Stochastic Processes and their Applications 34 (1990) 275-295
North-Holland

SELF-EXCITING COUNTING PROCESS SYSTEMS WITH FINITE STATE SPACE

Peter SPREIJ

Department of Econometrics, Free University, PO Box 7161, 1007 MC Amsterdam, The Netherlands

Received 31 October 1986
Revised 5 July 1989

Stochastic systems with counting process output and a finite state space are considered. This leads to studying processes with finite state space that are Markovian with respect to the flow of σ-algebras, that is generated by the counting process. It appears that there is a close relationship between the transition intensities of the Markov process and the intensity of the counting process. Some consequences for a stochastic realization problem are then studied.

AMS 1980 Subject Classifications: 93E03, 60G55, 60J75.

counting process * Markov process * stochastic realization theory

1. Introduction

Suppose that the dynamic behaviour of some phenomenon may be modelled by means of a counting process. It is then attractive to model the intensity of this counting process as a Markov process evolving on a finite state space. A practical situation where this model shows a very satisfactory behaviour is reported in e.g. Kemp (1986). In the case that one can observe the counting process, but not the associated finite state Markov process there exists a finite dimensional filter that estimates the Markov process. The existence of such a filter is one of the advantages of this model.

On the other hand it has been argued (see Boel, 1985) that in a situation where one cannot observe a state process and where there are no physical grounds that lead to an obvious choice of a state model, it is perhaps better to use self-exciting models for identification purposes.

Here, in a way, we adopt both these points and the question arises whether this yields an interesting model. To put it a little bit more precise, we want to characterize the class of counting processes that admit an intensity, which is a function of a finite state process which is Markov with respect to the flow of σ-algebras generated by such a counting process. Or, to formulate it in terms of a stochastic realization problem, given a counting process, under what conditions can be represented as the output of a stochastic system, where the state process assumes finitely many values, and is Markov with respect to the filtration generated by the output.

The purpose of this paper is to present a solution of the above stated problems. Thus we obtain a representation theory of counting processes via non-stationary Markov processes. In particular a detailed investigation will be made of finite state processes which are Markov with respect to some given counting process.

The paper is organized as follows. In Section 2 preliminary results for counting processes are reviewed. Section 3 contains results for finite state Markov processes. In particular, finite state Markov processes are characterized as solutions of certain stochastic differential equations. Section 4 reports the main results. A characterization of finite state processes which are Markov with respect to a counting process is given. In Section 5 the results of Section 4 will be used to solve a stochastic realization problem.

2. Basic results for counting processes

Good sources for the technical background of counting and jump processes are Brémaud (1981) or Boel, Varaiya and Wong (1975). Let (Ω, F, P) be a complete probability space. Let $n:\Omega\times[0,\infty)\to\mathbb{N}_0$ be a counting process and let $\mathscr{F}_t^n = \sigma\{n_s, s\leq t\}$ be the σ-algebra generated by the collection $(n_s, s\leq t\}$. Write $\mathbb{F}^n = \{\mathscr{F}_t^n, t\geq 0\}$. Assume that n admits the minimal decomposition

$$\mathrm{d}n_t = \lambda_t\,\mathrm{d}t + \mathrm{d}m_t$$

where $\lambda:\Omega\times[0,\infty)\to\mathbb{R}_+$ is the $\mathbb{F}^n$-predictable intensity process of n and $m:\Omega\times[0,\infty)\to\mathbb{R}$ is an $\mathbb{F}^n$-adapted martingale. The existence of an intensity is of crucial importance in this paper.

We recall the following result, known as the martingale representation theorem (see Brémaud, 1981, p. 76), since it plays a fundamental role.

Lemma 2.1. *Let $M:\Omega\times[0,\infty)\to\mathbb{R}$ be an $\mathbb{F}^n$-adapted martingale. Then there exists an $\mathbb{F}^n$-predictable process $k:\Omega\times[0,\infty)\to\mathbb{R}$ such that for all $t\geq 0$,*

$$M_t = M_0 + \int_0^t k_s(\mathrm{d}n_s - \lambda_s\,\mathrm{d}s).$$

The process k is $P(\mathrm{d}\omega)\lambda_t(\omega)\,\mathrm{d}t$ a.e. or equivalently $P(\mathrm{d}\omega)\,\mathrm{d}n_t(\omega)$ a.e. uniquely defined. □

In Section 4 the relation between two counting processes n and $\tilde{n}$ will be investigated. The following proposition will turn out to be useful there.

Proposition 2.2. *Let n and $\tilde{n}$ be two counting processes and let λ and $\tilde{\lambda}$ be their $\mathbb{F}^n$-, respectively $\mathbb{F}^{\tilde{n}}$-predictable intensities. Equivalent are:*

(i) *$\mathscr{F}_t^{\tilde{n}}\subset\mathscr{F}_t^n$, and $\mathscr{F}_\infty^{\tilde{n}}$ and $\mathscr{F}_t^n$ are conditionally independent given $\mathscr{F}_t^{\tilde{n}}$.*

(ii) *$\tilde{n}_t = \int_0^t 1_{\{\tilde{\lambda}_s>0\}}\,\mathrm{d}n_s$ and $\tilde{\lambda}_t = 1_{\{\tilde{\lambda}_t>0\}}\lambda_t$.*

The statements (i) and (ii) of this proposition can also be formulated as follows. Let $\{\tilde{T}_n\}$ and $\{T_n\}$ be the sequences of the jump times of $\tilde{n}$ and n. Then if $\tilde{T}_n < \infty$, there is k such that $\tilde{T}_n = T_k$ a.s., and whether for given T_k there is n such that $\tilde{T}_n = T_k$, depends only on $\sigma(T_k \wedge \tilde{T}_l,\ l \in \mathbb{N})$.

In the proof we will use the following lemma.

Lemma 2.3 (Brémaud and Yor, 1978). *Consider two filtrations $\mathbb{F}$ and $\mathbb{G}$, such that for all $t \geq 0$: $\mathcal{F}_t \subset \mathcal{G}_t$. Then there is equivalence between*:

(i) *Any $\mathbb{F}$-martingale is a $\mathbb{G}$-martingale.*

(ii) *$\mathcal{F}_\infty$ and $\mathcal{G}_t$ are conditionally independent given $\mathcal{F}_t$.* $\square$

Proof of Proposition 2.2. (i)$\Rightarrow$(ii): *Write* $\mathrm{d}\tilde{n}_t = \tilde{\lambda}_t\,\mathrm{d}t + \mathrm{d}\tilde{m}_t$, the Doob-Meyer decomposition of $\tilde{n}$ with respect to $\mathbb{F}^{\tilde{n}}$. From Lemma 2.3, $\tilde{m}$ is also an $\mathbb{F}^n$-martingale. Hence $\tilde{m}_t = \int_0^t h_s\,\mathrm{d}m_s$ for a $P(\mathrm{d}\omega)\,\mathrm{d}n_t(\omega)$ a.e. unique process h from Lemma 2.1. Then $\mathrm{d}\tilde{n}_t = (\tilde{\lambda}_t - \lambda_t h_t)\,\mathrm{d}t + h_t\,\mathrm{d}n_t$, which gives $\mathrm{d}\tilde{n}_t = h_t\,\mathrm{d}n_t$ and $\tilde{\lambda}_t = h_t\lambda_t$. Therefore on the jump times T_k of n we have $h^2_{T_k} = h_{T_k}$. Hence we can also write $\mathrm{d}\tilde{n}_t = h_t\,\mathrm{d}\tilde{n}_t = h_t\tilde{\lambda}_t\,\mathrm{d}t + h_t\,\mathrm{d}\tilde{m}_t$. From the fact that predictable intensities are unique, we find $\tilde{\lambda}_t = h_t\tilde{\lambda}_t$ a.s., which implies that $h_t 1_{\{\tilde{\lambda}_t > 0\}} = 1_{\{\tilde{\lambda}_t > 0\}}$. An obvious choice of h that satisfies this relation is $h'_t = 1_{\{\tilde{\lambda}_t > 0\}}$. It is certainly $\mathbb{F}^n$-predictable and

$$E\int_0^\infty 1_{\{h_t \neq 1_{\{\tilde{\lambda}_t > 0\}}\}}\,\mathrm{d}n_t = E\sum_{n\geq 1} 1_{\{h_{T_n} \neq 1_{\{\tilde{\lambda}_{T_n} > 0\}}\}}$$

$$= E\sum_{n\geq 1}\left[1_{\{h_{T_n}=1,\tilde{\lambda}_{T_n}=0\}} + 1_{\{h_{T_n}=0,\tilde{\lambda}_{T_n}>0\}}\right] = 0,$$

which can be seen as follows. It $h_{T_n} = 1$, then $\tilde{n}$ jumps at T_n, so that $\tilde{\lambda}_{T_n} > 0$, and if $h_{T_n} = 0$, then $\tilde{\lambda}_{T_n} = 0$ from $\tilde{\lambda}_t = h_t\tilde{\lambda}_t$. The uniqueness of the process h now gives the result.

(ii)$\Rightarrow$(i): Notice first that $\mathcal{F}^{\tilde{n}}_t \subset \mathcal{F}^n_t$, since by the assumption $\tilde{n}_t = \int_0^t 1_{\{\tilde{\lambda}_s > 0\}}\,\mathrm{d}n_s$, the sequence $\{\tilde{T}_k\}$ of jump times of $\tilde{n}$ is contained in the sequence $\{T_k\}$. In view of Lemma 2.3 it is now sufficient to prove that any $\mathbb{F}^{\tilde{n}}$-martingale is an $\mathbb{F}^n$-martingale. So let M be an $\mathbb{F}^{\tilde{n}}$-martingale. Then there is an $\mathbb{F}^{\tilde{n}}$-predictable process h such that $M_t = M_0 + \int_0^t h_s\,\mathrm{d}\tilde{m}_s$. Now

$$\lambda_t\,\mathrm{d}t + \mathrm{d}\tilde{m}_t = \mathrm{d}\tilde{n}_t = 1_{\{\tilde{\lambda}_t > 0\}}\,\mathrm{d}n_t = 1_{\{\tilde{\lambda}_t>0\}}\lambda_t\,\mathrm{d}t + 1_{\{\tilde{\lambda}_t>0\}}\,\mathrm{d}m_t = \tilde{\lambda}_t\,\mathrm{d}t + 1_{\{\tilde{\lambda}_t>0\}}\,\mathrm{d}m_t$$

by assumption. Because of $\mathcal{F}^{\tilde{n}}_t \subset \mathcal{F}^n_t$, the process $1_{\{\tilde{\lambda}_t > 0\}}$ is $\mathbb{F}^n$-predictable, hence $\tilde{m}$ is also an $\mathbb{F}^n$-martingale. But then the same conclusions holds for M.

Remark. The formulation of condition (ii) of Proposition 2.2 can be replaced by:

(ii)$'$ There exists an $\mathbb{F}^n$-predictable process u such that

$$\tilde{n}_t = \int_0^t u_s\,\mathrm{d}n_s \quad \text{and} \quad \tilde{\lambda}_t = u_t\lambda_t.$$

It then follows that one can identify u as $u_t = 1_{\{\tilde{\lambda}_t > 0\}}$.

3. Markov processes with a finite state space

3.1. Recall first that a stochastic process X with state space $(E, \mathscr{E})$ is Markov with respect to some filtration $\mathbb{F}=\{\mathscr{F}_t, t\geq 0\}$ (we will say that it is $\mathbb{F}$-Markov) if $\forall t\geq s$, $\forall B\in\mathscr{E}$,

$$P(X_t\in B\,|\,\mathscr{F}_s)=P(X_t\in B\,|\,\sigma(X_s)).$$

Or, equivalently, that for all bounded measurable functions f on E we have

$$E[f(X_t)\,|\,\mathscr{F}_s]=E[f(X_t)\,|\,\sigma(X_s)].$$

From now on we specialize to the case where the state space E is finite, $E=\{c_1,\dots,c_n\}$, and $\mathscr{E}$ is the power set of E. Define $Y:\Omega\times[0,\infty)\to\{0,1\}^n$ by its components $Y_{it}:=1_{\{X_t=c_i\}}$. Denote by $\Phi(t,s)$ the matrix of transition probabilities of X. That is for $t\geq s$, with the notation $z^+=z^{-1}1_{\{z\neq 0\}}$ and the understanding $0/0=0$,

$$\Phi_{ij}(t,s)=P(X_t=c_i\,|\,X_s=c_j)=(EY_{js})^+E(Y_{js}Y_{it}).$$

Then we have the following well known facts. Semigroup property: $\Phi(t,s)=\Phi(t,u)\Phi(u,s)$ for $t\geq u\geq s$. Assume that for all $t\geq 0$ the following limit exists:

$$A(t):=\lim_{h\downarrow 0}\frac{1}{h}[\Phi(t+h,t)-I].$$

So $A(t)$ has nonpositive diagonal elements, the other entries are nonnegative and the column sums are zero. Such a matrix will be called a Markov matrix. Then $(\partial/\partial t)\Phi(t,s)=A(t)\Phi(t,s)$. In particular $(\partial/\partial t)\Phi(t,0)=A(t)\Phi(t,0)$. From this equation we get $\det\Phi(t,0)=\exp(\int_0^t \operatorname{tr}A(s)\,\mathrm{d}s)$. Hence by definition of $A(t)$, we see that $\Phi(t,0)$ is invertible for all $t\geq 0$.

Proposition 3.1. *Define $Z:\Omega\times[0,\infty)\to\mathbb{R}^n$ by $Z_t=\Phi(t,0)^{-1}Y_t$. Then Z is an $\mathbb{F}$-martingale and Y satisfies the stochastic differential equation*

$$\mathrm{d}Y_t=A(t)Y_t\,\mathrm{d}t+\Phi(t,0)\,\mathrm{d}Z_t. \tag{3.1}$$

Proof. Using a representation of a conditional expectation when the conditioning σ-algebra is generated by a finite number of disjoint sets we get

$$\begin{aligned}E[Z_t\,|\,\mathscr{F}_s]&=\Phi(t,0)^{-1}E[Y_t\,|\,\mathscr{F}_s]=\Phi(t,0)^{-1}E(Y_t\,|\,\sigma(X_s)]\\&=\Phi(t,0)^{-1}E[Y_t\,|\,\sigma(Y_s)]=\Phi(t,0)^{-1}\sum_j E[Y_{js}]^+E[Y_tY_{js}]Y_{js}\\&=\Phi(t,0)^{-1}\Phi(t,s)Y_s=\Phi(s,0)^{-1}Y_s=Z_s.\end{aligned}$$

The second assertion can easily be proved by applying the stochastic differentiation rule to the product $Y_t=\Phi(t,0)Z_t$. □

Notice that $\int_0^t\Phi(s,0)\,\mathrm{d}Z_s$ apearing in (3.1) is again an $\mathbb{F}$-martingale since $\Phi(\cdot,0)$ is trivially predictable. Proposition 3.1 thus gives a representation of Markov processes in terms of a linear stochastic differential equation driven by a martingale. The next result gives a converse statement.

Proposition 3.2. *Let $X:\Omega\times[0,\infty)\to\{c_1,\dots,c_n\}$ be a stochastic process, $\mathbb{F}$-adapted, and let Y be associated with X as before. Assume that Y satisfies*

$$\mathrm{d}Y_t = A(t)\,Y_t\,\mathrm{d}t + \mathrm{d}m_t^Y. \tag{3.2}$$

Here $A:[0,\infty)\to\mathbb{R}^{n\times n}$ is a Lebesgue measurable function (deterministic !) and m^Y an $\mathbb{F}$-adapted martingale. Then X and Y are $\mathbb{F}$-Markov processes.

Proof. We have to prove that $E[f(X_t)\,|\,\mathcal{F}_s]=E[f(x_t)\,|\,\sigma(X_s)]$ for all $f:\{c_1,\dots,c_n\}\to\mathbb{R}$. Since $f(X_t)=\sum_j f(c_j)\,Y_{jt}$ we will only prove $E[Y_t\,|\,\mathcal{F}_s]=E[Y_t\,|\,\sigma(x_s)]$.

Let $B(t)$ be the solution of $\dot B(t)=A(t)B(t)$ with $B(0)=I$. Now we can write the solution Y_t of (3.2) as

$$Y_t = B(t)\,Y_0 + B(t)\int_0^t B^{-1}(s)\,\mathrm{d}m_s^Y.$$

Notice again that $\int_0^t B^{-1}(s)\,\mathrm{d}m_s^Y$ is an $\mathbb{F}$-martingale, $B(t)$ deterministic. Hence

$$\begin{aligned} E[Y_t\,|\,\mathcal{F}_s] &= B(t)\,Y_0 + B(t)\int_0^s B^{-1}(u)\,\mathrm{d}m_u^Y \\ &= B(t)\,Y_0 + B(t)[B^{-1}(s)\,Y_s - Y_0] = B(t)B^{-1}(s)\,Y_s. \end{aligned}$$

Since we have $\sigma(X_s)=\sigma(Y_s)\subset\mathcal{F}_s$ we get

$$\begin{aligned} E[Y_t\,|\,\sigma(X_s)] &= E[E[Y_t\,|\,\mathcal{F}_s]\,|\,\sigma(Y_s)] = E[B(t)B^{-1}(s)\,Y_s\,|\,\sigma(Y_s)] \\ &= B(t)B^{-1}(s)\,Y_s = E[Y_t\,|\,\mathcal{F}_s]. \qquad \square \end{aligned}$$

Concluding we see that the statement X and Y are $\mathbb{F}$-Markov is equivalent with saying that the indicator process Y satisfies equation (3.2).

Propositions 3.1 and 3.2 will play an important role in Section 4. Here is another illustration of the usefulness of this result.

Applying Propositions 3.1 and 3.2 to the case where X is a counting process and $\mathbb{F}=\mathbb{F}^n$, we easily obtain an intuitively appealing criterion to show, in terms of the predictable intensity, whether or not a counting process is Markov (see also Jacobsen, 1982). Of course we need a generalization of Propositions 3.1, 3.2 to include processes that assume countably many values, but this is straightforward in this situation, because of the special lower triangular form of the matrix $A(t)$ in the proof below.

Proposition 3.3. *Let n be a counting process, and λ its $\mathbb{F}^n$-predictable intensity process: $\mathrm{d}n_t=\lambda_t\,\mathrm{d}t+\mathrm{d}m_t$. Equivalent are:*

(i) *n is ($\mathbb{F}^n$-) Markov.*

(ii) *There exists a measurable $f:[0,\infty)\times\mathbb{N}_0\to[0,\infty)$ such that $\lambda_t=f(t,n_{t-})$.*

Proof. (i)⇒(ii): Let Y be the indicator process associated with n and let N^{T} be the vector $[0, 1, 2, \ldots]$. (Here and elsewhere T denotes transposition.) Then $n_t = N^{\mathrm{T}} Y_t$ and Y satisfies by assumption

$$\mathrm{d}Y_t = A(t) Y_t \,\mathrm{d}t + \mathrm{d}m_t^Y.$$

On the other hand we have immediately from the definition of Y:

$$\mathrm{d}Y_t = (J - I) Y_{t-} \,\mathrm{d}n_t,$$

where J is defined by its entries $J_{kl} = \delta_{k,l+1}$, $k, l \geq 0$ and $I_{kl} = \delta_{kl}$, $k, l \geq 0$. Then

$$\mathrm{d}Y_t = (J - I) Y_t \lambda_t \,\mathrm{d}t + (J - I) Y_{t-} \,\mathrm{d}m_t.$$

Since each component of Y is a special semimartingale we have from the uniqueness of the decomposition for all t,

$$\int_0^t (J - I) Y_{s-} \lambda_{s-} \,\mathrm{d}s = \int_0^t A(s) Y_{s-} \,\mathrm{d}s.$$

Since all processes above are left continuous we have for all $t > 0$: $(J - I) Y_{t-} \lambda_t = A(t) Y_{t-}$. After multiplying this equation by Y_{t-}^{T} we get

$$-\lambda_{t-} = Y_{t-}^{\mathrm{T}} A(t) Y_{t-} = -\sum_{i \geq 0} A_{ii}(t) Y_{it-}.$$

Define now f by $f(t, n) = -A_{nn}(t)$, then

$$\lambda_{t-} = f(t, n_{t-}).$$

Then λ, being predictable, is indistinguishable from $f(t, n_{t-})$.

(ii)⇒(i): Define $F(t) \in \mathbb{R}^{\mathbb{N}_0}$ by $F_n(t) = f(t, n)$. Hence

$$\lambda_t = F(t)^{\mathrm{T}} Y_t.$$

As in part (i) of the proof

$$\mathrm{d}Y_t = (J - I) Y_{t-} \,\mathrm{d}n_t.$$

Hence

$$\begin{aligned}\mathrm{d}Y_t &= (J - I) Y_t Y_t^{\mathrm{T}} F(t) \,\mathrm{d}t + (J - I) Y_{t-} \,\mathrm{d}m_t \\ &= (J - I) \mathrm{diag}(Y_t) F(t) \,\mathrm{d}t + (J - I) Y_{t-} \,\mathrm{d}m_t\end{aligned}$$

where $\mathrm{diag}(Y_t)$ is the diagonal matrix with entries $(\mathrm{diag}(Y_t))_{ij} = \delta_{ij} Y_{it}$.

Define $A(t) \in R^{\mathbb{N}_0 \times \mathbb{N}_0}$ by $A_{kl}(t) = (J - I)_{kl} F_l(t)$, then

$$A(t) Y_t = (J - I) \,\mathrm{diag}(Y_t) F(t)$$

and

$$\mathrm{d}Y_t = A(t) Y_t \,\mathrm{d}t + (J - I) Y_{t-} \,\mathrm{d}m_t,$$

which is of the form as in Proposition 3.2. □

3.2. From the equivalence of $\mathbb{F}$-Markov processes and solutions of certain linear stochastic differential equations (Propositions 3.1 and 3.2) it is easy to see when functions of a Markov chain again yield a Markov chain.

To be specific let as before X be an $\mathbb{F}$-Markov chain with state space $E = \{c_1, \ldots, c_n\}$. Let H be another set and $f: E \to H$ a function. Clearly $f(X)$ is again Markov if f is injective. To avoid trivialities let us assume that $H = \{h_1, \ldots, h_m\}$, $m < n$ and that f is onto. Write $Z_t = f(X_t)$. Associate with Z the indicator process W as usual:

$$W: \Omega \times [0, \infty) \to \{0, 1\}^m, \quad W_{it} = 1_{\{Z_t = h_i\}}.$$

Define $F \in \mathbb{R}^{m \times n}$ by $F_{ij} = 1_{\{f(e_j) = h_i\}}$. Notice that $\mathbf{1}_m^{\mathrm{T}} F = \mathbf{1}_n^{\mathrm{T}}$, where $\mathbf{1}_m$ is a column vector with as elements $+1$. Then $W_t = FY_t$. Observe that W, like Y, is a special semimartingale. Notice that because f is onto F has rank m, i.e. it has full row rank. Let $K \in \mathbb{R}^{n \times (n-m)}$ be a fixed matrix such that its columns span Ker F. Let as before $A(t)$ be the matrix of transition intensities of X. We have the following.

Theorem 3.4. *Let X be $\mathbb{F}$-Markov with finite state space E. Let $f: E \to H$. Then $f(X)$ is again $\mathbb{F}$-Markov iff $FA(t)K = 0$ where K is any matrix whose columns span* Ker F *and F is related to f as indicated above. If this condition is satisfied, then the matrix $B(t)$ of transition intensities of $f(X)$ is given by $B(t) = FA(t)\hat{F}$, where $\hat{F}$ is any right inverse of F.*

Proof. We have $\mathrm{d}Y_t = A(t)Y_t\,\mathrm{d}t + \mathrm{d}m_t^Y$. Hence

$$\mathrm{d}W_t = FA(t)Y_t\,\mathrm{d}t + F\,\mathrm{d}m_t^Y.$$

Now Z is $\mathbb{F}$-Markov iff $\mathrm{d}W_t = B(t)W_t\,\mathrm{d}t + \mathrm{d}m_t^W$ for some matrix-valued function B and an $\mathbb{F}$-martingale m^W. Hence we have Z is $\mathbb{F}$-Markov if and only if there is a $B(\cdot)$ such that $FA(t) = B(t)F$. Let $\hat{F}$ be a fixed right inverse of F. It exists, since F has full row rank. Then the last equation implies $B(t) = FA(t)\hat{F}$. Of course for B to be well defined it should not depend on the particular choice of $\hat{F}$.

Starting from $\hat{F}$ all other right inverses G of F are given by $G = \hat{F} + KX$, where $X \in \mathbb{R}^{(n-m)\times m}$ is an arbitrary matrix. Hence $B(t)$ is well defined iff $FA(t)\hat{F} = FA(t)(\hat{F} + KX)$ or iff $FA(t)K \equiv 0$. □

Remarks. (1) The determining condition $FA(t)K \equiv 0$ can be understood in two ways. Firstly for a given matrix $A(t)$ it tells us what functions f (if any) yield a Markov process $f(X)$. Secondly if one wants $f(X)$ to be Markov it gives a condition on $A(t)$ when this is indeed the case.

(2) The result as such is not new but can be found in a slightly different form in Kemeny and Snell (1960, p. 126) where Markov chains in discrete time are considered. However the proof given here is shorter.

4. $\mathbb{F}^n$-Markov processes

4.1. We will combine the results of Corollary 2.2 and Propositions 3.1, 3.2 applied to the situation where $\mathbb{F} = \mathbb{F}^n$ in order to find an integral representation of a finite state $\mathbb{F}^n$-Markov process in terms of its infinitesimal characteristics and the intensity of the counting process. Let us before $\lambda_t^+ = (1/\lambda_t)1_{\{\lambda_t>0\}}$, with the understanding that $0/0 = 0$.

Theorem 4.1. *Let X be an $\mathbb{F}^n$-Markov process with state space $\{c_1, \ldots, c_n\}$ and let Y be the indicator process associated to X as before. Then*

$$\text{(i)} \qquad Y_t = Y_0 + \int_0^t \lambda_s^+ A(s) Y_{s-} \, \mathrm{d}n_s. \tag{4.1}$$

(ii) *We have the following explicit expression for Y: If the T_k are the jumps times of n, then*

$$Y_t 1_{\{T_k \leq t < T_{k+1}\}} = \prod_{l=1}^{k} (\lambda_{T_l}^+ A(T_l) + I) Y_0 1_{\{T_k \leq t < T_{k+1}\}}.$$

Proof. (i) Y is a pure jump process satisfying $Y_t = Y_0 + \int_0^t A(s) Y_s \, \mathrm{d}s + m_t^Y$ where m^Y is an $\mathbb{F}^n$-martingale. Hence a multivariate extension of Lemma 2.1 applies, and one obtains for a certain $\mathbb{F}^n$-predictable process: $Y_t = Y_0 + \int_0^t k_s \, \mathrm{d}n_s$ and

$$\int_0^t A(s) Y_s \, \mathrm{d}s = \int_0^t k_s \lambda_s \, \mathrm{d}s.$$

Hence in order to ensure $\mathbb{F}^n$-predictability of k we have

$$A(t) Y_{t-} = k_t \lambda_t. \tag{4.2}$$

So $k_t = \lambda_t^+ A(t) Y_{t-} P(\mathrm{d}\omega)\lambda_t(\omega) \, \mathrm{d}t$ a.e. The proof of (ii) is now immediate. □

Example 4.1. Assume that the intensity process λ does not depend on t. Then $\lambda_t(\omega) = \lambda$ for some nonrandom constant λ since $\lambda_0(\cdot)$ is $\mathcal{F}_0^n$-measurable. Assume $\lambda > 0$. Assume further than X is a homogeneous Markov process. Then

$$Y_t 1_{\{T_k \leq t < T_{k+1}\}} = (\lambda^{-1} A + I)^k Y_0 1_{\{T_k \leq t < T_{k+1}\}}$$

or

$$Y_t = (\lambda^{-1} A + I)^{n_t} Y_0.$$

Since Y_t is a unit vector for all t, $\lambda^{-1} A + I$ is a semi-permutation matrix in the sense that each of its columns has exactly one $+1$ entry and the other entries are zero. Of course two $+1$ entries may occur in the same row. Consequently all the diagonal elements A_{ii} of A are either zero or equal to $-\lambda$. If some $A_{ii} = -\lambda$ then there is in the ith column A_i of A exactly one A_{ji} equal to $+\lambda$. All the other entries of A_i are zero. If $A_{ii} = 0$ for some i then the whole column $A_i = 0$.

A similar remark applies to the general expression Theorem 4.1(ii). We have for all i, $A_{ii}(T_l) \leq 0$. Then if $A_{ii}(T_l) < 0$ there is exactly one $j = j(i, T_l)$ such that $A_{ji}(T_l) = -A_{ii}(T_l)$. Since T_l can assume any value >0, we have that for each i and t there is exactly one $j = j(i, t)$ such that $A_{ji}(t) = -A_{ii}(t)$, all the other entries in the column $A_i(t)$ being zero.

From these considerations or directly by inspection of k_t we get the following. Suppose we have an $\mathbb{F}$-Markov process X with states $\{1, \ldots, p\}$. Then it can always be represented in the following way.

Consider p measurable functions $f_j : [0, \infty) \to \{1, \ldots, p\}$. Define $E_{ij} = \{t : f_j(t) = i\}$. Observe that for all j the collection $\{E_{ij}\}_{i=1}^p$ forms a partition of $[0, \infty)$, although some of the E_{ij} may be empty. Define the matrix $M(t)$ by $M(t)_{ij} = 1_{E_{ij}}(t)$. Then we have for Y the representation (like in Walrand and Varaiya, 1980),

$$\mathrm{d} Y_t = (M(t) - I) Y_{t-} \, \mathrm{d} n_t. \tag{4.3}$$

Clearly the interpretation of $M(t)$ is that $M(t)_{ij} = 1$ (or $f_j(t) = i$) iff at time t a transition $j \to i$ is possible.

Observe however that not all processes X for which the above representation (4.3) holds are $\mathbb{F}^n$-Markov.

A necessary and sufficient condition for this to hold in view of Propositions 3.1 and 3.2 is clearly

$$(M(t) - I) Y_{t-} \lambda_t = (M(t) - I) Y_{t-} E[\lambda_t \mid Y_{t-}]$$

or equivalently:

$$\exists \alpha : [0, \infty) \to \mathbb{R}^p : (M(t) - I) Y_{t-} \lambda_t = (M(t) - I) \operatorname{diag}(\alpha(t)) Y_{t-}.$$

4.2. The objective of this subsection is to study how λ and A are related. We also show that an $\mathbb{F}^n$-Markov process X automatically becomes $\mathbb{F}^{\tilde{n}}$-Markov, where $\tilde{n}$ counts the transitions of X. Conversely if X is $\mathbb{F}^{\tilde{n}}$-Markov and if n is another counting process that satisfies the conditional independence relation of Proposition 2.2, then it turns out that X is also $\mathbb{F}^n$-Markov.

Consider the first problem and observe that equation (4.2) relates the intensity λ_t of the counting process with the matrix $A(t)$ of transition inequalities of X by means of the intermediate process k. In this subsection we will study this relation a little further.

Multiply (4.3) by Y_{t-}^{T} to obtain

$$\lambda_t Y_{t-}^{\mathrm{T}} K_t = Y_{t-}^{\mathrm{T}} A(t) Y_{t-}. \tag{4.4}$$

At a jump time T_n of the counting process there are two possibilities. If X also jumps then $Y_{T_n} \neq Y_{T_n-} = Y_{T_{n-1}}$ and $Y_{T_{n-1}}^{\mathrm{T}} k_{T_n} = Y_{T_{n-1}}^{\mathrm{T}} (Y_{T_n} - Y_{T_{n-1}}) = -1$. If X does not jump then $Y_{T_{n-1}}^{\mathrm{T}} k_{T_n} = 0$. So we get from (4.4),

$$\lambda_{T_n} 1_{\{Y_{T_n} \neq Y_{T_{n-1}}\}} = -Y_{T_{n-1}}^{\mathrm{T}} A(T_n) Y_{T_{n-1}}. \tag{4.5}$$

This last equation (4.5) suggests a connection between λ and A. This connection will be studied in the sequel. First we need a definition. Define $\tilde{n}:\Omega\times[0,\infty)\to\mathbb{R}$ by $\tilde{n}_t=\sum_{s\leq t}1_{\{Y_s\neq Y_{s-}\}}$. Then $\tilde{n}_t=\frac{1}{2}[Y^{\mathrm{T}},Y]_t$. Here $[Y^{\mathrm{T}},Y]$ is the optional quadratic variation process of Y. It satisfies

$$Y_t^{\mathrm{T}}Y_t=Y_0^{\mathrm{T}}Y_0+2\int_0^t Y_{s-}^{\mathrm{T}}\,\mathrm{d}Y_s+[Y^{\mathrm{T}},Y]_t.$$

Since $Y_t^{\mathrm{T}}Y_t=Y_0^{\mathrm{T}}Y_0=1$, we have

$$\tilde{n}=-\int_0^t Y_{s-}^{\mathrm{T}}\,\mathrm{d}Y_s. \tag{4.6}$$

We now have the following proposition.

Proposition 4.2. (i) *$\tilde{n}$ is an $\mathbb{F}^Y$-adapted (and hence $\mathbb{F}^n$-adapted, since $\mathcal{F}_t^Y\subset\mathcal{F}_t^n$) counting process with $\mathbb{F}^n$ and $\mathbb{F}^Y$-predictable intensity $\tilde{\lambda}_t=-Y_{t-}^{\mathrm{T}}A(t)Y_{t-}$.*

(ii) *$n-\tilde{n}$ is also a counting process. It is only $\mathbb{F}^n$-adapted and has $\mathbb{F}^n$-predictable intensity $\lambda_t+Y_{t-}^{\mathrm{T}}A(t)Y_{t-}$.*

(iii) *$\tilde{n}_t=\int_0^t 1_{\{\tilde{\lambda}_s>0\}}\,\mathrm{d}n_s$ and $\tilde{\lambda}_t=1_{\{\tilde{\lambda}_t>0\}}\lambda_t$.*

(iii′) *$\mathcal{F}_t^n$ and $\mathcal{F}_\infty^n$ are conditionally independent given $\mathcal{F}_t^{\tilde{n}}$.*

(iv) *Let all the $A_{ii}(t)$ be strictly negative. Then $n=\tilde{n}$, $\mathcal{F}_t^Y=\mathcal{F}_t^n$ for all $t>0$ and $\lambda_t=-Y_{t-}^{\mathrm{T}}A(t)Y_{t-}$ and*

$$k_t=-(Y_{t-}^{\mathrm{T}}A(t)Y_{t-})^{-1}A(t)Y_{t-}=-\sum\frac{A_i(t)}{A_{ii}(t)}Y_{it-}.$$

Proof. (i) In view of equation (4.6) we have $\mathrm{d}\tilde{n}_t=-Y_t^{\mathrm{T}}A(t)Y_t\,\mathrm{d}t-Y_{t-}^{\mathrm{T}}\,\mathrm{d}m_t^Y$. By observing that $\int_0 Y_{u-}^T\,\mathrm{d}m_u^Y$ is again an $\mathbb{F}^n$ and $\mathbb{F}^Y$ martingale we get the desired result according to the definition of intensity.

(ii) Follows from (i).

(iii) Notice that $1_{\{Y_{T_k}\neq Y_{T_{k-1}}\}}=1_{\{\lambda_{T_k}>0,\tilde{\lambda}_{T_k}>0\}}=1_{\{\tilde{\lambda}_{T_k}>0\}}$, since $\tilde{\lambda}_{T_k}\leq\lambda_{T_k}$. Hence $\mathrm{d}\tilde{n}_t=1_{\{\tilde{\lambda}_t>0\}}\,\mathrm{d}n_t$. But then $\mathrm{d}\tilde{n}_t=1_{\{\tilde{\lambda}_t>0\}}\lambda_t\,\mathrm{d}t+1_{\{\tilde{\lambda}_t>0\}}\,\mathrm{d}m_t$, which shows that $1_{\{\tilde{\lambda}_t>0\}}\lambda_t$ is the $\mathbb{F}^n$-intensity of $\tilde{n}$ which is then also equal to λ_t by part (i).

(iii′) This is an alternative formulation of (iii) in view of Proposition 2.2.

(iv) From equation (4.4) we have $\lambda_{T_n}Y_{T_{n-1}}^{\mathrm{T}}k_{T_n}=Y_{T_{n-1}}^{\mathrm{T}}A(T_n)Y_{T_{n-1}}=\sum_i A_{ii}(T_n)1_{\{X_{T_{n-1}}=e_i\}}<0$. Hence $\lambda_{T_n}>0$ and $k_{T_n}\neq0$, which means that X always jumps as soon as n jumps. Hence $n=\tilde{n}$. Since always $\mathcal{F}_t^{\tilde{n}}\subset\mathcal{F}_t^Y\subset\mathcal{F}_t^n$ we now also have $\mathcal{F}_t^n=\mathcal{F}_t^Y$. Finally $n=\tilde{n}$ implies $\lambda_t=\tilde{\lambda}_t=-Y_{t-}^{\mathrm{T}}A(t)Y_{t-}$. Hence the expression for k_t follows from formula (4.1). □

It is appropriate to inspect the results of Proposition 4.2 a little closer. In general we have for all $t\geq0$ $\mathcal{F}_t^{\tilde{n}}\subset\mathcal{F}_t^Y\subset\mathcal{F}_t^n$. In the case described in Proposition 4.2(iv), we get equality of those σ-algebra's. Since now n is also the total number of jumps (or transitions) of the Markov chain and $\mathcal{F}_t^Y=\mathcal{F}_t^n$ it seems logical to expect that we have in the general situation (where $\tilde{n}$ counts the transitions of the chain) $\mathcal{F}_t^{\tilde{n}} = \mathcal{F}_t^Y$, which means in words that if we have a Markov chain adapted to a counting process then it is also adapted to the counting process that describes the total number of transitions of the chain. One could say that $\tilde{n}$ is sort of "minimal" counting processes to which X is adapted.

Next we show that the claim $\mathcal{F}_t^{\tilde{n}} = \mathcal{F}_t^Y$ holds true. It is a consequence of the following theorem.

Theroem 4.3. *Let X be finite state $\mathbb{F}^n$-Markov, then Y_t is $\mathcal{F}_t^{\tilde{n}}$-measurable, and hence X is finite state $\mathbb{F}^{\tilde{n}}$-Markov.*

Proof. Let $\tilde{T}_1, \tilde{T}_2, \ldots,$ be the possibly finite sequence of jump times of $\tilde{n}$. From the discussion leading to (4.5) we see that $\lambda_{\tilde{T}_j} = -Y^{\mathrm{T}}_{\tilde{T}_j-1} A(\tilde{T}_j) Y_{\tilde{T}_{j-1}} > 0$. Consider first $\tilde{T}_1$. Then $\lambda_{\tilde{T}_1}$ is a (measurable) function of $\tilde{T}_1$ only. Hence from $Y_{\tilde{T}_1} = (\lambda^{-1}_{\tilde{T}_1} A(\tilde{T}_1) + I) Y_0$, the random variable $Y_{\tilde{T}_1}$ is also a measurable function of $\tilde{T}_1$ only. But then by induction we find that $Y_{\tilde{T}_n} = (\lambda^{-1}_{\tilde{T}_n} A(\tilde{T}_n) + I) Y_{\tilde{T}_{n-1}}$ is a measurable function of $\tilde{T}_1, \ldots, \tilde{T}_n$, say $Y_{\tilde{T}_n} = y_n(\tilde{T}_1, \ldots, \tilde{T}_n)$.

Consequently, by right continuity of Y, we get

$$Y_t = Y_0 + \sum_{n=1}^{\infty} y_n(\tilde{T}_1, \ldots, \tilde{T}_n) 1_{\{\tilde{T}_n \leq t < \tilde{T}_{n+1}\}}.$$

Notice that y_n is an $\mathcal{F}^{\tilde{n}}_{\tilde{T}_n}$-measurable function since $\mathcal{F}^{\tilde{n}}_{\tilde{T}_n} = \sigma(\tilde{T}_1, \ldots, \tilde{T}_n)$. Now we invoke the fact that $\mathcal{F}^{\tilde{n}}_{\tilde{T}_n} \cap \{\tilde{T}_n \leq t < \tilde{T}_{n+1}\} = \mathcal{F}^{\tilde{n}}_t \cap \{\tilde{T}_n \leq t < \tilde{T}_{n+1}\}$ (see Brémaud, 1981, p. 308) to see that indeed Y_t is $\mathcal{F}^{\tilde{n}}_t$ measurable. Since a process that is Markov with respect to some filtration is also Markov with respect to any other smaller filtration to which it is adapted, X is also $\mathbb{F}^{\tilde{n}}$-Markov. $\square$

The statement of the theorem is sometimes immediately seen in specific cases. Consider for example the case where $\lambda_t \equiv \lambda > 0$ and A is a constant matrix (example 4.1). Then we have in fact $Y_t = (\lambda^{-1}A + I)^{\tilde{n}_t} Y_0$.

So far we have seen the following results. Given the fact that we have an $\mathbb{F}^n$-Markov process X, X is also $\mathbb{F}^{\tilde{n}}$-Markov and $\tilde{n}$ has intensity $\tilde{\lambda}_t = -Y^{\mathrm{T}}_{t-} A(t) Y_{t-}$, where $\tilde{n}$ is as before the process that counts all the transitions of X. As such these results form necessary conditions that follow from the existence of such processes. One might raise the question how to formulate sufficient conditions on a given Markov matrix function $A(\cdot)$ such that there exists an associated $\mathbb{F}^n$-Markov chain X.

Secondly, given that a process X is $\mathbb{F}^{\tilde{n}}$-Markov, what other counting processes n do exist such that X is also $\mathbb{F}^n$-Markov.

Answering the first question will be postponed until Section 5. Concerning the second one we have—as a converse of Proposition 4.2(iii)—the following proposition.

Proposition 4.4. *Let X be $\mathbb{F}^{\tilde{n}}$-Markov. Let n be another counting process with $\mathscr{F}^n$-predictable intensity λ such that*:

(i) $\tilde{n}_t = \int_0^t 1_{\{\tilde{\lambda}_s > 0\}} \, \mathrm{d}n_s.$

(ii) $\tilde{\lambda}_t = 1_{\{\tilde{\lambda}_t > 0\}} \lambda_t.$

Then X is also $\mathbb{F}^n$-Markov.

Proof. From Proposition 2.2, we see that $\mathscr{F}_t^{\tilde{n}} \subset \mathscr{F}_t^n$ and that $\mathscr{F}_\infty^{\tilde{n}}$ and $\mathscr{F}_t^n$ are conditionally independent given $\mathscr{F}_t^{\tilde{n}}$. Hence X is certainly $\mathbb{F}^n$-adapted.

Observe first that $\tilde{\lambda}_t = 0 \Leftrightarrow Y_{t-}^{\mathrm{T}} A(t) Y_{t-} = 0$ implies $A(t) Y_{t-} = 0$ as a result of the fact that $A(t)$ is a Markov-matrix. Since X is $\mathbb{F}^{\tilde{n}}$-Markov: $\mathrm{d}Y_t = \tilde{\lambda}_t^+ A(t) Y_{t-} \, \mathrm{d}\tilde{n}_t$ (Theorem 4.1). Hence

$$\mathrm{d}Y_t = \tilde{\lambda}_t^+ \lambda_t A(t) Y_t \, \mathrm{d}t + \tilde{\lambda}_t^+ A(t) Y_{t-} \, \mathrm{d}m_t = A(t) Y_t \, \mathrm{d}t + \tilde{\lambda}_t^+ A(t) Y_{t-} \, \mathrm{d}\tilde{m}_t.$$

From the conditional independence relation and Lemma 2.3, the last term is an $\mathbb{F}^n$-martingale. Therefore application of Proposition 3.2 completes the proof. □

Remark. In view of the remark following the proof of Proposition 2.2 one can replace conditions (i) and (ii) in Proposition 4.4 by $\tilde{n}_t = \int_0^t u_s \, \mathrm{d}n_s$ and $\tilde{\lambda}_t = u_t \lambda_t$ for some $\mathbb{F}^n$-predictable process u.

Until now we have studied processes X that are $\mathbb{F}^n$-Markov and thus $\mathbb{F}^n$-adapted. As mentioned before, one of the results is then, that X is also $\mathbb{F}^{\tilde{n}}$- Markov. Knowing this, one can prove all the results mentioned in the foregoing, such as $\tilde{\lambda}_t = Y_{t-}^{\mathrm{T}} A(t) Y_{t-}$ etc.

An interesting question is to see whether a process which is Markov with respect to its own flow of σ-algebras and which is $\mathbb{F}^n$-adapted, shares the same properties. In general this is not true. For instance if n is standard poisson process and X is defined by $X_t = n_{t/2}$, then X is $\mathbb{F}^X$-Markov, but not $\mathbb{F}^n$-Markov. Theorem 4.5 gives a sufficient condition for an affirmative answer. Let us first remark that any bounded process that is a semimartingale with respect to some filtration is special. (See Dellacherie and Meyer, 1980, Théorème VII.25.)

Theorem 4.5. *Let X be a finite state $\mathbb{F}^X$-Markov chain and assume that X is adapted to $\mathbb{F}^n$ for some counting process n. Assume moreover that the indicator process Y, being an $\mathbb{F}^n$-special semimartingale, admits a decomposition such that the predictable process of finite variation is continuous. Then $\mathscr{F}_t^X = \mathscr{F}_t^{\tilde{n}}$ $\forall t \geq 0$ and X is $\mathbb{F}^{\tilde{n}}$-adapted and thus $\mathbb{F}^{\tilde{n}}$-Markov.*

Proof. From Lemma 2.1 we get $\mathrm{d}Y_t = k_t \, \mathrm{d}n_t$ for some $\mathbb{F}^n$-predictable process k. By definition of $\tilde{n}$ we have $\mathrm{d}\tilde{n}_t = \frac{1}{2} \mathrm{d}[Y^{\mathrm{T}}, Y]_t = \frac{1}{2} k_t^{\mathrm{T}} k_t \, \mathrm{d}n_t$. So $\Delta\tilde{n}_t = 0$ iff $k_t = 0$. Therefore we can write $\mathrm{d}Y_t = k_t \, \mathrm{d}\tilde{n}_t$. Observe that $\tilde{n}$ is $\mathbb{F}^Y$-adapted. As in Brémaud (1981, p.

2.13), we can interpret k_t as a Radon–Nikodym derivative $\mathrm{d}Y_t/\mathrm{d}\tilde{n}_t$ on the $\mathbb{F}^Y$-predictable sets. Therefore we may take k to be $\mathbb{F}^Y$-predictable. For $\tilde{n}$ we have by its definition

$$\mathrm{d}\tilde{n}_t = -Y_{t-}^{\mathrm{T}}\,\mathrm{d}Y_t = -Y_{t-}^{\mathrm{T}}A(t)Y_{t-}\,\mathrm{d}t - Y_{t-}\,\mathrm{d}m_t^Y$$

so

$$\mathrm{d}Y_t = k_t\,\mathrm{d}\tilde{n}_t = -k_tY_{t-}^{\mathrm{T}}A(t)Y_{t-}\,\mathrm{d}t - k_tY_{t-}^{\mathrm{T}}\,\mathrm{d}m_t^Y \tag{4.7}$$

but on the other hand,

$$\mathrm{d}Y_t = A(t)Y_t\,\mathrm{d}t + \mathrm{d}m_t^Y. \tag{4.8}$$

Since all processes in (4.7) and (4.8) are $\mathbb{F}^Y$-adapted, we have from the uniqueness of the decomposition of a special semimartingale that $-k_tY_{t-}^{\mathrm{T}}A(t)Y_{t-} = A(t)Y_{t-}$ a.s., which then leads to $k_t = -(Y_{t-}^{\mathrm{T}}A(t)Y_{t-})^{+}A(t)Y_{t-}$. As in the proof of the Theorem 4.3 we can conclude that Y is $\mathbb{F}^{\tilde{n}}$-measurable. Therefore $\mathcal{F}_t^Y \subset \mathcal{F}_t^{\tilde{n}} \subset \mathcal{F}_t^Y$. Hence X is $\mathbb{F}^X$-Markov is now equivalent to X is $\mathbb{F}^Y = \mathbb{F}^{\tilde{n}}$-Markov. □

Remark. The statement of Theorem 4.5 indicates why $n_{t/2}$ cannot be $\mathbb{F}^n$-Markov. This is immediately seen by noting that $n_{t/2}$ is $\mathbb{F}^n$-predictable. Hence its dual predictable projection with respect to $\mathbb{F}^n$ is the process itself, which is discontinuous.

4.3. In this subsection we mention some consequences of the foregoing for the case where X is a homogeneous chain. Some of the results can also be derived from Davis and Varaiya (1974).

Corollary 4.6. *Assume that X is a homogeneous chain.*

(i) *If $A_{ii} < 0$, then in the corresponding column A_i of A there is exactly one $j = j(i)$ such that $A_{ji} = -A_{ii}$ and all other A_{ki}'s are zero. If $A_{ii} = 0$ then the whole column $A_i = 0$.*

(ii) *k is now a left continuous piecewise constant process and satisfies*

$$k_t 1_{\{T_n \le t \le T_{n+1}\}} = -\sum_i A_{ii}^{+} A_i 1_{\{X_{T_n} = e_i\}} 1_{\{T_n < t \le T_{n+1}\}}.$$

(iii) *The sampled chain $\hat{X}_n := X_{T_n}$ is now a deterministic process and completely known given the initial state $\hat{X}_0 = X_0$.*

(iv) *If there are no absorbing states, then the process λ assumes only a finite number of values. Specifically $\lambda_t \in \{-A_{11}, \ldots, -A_{nn}\}$.*

Proof. (i), (iii) and (iv) follow immediately from the explicit expression in Corollary 4.2 (ii) requires a little work. Recall that we have $k_t = \lambda_t^{+}AY_{t-}$. Let T be the absorption time of the chain. Then $AY_{t-}1_{\{t>T\}} \equiv 0$. Hence $\tilde{\lambda}_t > 0 \Leftrightarrow t \le T$. Therefore $\lambda_t 1_{\{t\le T\}} = \tilde{\lambda}_t 1_{\{t \le T\}} = -Y_{t-}^{\mathrm{T}}AY_{t-}1_{\{t\le T\}}$. Hence $k_t = -\sum_i A_{ii}^{+}A_iY_{it-}1_{\{t\le T\}} = -\sum_i A_{ii}^{+}A_iY_{it-}$, because $A_iY_{it-}1_{\{t>T\}} = 0$. □

At this point one might raise the question in virtue of Corollary 4.6(iv) whether λ is also a Markov process. Clearly this is the case if all the A_{ii} are different or when they are all the same. Interesting is the case when there exists at least one pair (i, j) such that $A_{ii} = A_{jj}$. We will answer this question by means of Theorem 3.4. Assume that there are $2 \leq m \leq n-1$ distinct values among the A_{ii}. Call these $a_1, \ldots, a_m$ and denote for all $j = 1, \ldots, m$ by E_i the set of all j such that $A_{jj} = a_i$. Define $F \in \mathbb{R}^{m \times n}$ by $F_{ij} = 1_{\{j \in E_i\}}$. We have the following result in the terminology of Theorem 3.4.

The process λ is an $\mathbb{F}^n$-Markov chain iff $FAK = 0$. If the last condition is satisfied then the matrix B of transition intensities of λ is given by $FA\hat{F}$.

Example 4.2. (i) If

$$A = \begin{bmatrix} -a & 0 & 0 & b \\ a & -b & 0 & 0 \\ 0 & b & -a & 0 \\ 0 & 0 & a & -b \end{bmatrix},$$

then λ is Markov with

$$B = \begin{bmatrix} -a & b \\ a & -b \end{bmatrix}$$

and state space $\{-a, -b\}$. Here we should take

$$F = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}.$$

(ii) If

$$A = \begin{bmatrix} -a & 0 & 0 & b \\ a & -a & 0 & 0 \\ 0 & a & -b & 0 \\ 0 & 0 & b & -b \end{bmatrix},$$

then λ is not Markov, which is seen by calculating

$$FAK = \begin{bmatrix} a & -b \\ -a & b \end{bmatrix},$$

with

$$F = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} \quad \text{and} \quad K^{\mathrm{T}} = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}.$$

Remarks. (i) Although it might happen that λ is not Markov of course (λ, X_-) is jointly Markov.

(ii) Since it follows from Corollary 4.6(iv) that the number of values that λ can assume is always at most the number of states that X can assume, we see that a necessary condition for a process X to be Markov is, that it takes values in a set which is at least as big as the set of values of the process λ: So $n \geqslant \#\{\lambda_t : t \geqslant 0\}$. Hence a homogeneous chain X cannot have a finite state space if λ has a continuously varying component. In the same way as checking, whether λ is $\mathbb{F}^n$-Markov one can investigate whether there exist Markov processes X^1 with a smaller state space than X by considering all possible choices of F. Thus obtaining a description of a "minimal" Markov process. This is of some relevance in connection with the stochastic realization problem to be posed in Section 5.

(iii) The case where λ is $\mathbb{F}^n$-Markov itself implies here that it changes value as soon as n jumps. Thus we can immediately see from the A-matrix whether λ is $\mathbb{F}^n$-Markov or not. In Example 4.2(i) we see that at jump times λ switches from a to b or conversely which is an agreement with the fact that it is Markov. In Example 4.2(ii) we see that it is possible that λ stays in a even when n jumps.

4.4. In the previous subsection we have seen that the existence of a homogeneous $\mathbb{F}^n$-Markov chain X does not necessarily imply that λ is also $\mathbb{F}^n$-Markov. Here after we describe some consequences of the situation where indeed λ is an $\mathbb{F}^n$-Markov process with finite state space. Since in this case λ assumes only a finite number of values it follows that λ (being predictable) may be taken as a left continuous process. Write $X_t = \lambda_{t+}$, the right continuous version of λ. We will apply the previous results to this particular choice of X.

Denote by $\{\lambda_1, \ldots, \lambda_n\}$ the state space of X. If there are no absorbing states then $A_{ii} < 0$ and we have that $\lambda_i = -A_{ii}$ for all i in view of Corollary 4.6(iv). So all $\lambda_i > 0$.

For reasons of completeness we will show what happens if some of the A_{ii} are equal to zero or if one of the λ_i equals zero. The latter case clearly implies that the corresponding $A_{ii} = 0$. Hence this case is covered by the first one. Define $B \subset \{1, \ldots, n\}$ to be the set of integers i such that λ_i is an absorbing state. Define also $T = \inf\{t \geqslant 0 : X_t \in \{\lambda_i, i \in B\}\}$.

Notice that $T < \infty$ a.s. if and only if $B = \emptyset$, and for $i \in B$ we have $A_{ii}(t) \equiv 0$, and hence the whole column $A_i(t) \equiv 0$. The principal result of this subsection is the next proposition which tells that for $t \leqslant T$ we can more or less identify the intensity λ_t as one of the $A_{ii}(t)$'s, and that $A_{ii}(t)$ only assumes the values $-\lambda_i$ or 0.

Proposition 4.7. *Assume that λ is $\mathbb{F}^n$-Markov with state space $\{\lambda_1, \ldots, \lambda_n\}$ and transition intensity matrix $A(t)$. Let T be the absorption time as defined above and B the set of integers corresponding to the absorbing states. Then*

$$\lambda_t = \lambda_T 1_{\{t > T\}} + \sum_{i \in B^c} \lambda_i 1_{\{\lambda_t = \lambda_i\}} 1_{\{A_{ii}(t) = 0\}} - \sum_{i \in B^c} A_{ii}(t) 1_{\{\lambda_t = \lambda_i\}}$$

and for $i \in B^c$: $A_{ii}(t) = -\lambda_i$ if $A_{ii}(t) < 0$.

Proof. Let $X_t = \lambda_{t+}$, then $Y_{it} = 1_{\{X_t = \lambda_i\}}$ and $Y_{it-} = 1_{\{\lambda_t = \lambda_i\}}$. In the notation that we have used previously, $\tilde{n}$ has rate $\tilde{\lambda}_t = -Y_{t-}^{\mathrm{T}} A(t) Y_{t-} = -\sum_{i \in B^c} A_{ii}(t) 1_{\{\lambda_t = \lambda_i\}} 1_{\{t \leqslant T\}}$.

Since $\tilde{\lambda}_t = 1_{\{\tilde{\lambda}_t>0\}}\lambda_t$ (Proposition 4.2(iii)), we have

$$\tilde{\lambda}_t 1_{\{t\leq T\}} = 1_{\{\tilde{\lambda}_t>0\}}1_{\{t\leq T\}}\lambda_t + 1_{\{\tilde{\lambda}_t>0\}}1_{\{t>T\}}\lambda_t = 1_{\{\tilde{\lambda}_t>0\}}\lambda_t,$$

since $\tilde{\lambda}_t > 0$ implies $t \leq T$ and conversely $t > T$ implies $\tilde{\lambda}_t = 0$. Hence

$$-\sum_i A_{ii}(t)1_{\{\lambda_t=\lambda_i\}}1_{\{t\leq T\}} = 1_{\{\tilde{\lambda}_t>0\}}\sum_i \lambda_i 1_{\{\lambda_t=\lambda_i\}}.$$

Now let $i \in B^{\mathrm{c}}$. Then

$$-A_{ii}(t)1_{\{\lambda_t=\lambda_i\}}1_{\{t\leq T\}} = 1_{\{\tilde{\lambda}_t>0\}}1_{\{\lambda_t=\lambda_i\}}\lambda_i.$$

Observe that

$$1_{\{\lambda_t>0\}}1_{\{\lambda_t=\lambda_i\}} = 1_{\{A_{ii}(t)<0\}}1_{\{\lambda_t=\lambda_i\}}$$

and for $i \in B^{\mathrm{c}}\lambda_t = \lambda_i$ implies $t \leq T$. Hence we get

$$-A_{ii}(t)1_{\{\lambda_t=\lambda_i\}} = 1_{\{A_{ii}(t)<0\}}1_{\{\lambda_t=\lambda_i\}}\lambda_i.$$

Since we may assume that $P(\lambda_t=\lambda_i)>0$ we now get by taking expectations

$$-A_{ii}(t) = 1_{\{A_{ii}(t)<0\}}\lambda_i$$

which proves the second assertion of the proposition. Furthermore

$$\begin{aligned}
\lambda_t &= \lambda_T 1_{\{t>T\}} + \lambda_t 1_{\{t\leq T\}} \\
&= \lambda_T 1_{\{t>T\}} + 1_{\{t\leq T\}} \sum_{i\in B^{\mathrm{c}}} 1_{\{\lambda_t=\lambda_i\}}\lambda_i \\
&= \lambda_T 1_{\{t\geq T\}} + \sum_{i\in B^{\mathrm{c}}} 1_{\{\lambda_t=\lambda_i, A_{ii}(t)=0\}}\lambda_i + \sum_{i\in B^{\mathrm{c}}} 1_{\{\lambda_t=\lambda_i, A_{ii}(t)<0\}}\lambda_i \\
&= \lambda_T 1_{\{t>T\}} + \sum_{i\in B^{\mathrm{c}}} 1_{\{\lambda_t=\lambda_i, A_{ii}(t)=0\}}\lambda_i - \sum_{i\in B^{\mathrm{c}}} A_{ii}(t)1_{\{\lambda_t=\lambda_i\}}
\end{aligned}$$

which proves the first assertion. □

Remarks. (i) If λ is a homogeneous $\mathbb{F}^n$-Markov chain, then A is a constant matrix and we have for $i \in B^{\mathrm{c}}$ the identity $A_{ii}(t) \equiv -\lambda_i$. Hence

$$\lambda_t = \lambda_T 1_{\{t>T\}} - \sum_{i\in B^{\mathrm{c}}} A_{ii}1_{\{\lambda_t=\lambda_i\}}.$$

And of course if there are no absorbing states or if the value zero is the only one, then $A_{ii}(t) \equiv -\lambda_i$ for all i and $\lambda_t = -\sum_{i=1}^n A_{ii}1_{\{\lambda_t=\lambda_i\}}$.

(ii) Now it is easy to see that for any function f which is not injective or constant $f(\lambda)$ cannot be a $\mathbb{F}^n$-Markov chain, since we have tacitly assumed that all the λ_i are different. Hence the number of states of λ is now the minimal number of elements that a set should have in order that it can serve as a state space for some $\mathbb{F}^n$-Markov process. In this sense one can say that λ, if it is $\mathbb{F}^n$-Markov, is the minimal $\mathbb{F}^n$-Markov chain.

5. Stochastic realization

The purpose of this section is to solve a certain stochastic realization problem, to be stated in Subsection 5.2. The solution involves a technical result on the existence of $\mathbb{F}^n$-Markov process which is formulated in Subsection 5.1.

5.1. It is known that given a Markov-matrix function $A:[0,\infty)\to\mathbb{R}^{n\times n}$, one can always construct a probability space $(\Omega, \mathscr{F}, P)$ and a Markov process $X:\Omega\times[0,\infty)\to\{1,\dots,n\}$, such that its transition probabilities are generated by A.

In this section we are concerned with a restrictive version of this problem, namely given a complete probability space $(\Omega, \mathscr{F}, P)$ a counting process $n:\Omega\times[0,\infty)\to\mathbb{N}_0$ and a Markov matrix function $A:[0,\infty)\to\mathbb{R}^{n\times n}$, does there exist an $\mathbb{F}^n$-Markov process $X:\Omega=[0,\infty)\to\{1,\dots,n\}$ such that A generates its transition probabilities. We know from previous results that given such a process we have the identities $\tilde\lambda_t=-Y_{t-}^{\mathrm{T}}A(t)Y_{t-}$ and $\tilde\lambda_t=\lambda_t 1_{\{\tilde\lambda_t>0\}}$ and that for each (i,t) such that $A_{ii}(t)<0$, there exists only one j such that $A_{ji}(t)=-A_{ii}(t)$. Hence for the existence of such a process X this imposes some necessary conditions on the matrix $A(t)$. In Theorem 5.1 we present a set of sufficient conditions on both $A(t)$ and λ_t that implies the existence of such a desired process X, and we also give a construction for X. Before stating the theorem let us emphasize that one should not overestimate its content, since in a sense it looks like a tautology. On the other hand it shows how one can extract an $\mathbb{F}^n$-Markov process that is hidden in a suitable matrix function A. After having proved the theorem we give an example, how to use the construction of X.

Theorem 5.1. *Given a counting process n with $\mathbb{F}^n$-predictable intensity λ and a Markov matrix function $A:[0,\infty)\to\mathbb{R}^{n\times n}$. There exists an $\mathbb{F}^n$-Markov process $X:\Omega\times[0,\infty)\to\{1,\dots,n\}$ with A as its infinitesimal generator if there is a unique sequence of random variables $\{x_m\}_{m\geq 0}$, $x_m:\Omega\to\{1,\dots,n\}$ such that the following two conditions hold:*

(i) $A_{x_mx_m}(T_m)(A_{x_mx_m}(T_m)+\lambda_{T_m}I)=0\ \forall_m$.

(ii) *If* $A_{x_mx_m}(T_m)<0$ *then* x_{m+1} *is such that* $A_{x_{m+1}x_m}(T_m)=-A_{x_mx_m}(T_m)$ *and if* $A_{x_mx_m}(T_m)=0$, *then* $x_{m+1}=x_m$.

Proof. Let us define a process $Y^-:\Omega\times[0,\infty)\to\{0,1\}^n$ by requiring that $Y_t^-1_{\{T_{m-1}<t\leq T_m\}}=Y_{T_m}^-1_{\{T_{m-1}<t\leq T_m\}}$ and $Y_{iT_m}^-=1_{\{x_m=i\}}$. Then

$$\begin{aligned}
\lambda_{T_m}^+\sum_j A_{ij}(T_m)Y_{jT_m}^- &= \lambda_{T_m}^+\sum_j A_{ij}(T_m)1_{\{x_m=j\}}\\
&=\lambda_{T_m}^+A_{ix_m}(T_m)\\
&=\lambda_{T_m}^+A_{x_{m+1}x_m}(T_m)1_{\{x_{m+1}=i\}}+\lambda_{T_m}^+A_{x_mx_m}(T_m)1_{\{x_m=i\}}\\
&\quad+\lambda_{T_m}^+A_{ix_m}(T_m)1_{\{x_m\neq i,x_{m+1}\neq i\}}\\
&=-\lambda_{T_m}^+A_{x_mx_m}(T_m)1_{\{x_{m+1}=i\}}1_{\{A_{x_mx_m}(T_m)<0\}}\\
&\quad+\lambda_{T_m}^+A_{x_mx_m}(T_m)1_{\{x_m=i\}}+0\\
&=-\lambda_{T_m}^+A_{x_mx_m}(T_m)[1_{\{x_{m+1}=i\}}-1_{\{x_m=i\}}]\\
&=1_{\{x_{m+1}=i\}}-1_{\{x_m=i\}}=Y_{iT_{m+1}}^--Y_{iT_m}^-.
\end{aligned}$$

So in vector notation we have

$$Y^-_{T_{m+1}} - Y^-_{T_m} = \lambda^+_{T_m} A(T_m) Y^-_{T_m}. \tag{5.1}$$

Notice that $\lambda_{T_m} = 0$ implies $A(T_m) Y^-_{T_m} = 0$. Therefore with the usual convention that $0/0 = 0$ we have from (5.1),

$$Y^-_{T_{m+1}} - Y^-_{T_m} = \lambda^{-1}_{T_m} A(T_m) Y^-_{T_m}. \tag{5.2}$$

Define now $Y : \Omega \times [0, \infty) \to \{0, 1\}^n$ by $Y_t = Y^-_{t+}$. Then $Y^-_{T_{m+1}} = Y_{T_m}$. Hence (5.1) reads

$$Y_{T_m} - Y_{T_{m-1}} = \lambda^{-1}_{T_m} A(T_m) Y_{T_{m-1}} \tag{5.3}$$

which can be rephrased as

$$\mathrm{d} Y_t = \lambda_t^{-1} A(t) Y_{t-} \, \mathrm{d} n_t \tag{5.4}$$

or

$$\mathrm{d} Y_t = A(t) Y_{t-} \, \mathrm{d} t + \lambda_t^+ A(t) Y_{t-} \, \mathrm{d} m_t.$$

We now want to apply Proposition 3.2. Therefore we have to verify that Y_{t-} is $\mathbb{F}^n$-predictable. Observe that

$$Y_{t-} 1_{\{T_m < t \leq T_{m+1}\}} = Y_{T_m} 1_{\{T_m < t \leq T_{m+1}\}}. \tag{5.5}$$

Now the sequence $\{x_m\}_{m \geq 0}$ is such that x_{m+1} is selected on the basis of knowing x_m and T_m, or iteratively is selected on the knowledge of $\{T_1, \ldots, T_m\}$. Therefore $Y_{iT_m} = Y^-_{iT_{m+1}} = 1_{\{x_{m+1} = i\}}$ only depends on $\{T_1, \ldots, T_m\}$. From (5.5) and Lemma 2.3 we now find the desired result. □

Example 5.1. Let λ be constant between the jump times T_i and envolve according to $\lambda_1, \lambda_2, \lambda_3, \lambda_1, \lambda_2, \lambda_3, \ldots$, etc. Let

$$A_1 = \begin{bmatrix} -\lambda_1 & 0 & \lambda_2 \\ \lambda_1 & -\lambda_3 & 0 \\ 0 & \lambda_3 & -\lambda_2 \end{bmatrix}.$$

Then we see that A_1 cannot be a transition matrix of an $\mathbb{F}^n$-Markov chain $X : \Omega \times [0, \infty) \to \{1, 2, 3\}$. Because from condition (i) of Theorem 5.1 we see that $X_t = 1$ iff $\lambda_t = \lambda_1$, $X_t = 3$ iff $\lambda_t = \lambda_1$ and $X_t = 2$ iff $\lambda_t = \lambda_3$. From $X_t = 1$ it can only jump to 2 according to A_1. But from the given sequence of λ's it should jump from 1 to 3. However

$$A_2 = \begin{bmatrix} -\lambda_1 & 0 & \lambda_3 \\ \lambda_1 & -\lambda_2 & 0 \\ 0 & \lambda_2 & -\lambda_3 \end{bmatrix}$$

is compatible with the sequence of λ's as one can easily verify and thus A_2 can act as the transition matrix of an $\mathbb{F}^n$-Markov chain $X : \Omega \times [0, \infty) \to \{1, 2, 3\}$.

5.2. In this section we will address a certain stochastic realization problem, and see how we can solve it by means of Theorem 5.1. Let us state the problem precisely.

We are given a complete filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}^n, P)$, where the filtration $\mathbb{F}^n$ is generated by a counting process satisfying $dn_t = \lambda_t\, dt + dm_t$, where λ is the $\mathbb{F}^n$-predictable intensity process and m an $\mathbb{F}^n$-martingale.

We pose the following question. Does there exist a homogeneous $\mathbb{F}^n$-Markov process X with finite state space E and a (measurable) function $f: E \to \mathbb{R}^+$ such that $\lambda_t = f(X_{t-})$?

One can reformulate this question in terms that are used in stochastic realization theory. The concepts involved are then stochastic system, state process, output process. However it seems that there is no consensus on how to define in abstract terms, what a stochastic system is. One approach can be found in Van Schuppen (1979). We will not touch upon all the difficulties that are inherent to this problem.

We will give a definition that suffices for our purpose. Suppose that we are given an object, to be called a stochastic system, with output process y. Then from the intuitive interpretation of state a process X that should play the role of state process has to satisfy at least the following requirement: the conditional distribution of X_{t+v}, $v > 0$ given all X_s and y_s for $s \leq t$ is the same as the conditional distribution of X_{t+v} given X_t alone.

Here we are interested in systems with a counting process output only. The above considerations are captured in the next definition, which is probably not the most general one.

Definition 5.2. A stochastic state space system with counting process output is a complete filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$ together with an adapted stochastic process $X: \Omega \times [0, \infty) \to E$, an adapted counting processes n and a measurable function $f: \mathbb{R} \times E \to \mathbb{R}_+$ such that X is $\mathbb{F}$-Markov and $f(t, X_{t-})$ is the $\mathbb{F}$-predictable intensity of n.

In this section we are concerned with state processes X that assume finitely many values and with self-exciting counting processes n, like in previous sections. This amounts to studying stochastic systems such that X is $\mathbb{F}^n$-adapted. By taking $\mathbb{F} = \mathbb{F}^n$ in Definition 5.2, we have that the state X is even $\mathbb{F}^n$-Markov. As already mentioned in the introduction, the property that the state process is $\mathbb{F}^n$-Markov was a motivation for studying $\mathbb{F}^n$-Markov processes.

Remark. Observe that we can take n as a state process if and only if n is Markov, which is the case if and only if its predictable intensity is of the form $f(t, n_{t-})$ (Proposition 3.3).

An alternative formulation of the question that we posed in the beginning of this section is the following. Given a counting process n on $(\Omega, \mathcal{F}, P)$ can we find a stochastic system on $(\Omega, \mathcal{F}, \mathbb{F}^n, P)$ such that its state process X is homogeneous and has finite state space E and such that the output processes is n with $\mathbb{F}^n$-predictable intensity $f(X_{t-})$ for some $f: E \to \mathbb{R}_+$.

Let us suppose that we can affirmatively answer this question. From Corollary 4.6 we see that the sequence $\{\lambda_{T_m}\}$ is eventually constant or periodic. This observation also gives us a sufficient condition for solving the problem, which is the content of the next theorem.

Theorem 5.3. *There exists on $(\Omega, \mathscr{F}, \mathbb{F}^n, P)$ a finite state $\mathbb{F}^n$-Markov process X with state space E and a function $f: E \to \mathbb{R}^+$ such that $\lambda_t = f(X_{t-})$ if and only if there exist a $k \in \mathbb{N}$ such that the sequence $\{\lambda_{T_n}\}$ for $n \geqslant k$ is either constant or periodic.*

Proof. We only have to prove that this condition on λ is sufficient for the existence of X.

(i) Consider first the case where $\{\lambda_{T_N}\}$ is eventually cyclic, which means that there exist integers N' and p' such that $\lambda_{T_{i+p'}} = \lambda_{T_i}$ for $i \geqslant N'$. Let N and p be the smallest of such integers. Now we can construct an $\mathbb{F}^n$-Markov process X with state space $\{1, \ldots, N+p\}$ as follows. Define $A \in \mathbb{R}^{(N+p)\times(N+p)}$ by $A_{ii} = -\lambda_{T_{i-1}}$ for $i = 1, \ldots, N+p$, $A_{i+1,i} = -A_{ii} = \lambda_{T_{i-1}}$ for $i = 1, \ldots, N+p-1$ and $A_{N+1,N+p} = \lambda_{T_{N+p-1}}$. All other A_{ij} are zero.

$$A = \begin{bmatrix} -\lambda_0 & & & \\ \lambda_0 & & & \\ & -\lambda_{T_N} & & +\lambda_{T_{N+p-1}} \\ & +\lambda_{T_N} & & \\ & & -\lambda_{T_{N+p-2}} & \\ & & +\lambda_{T_{N+p-2}} & -\lambda_{T_{N+p-1}} \end{bmatrix}.$$

The existence of the X we are looking for is guaranteed by Theorem 5.1 (take $X_m = m$, etc.) and f is defined by $f(i) = \lambda_{T_{i-1}}$, $i = 1, \ldots, N+p$.

(ii) If λ is eventually constant, take $p = 1$ in case (i). Then $\lambda_{T_n} = 0$. $\square$

Remark. The behaviour of the system for $t \leqslant T_N$ (T_N as defined in the proof of Theorem 5.3) can be considered as the transient behaviour of the system. If one would assume that time runs from minus infinity, instead from zero, then the necessary and sufficient condition in Theorem 5.3 would read: The sequence $\{\lambda_{T_n}\}$ is either periodic or constant.

One other problem that remains to be solved is that of minimality of the solution of the realization problem. In our context minimality means minimality of the number of elements of the state space E. We have the following result.

Corollary 5.4. *The solution of the stochastic realization problem as presented in the proof of Theorem 5.3 is minimal.*

Proof. In principle one can prove the corollary by applying the $FAK = 0$ criterion of Theorem 3.4. Here we give an alternative proof. Consider first the case where

$\{\lambda_{T_n}\}$ is eventually constant. Assume that there exists a function g such that $g(X)$ is Markov and a function h such that $h(g(X_t)) = f(X_t) = \lambda_t$. Consider a state j of X, $j \leqslant N$. Then there is no $i \leqslant j-1$ such that $g(i) = g(j)$, otherwise the sequence $\{\lambda_{T_n}\}$ would reach a loop, which is forbidden by assumption. Similarly there is no $i \leqslant N$ such that $g(i) = g(N+1)$, otherwise the absorption time would be smaller than T_N, which is minimal by construction. This shows that g is injective, so that E is minimal. A similar argument applies to the other case. Assume again that there is a function g such that $g(X)$ is Markov. For the transient states we have the same argument as in case (i). For the cyclic part of the chain we have for each recurrent state j that there is by definition no transient state $i < j$ such that $g(i) = g(j)$, but also no recurrent state $i < j$ such that $g(i) = g(j)$, because that would contradict the minimality of the number (period) p. Again g is injective. $\square$

6. Conclusions

The object that we have studied in this paper was a stochastic process X that is $\mathbb{F}^n$-Markov, where $\mathbb{F}^n$ denotes the filtration that was generated by some given counting process n, and has finite state space. The additional requirement homogeneous resulted in the fact that then X has to be eventually either cyclic or constant. Consequently the idea of viewing n as the output of a stochastic system, with such a process X as state process, leads to a rather restricted class of counting processes that satisfy this requirement. This partly negative result answers a question posed in the introduction, namely whether we get an interesting class of counting processes that obeys the afore mentioned conditions.

References

R.K. Boel, Modelling, estimation and prediction for jump processes, in: Advance in Statistical Signal Processing, Vol. 1 (JAI Press, 1985).

R.K. Boel, P. Varaiya and E. Wong, Martingales on jump processes. Part I, SIAM J. Control Optim. 13 (1975) 999–1021

P. Brémaud, Point Processes and Queues (Springer, New York, 1981).

P. Brémaud and M. Yor, Changes of filtration of probability measures, Z. Wahrsch. Verw. Gebiete 45 (1978) 269–295.

M.H.A. Davis and P. Varaiya, The multiplicity of an increasing family of σ-fields, Ann. Probab. 2 (1974) 952–963.

C. Dellacherie and P.A. Meyer, Probabilités et Potential (Hermann, Paris, 1980).

M. Jacobsen, Statistical analysis of counting processes, in: Lecture Notes in Statist., Vol. 12 (Springer, New York, 1982).

J.G. Kemeny and J.L. Snell, Finite Markov Chains (Van Nostrand Reinhold, New York, 1960).

B. Kemp, An optimal monitor of the rapid-eye-movement brain state, Biol. Cybernet. 54 (1986) 133–139.

J.H. van Schuppen, Stochastic filtering theory, a discussion of concepts, methods and results, in: M. Kohlmann and W. Vogel, eds., Stochastic Control Theory and Stochastic Differential Systems, Proc. Workshop Deutsch. Forschungsgemeinsch. (Univ. Bonn, Bad Honnef, 1979) pp. 209–226.

J. Walrand and P. Varaiya, Interconnections of Markov chains and quasi-reversible queuing networks, Stochastic Process. Appl. 10 (1980) 209–219.