cap \{T_n > t\} \subset \mathcal{G}_\infty \cap \{T_n > t\}$.

REMARK: We have seen above that in general $\mathcal{H}_{T_n}$ is a genuine subset of $\mathcal{G}_n$. So one can also construct a filtration $\bar{\mathbb{H}}$ by letting the $\mathcal{H}_{T_n}$ take the place of the $\mathcal{G}_n$. It trivially follows that $\bar{\mathcal{H}}_t \subset \mathcal{H}_t$, but also the opposite inclusion is true, because if $H \in \mathcal{H}_t$, then by definition of $\mathcal{H}_t$ it holds that $H \cap \{T_{n+1} > t\} \in \mathcal{G}_n \cap \mathcal{H}_\infty = \mathcal{H}_{T_n}$ and so $H \cap \{T_{n+1} > t\} \in \mathcal{H}_{T_n} \cap \{T_{n+1} > t\}$. So $\bar{\mathbb{H}} = \mathbb{H}$.

Assume furthermore that there is a sequence of random variables x_n, $n \in \bar{\mathbb{N}}$ which take their values in some other measurable space, and that x_n is $\mathcal{G}_n$-measurable for each $n \in \mathbb{N}$, x_∞ is $\mathcal{F}$-measurable. Define

$$X_t = \sum_{n=0}^{\infty} x_n 1_{\{T_n \leq t < T_{n+1}\}} + x_\infty 1_{\{t \geq T_\infty\}} \tag{3.3}$$

Proposition 3.6 *The process X defined by equation* (3.3) *is $\mathbb{H}$-adapted.*

PROOF: Let $B \in B^m$. We have to show that for all t the set $\{X_t \in B\}$ belongs to $\mathcal{H}_t$. Introduce the sets $A_k = \bigcup_{n=0}^{k-1} (\{x_n \in B\} \cap \{T_n \leq t < T_{n+1}\})$ and $B_k = \{x_k \in B\} \cap \{T_k \leq t\}$. Notice that both these sets belong to $\mathcal{G}_k$.
Consider now

$$\{X_t \in B\} \cap \{T_{k+1} > t\} =$$

$$\bigcup_{n=0}^{\infty} (\{X_t \in B\} \cap \{T_n \leq t < T_{n+1}\} \cap \{T_{k+1} > t\}) \cup$$

$$(\{X_t \in B\} \cap \{T_\infty \le t\} \cap \{T_{k+1} > t\})$$

$$\bigcup_{n=0}^{k} (\{x_n \in B\} \cap \{T_n \le t < T_{n+1}\} \cap \{T_{k+1} > t\}) =$$

$$(A_k \cap \{T_{k+1} > t\}) \cup (B_k \cap \{T_{k+1} > t\}) =$$

$$(A_k \cup B_k) \cap \{T_{k+1} > t\}.$$

Since $A_k \cup B_k \in \mathcal{G}_k$, the proof is complete. □

REMARK: Now we return to the question posed in the last remark of the previous section. The answer is given by taking $\mathcal{G}_n = \mathcal{F}_{T_n}$ and the minimal filtration one looks for is nothing else but the resulting $\mathbb{H}$, for which equality in (2.4) holds and to which X is adapted.

3.3 Multivariate point processes

In this subsection we discuss an application of the obtained results to multivariate point processes. Consider next to the T_n sequence a sequence of random variables Z_n^*, taking values in some auxiliary measurable space $(E, \mathcal{E})$. Define $Z_n = Z_n^* 1_{\{T_n < \infty\}}$, assuming that the product makes sense in E. Define then the $\mathcal{G}_n$ as $\sigma(Z_0, Z_1, T_1, \ldots, Z_n, T_n)$. Let now $x_n^* = (T_n, Z_n^*)$, $x_n = (T_n, Z_n)$ and X_t is defined for $t \in [0, \infty)$ by $X_t = \sum_{n=0}^\infty x_n 1_{\{T_n \le t < T_{n+1}\}} = \sum_{n=0}^\infty x_n^* 1_{\{T_n \le t < T_{n+1}\}}$. Then X can be considered as a multivariate point process with the Z_n-sequence as marks. Following the usual convention all the events of X take place before T_∞. We claim the following:

Proposition 3.7 *The filtration $\mathbb{H}$, as defined in the previous subsection with the $\mathcal{G}_n$ as above, is identical with $\mathbb{F}^X$, the filtration generated by X.*

PROOF: Since X is $\mathbb{H}$-adapted, one trivially has $\mathcal{F}_t^X \subset \mathcal{H}_t$. The T_n are the jump times of X and hence stopping times for $\mathbb{F}^X$. Consequently $\mathcal{F}_{T_n}^X \subset \mathcal{H}_{T_n}$. But $x_n 1_{\{T_n < \infty\}} = (T_n 1_{\{T_n < \infty\}}, Z_n) = X_{T_n} 1_{\{T_n < \infty\}} \in \mathcal{F}_{T_n}^X$. Hence we obtain the following chain of inclusions: $\mathcal{G}_n \subset \mathcal{F}_{T_n}^X \subset \mathcal{H}_{T_n} = \mathcal{H}_\infty \cap \mathcal{G}_n \subset \mathcal{G}_n$. It follows that $\mathcal{F}_{T_n}^X = \mathcal{G}_n$. Using this identity, one easily verifies that the traces of $\mathcal{H}_t$ and $\mathcal{F}_t^X$ coincide on the sets $\{T_n \le t < T_{n+1}\}$. Let now $H \in \mathcal{H}_t$. Then there exist sets $F_{nt} \in \mathcal{F}_t^X$ such that $H \cap \{t < T_\infty\} = \bigcup_{n=0}^\infty (H \cap \{T_n \le t < T_{n+1}\}) = \bigcup_{n=0}^\infty (F_{nt} \cap \{T_n \le t < T_{n+1}\}) \in \mathcal{F}_t^X$, by the fact that the T_n are $\mathbb{F}^X$-stopping times.
Furthermore we know that $\mathcal{F}_{T_\infty}^X = \bigvee_{n=0}^\infty \mathcal{F}_{T_n}^X = \bigvee_{n=0}^\infty \mathcal{G}_n = \mathcal{G}_\infty \supset \mathcal{H}_\infty \supset \mathcal{H}_t$. So $H \in \mathcal{F}_{T_\infty}^X$, from which it follows that $H \cap \{t \ge T_\infty\} \in \mathcal{F}_t^X$. So we proved that $\mathcal{H}_t \subset \mathcal{F}_t^X$, and hence $\mathcal{H}_t = \mathcal{F}_t^X$. □

REMARK: Notice that there is a little difference with for instance T30 in Brémaud [3], page 307, where (in our notation) $\mathcal{F}_{T_n}^X = \sigma\{Z_0^*, T_1, Z_1^*, \ldots, T_n, Z_n^*\}$. We cannot have this result here, since if for instance $T_1 = \infty$, then for all $n \ge 1$ one has $\mathcal{F}_{T_n}^X = \mathcal{G}_n = \sigma\{Z_0^*\} \ne \sigma\{Z_0^*, T_1, Z_1^*, \ldots, T_n, Z_n^*\}$. Of course the difference disappears if all the T_n are finite.

Finally we observe that the constructions in this section allow for a generalization of the notion of a multivariate or marked point process as a sequence of pairs of random times and σ-algebras $\{(T_n, \mathcal{G}_n)\}$, where the T_n and the $\mathcal{G}_n$ satisfy the assumptions of the previous subsection.

3.4 The Markov property

In addition to the assumptions made in the previous two subsections we impose the following conditions on the random variables x_n and T_n. Each x_n assumes its values in the set B^m (see the introduction) and the sequence $\{x_n\}$ is Markov w.r.t. the filtration $\mathbb{G}$. Denote by A' the matrix of transition probabilities of x, so $P(x_{n+1} = b_i | x_n = b_j) = A'_{ij}$. Unlike Jacobsen [11] we do not assume that the A'_{ii} are zero, which is a necessary property of an embedded Markov chain (see the previous section).

Define $S_{n+1} = T_{n+1} - T_n$ and assume that S_{n+1} has, conditionally on $\mathcal{G}_n$, an exponential distribution with density on $(0, \infty)$

$$\mathbf{1}^T \Lambda \exp(-\Lambda s) x_n,$$

where Λ is a diagonal matrix with entries $\Lambda_{ii} = \lambda_i$ and all $\lambda_i \geq 0$. Furthermore S_{n+1} and x_{n+1} are assumed to be conditionally independent given $\mathcal{G}_n$.

The main result of this subsection is that the process X is a continuous time Markov process w.r.t. the filtration $\mathbb{H}$. In order to show that we use the following two lemmas. The first of these tells us how to compute certain conditional expectations given $\mathcal{H}_t$ in terms of conditional expectations given $\mathcal{G}_n$.

Lemma 3.8 *Let Z be an integrable random variable. Then, with the convention $\frac{0}{0} = 0$,*

$$E[Z 1_{\{T_n \leq t < T_{n+1}\}} | \mathcal{H}_t] = 1_{\{T_n \leq t < T_{n+1}\}} \frac{E[Z 1_{\{T_n \leq t < T_{n+1}\}} | \mathcal{G}_n]}{E[1_{\{T_n \leq t < T_{n+1}\}} | \mathcal{G}_n]}, \tag{3.4}$$

with

$$E[1_{\{T_n \leq t < T_{n+1}\}} | \mathcal{G}_n] = 1_{\{T_n \leq t\}} x_n^T \exp(-\Lambda(t - T_n)) \mathbf{1}. \tag{3.5}$$

PROOF: Introduce another measure Q, which is absolutely continuous with respect to P by its Radon-Nikodym derivative

$$\frac{dQ}{dP} = \mu = \frac{1_n(t)}{P(\{T_n \leq t < T_{n+1}\})},$$

where $1_n(t) = 1_{\{T_n \leq t < T_{n+1}\}}$. Denote (conditional) expectations under Q by E_Q. Then we have by the Kallianpur-Striebel formula (see equation (3.5') on page 171 of

Brémaud [3] or lemma A.5 in the appendix) for the conditional expectation P and Q a.s. on $\{T_n \leq t < T_{n+1}\}$

$$E_Q[Z1_n(t)|\mathcal{G}_n] = \frac{E[Z1_n(t)\mu|\mathcal{G}_n]}{E[\mu|\mathcal{G}_n]} = \frac{E[Z1_n(t)|\mathcal{G}_n]}{E[1_n(t)|\mathcal{G}_n]}. \tag{3.6}$$

Let $H \in \mathcal{H}_t$. Then there is a $G \in \mathcal{G}_n$ such that $H \cap \{T_n \leq t < T_{n+1}\} = G \cap \{T_n \leq t < T_{n+1}\}$. Abbreviate $1_{\{T_n \leq t < T_{n+1}\}}$ by $1_n(t)$ and consider the following string of equalities:

$$\begin{aligned}
&\int_H E[Z1_n(t)|\mathcal{H}_t]dP = \\
&\int_H Z1_n(t)dP = \\
&\int_{H\cap\{T_n \leq t < T_{n+1}\}} ZdP = \\
&\int_{G\cap\{T_n \leq t < T_{n+1}\}} ZdP = \\
&P(\{T_n \leq t < T_{n+1}\}) \int_{G\cap\{T_n \leq t < T_{n+1}\}} ZdQ = \\
&P(\{T_n \leq t < T_{n+1}\}) \int_G 1_n(t)ZdQ = \\
&P(\{T_n \leq t < T_{n+1}\}) \int_G E_Q[1_n(t)Z|\mathcal{G}_n]dQ = \\
&P(\{T_n \leq t < T_{n+1}\}) \int_{G\cap\{T_n \leq t < T_{n+1}\}} E_Q[1_n(t)Z|\mathcal{G}_n]dQ = \\
&\int_{G\cap\{T_n \leq t < T_{n+1}\}} E_Q[1_n(t)Z|\mathcal{G}_n]dP = \\
&\int_{H\cap\{T_n \leq t < T_{n+1}\}} E_Q[1_n(t)Z|\mathcal{G}_n]dP = \\
&\int_H 1_n(t)E_Q[1_n(t)Z|\mathcal{G}_n]dP.
\end{aligned}$$

By construction of $\mathcal{H}_t$ the last integrand is $\mathcal{H}_t$-measurable. Equation (3.4) now follows from equation (3.6). Finally equation (3.5) results from the assumption that S_{n+1} has an exponential distribution given $\mathcal{G}_n$. □

Lemma 3.9 *The following equation holds for $n, k \geq 0$, $t \geq s \geq 0$:*

$$E[x_{n+k}1_{n+k}(t)1_n(s)|\mathcal{H}_s] = p_k(t-s)x_n1_n(s), \tag{3.7}$$

where the matrix valued functions p_k are recursively defined by

$$p_0(t) = e^{-\Lambda t} \tag{3.8}$$

$$p_k(t) = \int_0^t p_{k-1}(u)A'\Lambda e^{\Lambda(u-t)}du, \tag{3.9}$$

for $k \geq 1$.

PROOF: Consider first the case where $k = 0$. Use lemma 3.8 to write

$$E[x_n 1_n(t)1_n(s)|\mathcal{H}_s] = 1_n(s)E[x_n 1_n(t)1_n(s)|\mathcal{G}_n]x_n^T e^{\Lambda(s-T_n)}\mathbf{1} =$$

$$1_n(s)1_{\{T_n\leq s\}}E[1_{\{T_{n+1}>t\}}|\mathcal{G}_n]x_n x_n^T e^{\Lambda(s-T_n)}\mathbf{1} =$$

$$1_n(s)E[1_{\{T_{n+1}>t\}}|\mathcal{G}_n]e^{\Lambda(s-T_n)}x_n =$$

$$1_n(s)\mathbf{1}^T e^{-\Lambda(t-T_n)}x_n e^{\Lambda(s-T_n)}x_n =$$

$$e^{\Lambda(s-T_n)}x_n x_n^T e^{-\Lambda(t-T_n)}\mathbf{1} =$$

$$1_n(s)e^{\Lambda(t-s)}x_n = 1_n(s)p_0(t-s)x_n.$$

We repeatedly used the fact that $x_n x_n^T$, $e^{\Lambda(s-T_n)}$ etc. are diagonal matrices and that hence their products commute.
Before proceeding to the case $k \geq 1$ we introduce some auxiliary functions. Define recursively the functions $F_k : \mathbb{R} \to \mathbb{R}^{m\times m}$ for $k \geq 0$ by (I is the identity matrix)

$$\begin{aligned} F_0(t) &= 1_{\{t\geq 0\}}I \\ F_k(t) &= 1_{\{t\geq 0\}}\int_0^t e^{\Lambda u}F_{k-1}(t-u)A'\Lambda e^{-\Lambda u}du, \end{aligned}$$

for $k \geq 1$. Notice that all the F_k are zero for negative arguments. Then the following statement is true for all $n, k, j \geq 0$:

$$E[e^{-\Lambda(t-T_{n+k})}F_j(t-T_{n+k})x_{n+k}|\mathcal{G}_n] = e^{-\Lambda(t-T_n)}F_{j+k}(t-T_n)x_n. \tag{3.10}$$

Equation (3.10) can be proved as follows. Assume that $k \geq 1$, otherwise there is almost nothing to prove. Recall that S_{k+1} given $\mathcal{G}_k$ has an exponential distribution with density $\mathbf{1}^T\Lambda e^{-\Lambda s}x_k$, that x_{k+1} and S_{k+1} are conditionally independent given $\mathcal{G}_k$ and that diagonal matrices (as above) commute.

$$E[e^{-\Lambda(t-T_{n+k})}F_j(t-T_{n+k})x_{n+k}|\mathcal{G}_{n+k-1}] =$$

$$E[e^{-\Lambda(t-T_{n+k-1}-S_{n+k})}F_j(t-T_{n+k-1}-S_{n+k})x_{n+k}|\mathcal{G}_{n+k-1}] =$$

$$e^{-\Lambda(t-T_{n+k-1})}\int_0^{t-T_{n+k-1}} e^{\Lambda s}\mathbf{1}^T\Lambda e^{-\Lambda s}x_{n+k-1}F_j(t-T_{n+k-1}-s)dsA'x_{n+k-1} =$$

$$e^{-\Lambda(t-T_{n+k-1})}\int_0^{t-T_{n+k-1}} e^{\Lambda s}F_j(t-T_{n+k-1}-s)dsA'\Lambda e^{-\Lambda s}x_{n+k-1}ds =$$

$$e^{-\Lambda(t-T_{n+k-1})}F_{j+1}(t-T_{n+k-1})x_{n+k-1}.$$

Iterate this procedure (and use repeated conditioning) another $k-1$ times to obtain equation (3.10).

Consider now $E[x_{n+k}1_{n+k}(t)1_n(s)|\mathcal{H}_s]$ for $k \geq 1$. In view of lemma 3.8 we can write this as the product

$$E[x_{n+k}1_{n+k}(t)1_n(s)|\mathcal{G}_n]1_n(s)x_n^T e^{\Lambda(s-T_n)}\mathbf{1}. \tag{3.11}$$

Consider the last conditional expectation. It equals

$$E[x_{n+k}1_n(s)E[1_{n+k}(t)|\mathcal{G}_{n+k}]|\mathcal{G}_n] =$$

$$E[x_{n+k}1_n(s)1_{\{T_{n+k}\leq t\}}x_{n+k}^T e^{-\Lambda(t-T_{n+k})}\mathbf{1}|\mathcal{G}_n] =$$

$$E[1_n(s)1_{\{T_{n+k}\leq t\}}e^{-\Lambda(t-T_{n+k})}x_{n+k}|\mathcal{G}_n] =$$

$$E[1_n(s)e^{-\Lambda(t-T_{n+k})}F_0(t-T_{n+k})x_{n+k}|\mathcal{G}_n] =$$

$$E[1_n(s)E[e^{-\Lambda(t-T_{n+k})}F_0(t-T_{n+k})x_{n+k}|\mathcal{G}_{n+1}|\mathcal{G}_n] = \text{(by equation (3.10))}$$

$$E[1_n(s)e^{-\Lambda(t-T_{n+1})}F_{k-1}(t-T_{n+1})x_{n+1}|\mathcal{G}_n] =$$

$$E[1_n(s)e^{-\Lambda(t-T_n-S_{n+1})}F_{k-1}(t-T_n-S_{n+1})x_{n+1}|\mathcal{G}_n] =$$

$$E[1_n(s)e^{-\Lambda(t-T_n-S_{n+1})}F_{k-1}(t-T_n-S_{n+1})|\mathcal{G}_n]E[x_{n+1}|\mathcal{G}_n] =$$

$$1_{\{T_n\leq s\}}\int_{s-T_n}^{t-T_n} e^{-\Lambda(t-T_n-u)}F_{k-1}(t-T_n-u)\mathbf{1}^T\Lambda e^{-\Lambda u}x_n du A' x_n =$$

$$1_{\{T_n\leq s\}}\int_{s-T_n}^{t-T_n} e^{-\Lambda(t-T_n-u)}F_{k-1}(t-T_n-u)A'\Lambda e^{-\Lambda u}x_n du =$$

$$1_{\{T_n\leq s\}}\int_0^{t-s} e^{-\Lambda(t-s-v)}F_{k-1}(t-s-v)A'\Lambda e^{-\Lambda v}dv e^{-\Lambda(s-T_n)}x_n =$$

$$1_{\{T_n\leq s\}}e^{-\Lambda(t-s)}F_k(t-s)e^{-\Lambda(s-T_n)}x_n.$$

Use this result together with equation (3.5) to obtain that equation (3.11) becomes

$$1_n(s)e^{-\Lambda(t-s)}F_k(t-s)e^{-\Lambda(s-T_n)}x_n x_n^T e^{\Lambda(s-T_n)}\mathbf{1} =$$

$$1_n(s)e^{-\Lambda(t-s)}F_k(t-s)e^{-\Lambda(s-T_n)}x_n.$$

Define then for $k \geq 0$ the functions p_k by $p_k(t) = e^{-\Lambda t}F_k(t)$. Then one obtains from the definition of the F_k that $p_k(t) =$

$$p_k(t) = e^{-\Lambda t}\int_0^t e^{\Lambda u}F_{k-1}(t-u)A'\Lambda e^{-\Lambda u}du =$$

$$\int_0^t e^{-\Lambda(t-u)}F_{k-1}(t-u)A'\Lambda e^{-\Lambda u}du =$$

$$\int_0^t p_{k-1}(t-u)A'\Lambda e^{-\Lambda u}du =$$

$$\int_0^t p_{k-1}(s)A'\Lambda e^{-\Lambda(t-s)}ds,$$

which proves the lemma for $k \geq 1$.

REMARK: Actually we proved by other methods in a slightly more general setting equation (1.14) in section VI.1 of [6]. The difference in approach is that we started from discrete time processes, whereas in [6] use is made of the fact that X is Markov process in continuous time, which we still have to prove. This is the content of the next

Theorem 3.10 *For $t \geq s \geq 0$ it holds that*

$$E[X_t|\mathcal{H}_s] = e^{A(t-s)}X_s, \tag{3.12}$$

with $A = (A' - I)\Lambda$. So X is a Markov process with respect to the filtration $\mathbb{H}$ with transition intensities given by A.

PROOF: Define $P(t) = \sum_0^\infty p_k(t)$. Then $P(0) = I$, and from lemma 3.9 it follows that P satisfies the linear differential equation $\dot{P} = PA$. So $P(t) = \exp(At)$. (This equation is actually equivalent to equation (4) on page 203 in [7]).
Consider now

$$E[X_t|\mathcal{H}_s] = \sum_{j=0}^\infty E[x_j 1_j(t)|\mathcal{H}_s] = \sum_{j=0}^\infty \sum_{n=0}^\infty E[x_j 1_j(t)1_n(s)|\mathcal{H}_s] =$$

$$\sum_{n=0}^\infty \sum_{j=n}^\infty E[x_j 1_j(t)1_n(s)|\mathcal{H}_s],$$

because $1_j(t)1_n(s) = 0$ for $j < n$ if $s < t$. Use lemma 3.9 to write the last expression as

$$\sum_{n=0}^\infty \sum_{k=0}^\infty p_k(t-s)x_n 1_n(s) =$$

$$P(t-s)X_s = e^{A(t-s)}X_s,$$

from the first part of the proof. This establishes the theorem. □

Notice that the Markov process X constructed here has jump times that in general differ from the T_k, since it is not explicitly assumed that the diagonal elements of A' are zero. Denote by $\tilde{T}_k$ the jump times of X. Then of course for almost all ω the sequence $\{\tilde{T}_k(\omega)\}$ is a subsequence of $\{T_k(\omega)\}$ and $\tilde{T}_k(\omega) = \inf\{t > \tilde{T}_{k-1} : x_t \neq x_{\tilde{T}_{k-1}}\}$. In order to avoid uninteresting complications, we assume that x has no absorbing states, so all the $A'_{ii} \neq 1$. If we denote in this case the embedded Markov chain by $\tilde{x}$, then it follows from the results of the previous section, that $\tilde{x}$ has the transition matrix $\tilde{A}$ and the interarrival times $\tilde{S}_{k+1} = \tilde{T}_{k+1} - \tilde{T}_k$ are exponentially distributed given $\mathcal{H}_{\tilde{T}_k}$ with mean $\mathbf{1}^T\tilde{\Lambda}^{-1}\tilde{x}_{\tilde{T}_k}$, where the matrix $\tilde{A}$ has as its entries zeros on the diagonal and outside it

$$\tilde{A}_{ij} = \frac{A'_{ij}}{1 - A'_{jj}},$$

and the diagonal matrix $\tilde{\Lambda}$ has entries $\tilde{\Lambda}_{ii} = \lambda_i(1 - A'_{ii})$. (Notice that $\tilde{\Lambda} = \Lambda$ and $\tilde{A} = A'$ if all the A'_{ii} are zero).

By letting the $\mathcal{H}_{\tilde{T}_n}$ play the role of the $\mathcal{G}_n$, one can together with the $\tilde{T}_n$ sequence instead of the T_n, construct a filtration $\tilde{\mathbb{H}}$. The following statement is true: $\forall t \geq 0 :$ $\tilde{\mathcal{H}}_t \subset \mathcal{H}_t$.
This can be proved as follows. Let $H \in \tilde{\mathcal{H}}_t$. Then $H \cap \{\tilde{T}_\infty \leq t\} \in \mathcal{H}_t$, because $\tilde{T}_\infty \geq T_\infty$. Furthermore, for all n there is $H_n \in \mathcal{H}_{\tilde{T}_n}$ such that $H \cap \{\tilde{T}_n \leq t < \tilde{T}_{n+1}\} = H_n \cap \{\tilde{T}_n \leq t < \tilde{T}_{n+1}\}$ which belongs to $\mathcal{H}_t$, because of the fact that the $\tilde{T}_n$ are $\mathbb{H}$-stopping times and the definition of $\mathcal{H}_{\tilde{T}_n}$. By taking the union over all n we obtain that $H \cap \{\tilde{T}_\infty > t\} \in \mathcal{H}_t$. So we conclude that $H \in \mathcal{H}_t$.
Since we can alternatively write $X_t = \sum_{n=0}^{\infty} \tilde{x}_n 1_{\{\tilde{T}_n \leq t < \tilde{T}_{n+1}\}}$, it follows that X is also $\tilde{\mathbb{H}}$-Markov. This filtration $\tilde{\mathbb{H}}$ enjoys some minimality properties, but is in general still bigger than the filtration generated by X itself.

References

[1] R. Boel, P. Varaiya, E. Wong, Martingales on Jump Processes. I: Representation Results, SIAM J. Control vol. 13, no. 5, pp 999-1021.

[2] R. Boel, P. Varaiya, E. Wong, Martingales on Jump Processes. II: Applications, SIAM J. Control vol. 13, no. 5, pp 1022-1061.

[3] P. Brémaud, *Point processes and queues*, Springer.

[4] E. Çinlar, *Introduction to stochastic processes*, Prentice Hall.

[5] C. Dellacherie, P.A. Meyer, *Probabilités et Potentiel*, Hermann.

[6] J.L. Doob, *Stochastic Processes*, Wiley.

[7] I.I. Gihman, A.V. Skorohod, *The Theory of Stochastic Processes II*, Springer.

[8] R.D. Gill, S. Johansen, *A survey of product-integration with a view toward application in survival analysis*, Ann. Stat., Vol. 18, No. 4, pp 1501-1555.

[9] S.W. He, J.G. Wang, *Some results on jump processes*, Sém. Prob. XVIII, Springer LNM 1059, pp. 256-267.

[10] M. Itmi, *Processus ponctuels marqués stochastiques*, Sém. Prob. XV, LNM 850, pp. 618-626.

[11] M. Jacobsen, *A characterization of minimal Markov jump processes*, Z. Wahrscheinlichkeitstheorie ver. Geb. 23, pp 32-46.

[12] J. Jacod, *Calcul Stochastique et Problèmes de Martingales*, Springer LNM 714.

[13] J. Jacod, A.N. Shiryaev, *Limit Theorems for Stochastic Processes*, Springer.

[14] J. Jacod, A.V. Skorohod, *Jumping filtrations and martingales with finite variation*, Prépublication 164 du laboratoire de probabilités de l'université Paris VI.

[15] D. Lépingle, P.A. Meyer, M. Yor, *Extrémalité et remplissage de tribus pour certaines martingales purement discontinus*, Sém. Prob. XV, Springer LNM 850, pp. 604-617.

[16] R.S. Liptser, A.N. Shiryaev, *Theory of Martingales*, Kluwer.

[17] P.J.C. Spreij (1990), *Self-exciting counting process systems with finite state space*, Stoch. Proc. Appl. 34, pp. 275-295.

A Appendix

If $\mathcal{F}$ is a σ-algebra on a set Ω and A an arbitrary subset of Ω, then we denote by $\mathcal{F} \cap A$ the induced *trace σ-algebra* on A. So $\mathcal{F} \cap A$ is the collection $\{F \cap A : F \in \mathcal{F}\}$.

Lemma A.1 *Let $\mathcal{F}$ and $\mathcal{G}$ be two σ-algebras on a set Ω. Let $A_i \subset \Omega$ for each $i \in I\!N$ such that $A_i \subset A_{i+1}$ and $\bigcup_{i=0}^{\infty} A_i = \Omega$. If for all $i \in I\!N$ the inclusion $\mathcal{F} \cap A_i \subset \mathcal{G} \cap A_i$ holds, then also $\mathcal{F} \subset \mathcal{G}$.*

For a proof see Brémaud [3], lemma L29. Notice that no measurability properties of the A_i are required. Essential in this lemma is however the fact that the A_i are nested. Without this property the assertion of the lemma is in general false as is shown by the next example. Take $\Omega = \{1,2,3,4,5,6\}$, $F_1 = \{1,2,4\}, F_2 = \{3,5,6\}, G_1 = \{1,4,5\}, G_2 = \{2,3,6\}$, $\mathcal{F} = \{\emptyset, F_1, F_2, \Omega\}$ and $\mathcal{G} = \{\emptyset, G_1, G_2, \Omega\}$. Define $A_1 = \{1,4\}$, $A_2 = \{2,5\}$ and $A_3 = \{3,6\}$. Then for $i = 1,2,3$ the equality $\mathcal{F} \cap A_i = \mathcal{G} \cap A_i$ holds, which is easy to check. But clearly there is no inclusion relation between $\mathcal{F}$ and $\mathcal{G}$.

The next proposition below generalizes proposition 2.1 . So let $(\Omega, \mathcal{F}, \mathbb{F}, P)$ be a filtered probability space and $X : \Omega \times [0,\infty) \to B^m$ be an $\mathbb{F}$-adapted cadlag stochastic process. Denote by $\Phi(t,s)$ the matrix with entries $\Phi_{ij}(t,s) = P(X_t = b_i | X_s = b_j)$. The following holds true.

Lemma A.2 *If X is $\mathbb{F}$-Markov, then there is a function of bounded variation Q with values in $\mathbb{R}^{m\times m}$ such that $\Phi(t,s) = \mathcal{E}(Q(s+.))_{t-s}$ for all $t \geq s$, where $\mathcal{E}$ means the Doléans exponent.*

PROOF: This follows from Gill & Johansen [8]. Using their theorem 15, we have $Q(t)=\int_{(0,t]}d(\Phi-I)$, and $\Phi(t,s)=\prod_{(s,t]}(I+dQ)$, where the $\prod$ here stands for the product-integral defined as a limit of matrix products, in which the ordering of the product is the *opposite* of the one in [8]. As a consequence their formula (40) now takes the form

$$\Phi(t,s)=I+\int_{(s,t]}dQ(u)\Phi(u-,s) \tag{A.1}$$

from which the assertion follows. □

REMARK: The form of the function Q follows from results on product-integration. However, in two extreme cases it is easy to define Q without the theory of product-integration. Consider first the case in which $\Phi(.,0)$ is differentiable. Then $\Phi(t,0)$ is invertible and $Q(t)$ is simply $\int_0^t \dot{\Phi}(u,0)\Phi(u,0)^{-1}du$.
In the other case we assume that X is a Markov chain in discrete time on the integers with $\Phi(t+1,t)=A(t)$. Then $Q(t)=\sum_{k=0}^{t-1}(A(k)-I)$.

Proposition A.3 *If X is $\mathbb{F}$-Markov, then there is a bounded variation function Q such that M defined by*

$$M_t=X_t-X_0-\int_{(0,t]}dQ(s)X_{s-} \tag{A.2}$$

is an $\mathbb{F}$-martingale.
Conversely, if there is a martingale M and a bounded variation function Q such that X is a solution of equation (A.2)*, then X is $\mathbb{F}$-Markov with transition probabilities as in lemma A.2.*

PROOF: Using the fact that $E[X_t|\mathcal{F}_s]=\Phi(t,s)X_s$, the definition of M and lemma A.2 we compute the conditional expectation

$$\begin{aligned}
E[M_t|\mathcal{F}_s] &= E[X_t|\mathcal{F}_s]-X_0-\int_{(0,s]}dQ(u)X_{u-}-E[\int_{(s,t]}dQ(u)X_{u-}|\mathcal{F}_s] \\
&= \Phi(t,s)X_s+M_s-X_s-\int_{(s,t]}dQ(u)\Phi(u-,s)X_s \\
&= M_s+[\Phi(t,s)-I-\int_{(s,t]}dQ(u)\Phi(u-,s)]X_s
\end{aligned}$$

and the result follows from equation (A.1).
For the proof of the converse statement we compute

$$\begin{aligned}
E[X_t|\mathcal{F}_s] &= X_0+\int_{(0,s]}dQ(u)X_{u-}+E[\int_{(s,t]}dQ(u)X_{u-}|\mathcal{F}_s]+M_s \\
&= X_s+\int_{(s,t]}dQ(u)E[X_{u-}|\mathcal{F}_s]
\end{aligned}$$

So, $E[X_t|\mathcal{F}_s]$ satisfies a Volterra in equation t which has a unique solution given by $E[X_t|\mathcal{F}_s]=\Phi(t,s)X_s$ in view of the fact that $\Phi(t,s)$ satisfies equation (A.1). □

A closer inspection of the second part of this proof yields that if the process X is $\mathbb{F}$-Markov, it also satisfies the Strong Markov property:

Proposition A.4 *Let X be $\mathbb{F}$-Markov and T an a.s. finite $\mathbb{F}$-stopping time. Then $E[X_{T+t}|\mathcal{F}_T] = \Phi(T+t,T)X_T$ for all $t \geq 0$.*

PROOF: Since we know that X satisfies equation (A.2) we find by application of Fubini's theorem for conditional expectations that $E[X_{T+t}|\mathcal{F}_T] =$
$X_T + E[\int_{(T,T+t]} dQ(u)X_{u-}|\mathcal{F}_T] = X_T + E[\int dQ(u)1_{\{T<u\leq T+t\}}X_{u-}|\mathcal{F}_T] =$
$X_T + \int dQ(u)E[1_{\{T<u\leq T+t\}}X_{u-}|\mathcal{F}_T] = X_T + \int_{(T,T+t]} E[X_{u-}|\mathcal{F}_T]$. So we obtain again a Volterra equation, the solution of which is the desired expression. $\square$

For convenience we here state what is also known under the name Kallianpur-Striebel formula.

Lemma A.5 *Let $(\Omega, \mathcal{F}, P)$ be a probability space. Let Q be another probability measure which is assumed to be absolutely continuous w.r.t. P. Let ξ be a version of the Radon-Nikodym derivative $\frac{dQ}{dP}$. Let Z be a random variable that is integrable w.r.t. Q and let $\mathcal{G}$ be a sub σ-algebra of $\mathcal{F}$. Then Q-a.s.*

$$E_Q[Z|\mathcal{G}] = \frac{E_P[Z\xi|\mathcal{G}]}{E_P[\xi|\mathcal{G}]}$$

PROOF: Let $G \in \mathcal{G}$ and use the definition of conditional expectation and Radon-Nikodym derivative to verify the following string of equalities.
$\int_G E_P[Z\xi|\mathcal{G}]dP = \int_G Z\xi dP = \int_G ZdQ = \int_G E_Q[Z|\mathcal{G}]dQ = \int_G E_Q[Z|\mathcal{G}]\xi dP =$
$\int_G E_P[E_Q[Z|\mathcal{G}]\xi|\mathcal{G}]dP = \int_G E_Q[Z|\mathcal{G}]E_P[\xi|\mathcal{G}]dP$.
Since this holds for all $G \in \mathcal{G}$, we obtain $E_P[Z\xi|\mathcal{G}] = E_Q[Z|\mathcal{G}]E_P[\xi|\mathcal{G}]$, P-a.s. The assertion follows by noting that $Q(E_P[\xi|\mathcal{G}] = 0) = 0$. $\square$