SCHOOL OF OPERATIONS RESEARCH
AND INDUSTRIAL ENGINEERING
COLLEGE OF ENGINEERING
CORNELL UNIVERSITY
ITHACA, NY 14853-3801

TECHNICAL REPORT NO. 1096

May 1994

Markov-Additive Processes Arising in Storage Models for Communications Systems

by

António Pacheco

MARKOV-ADDITIVE PROCESSES ARISING IN STORAGE MODELS FOR COMMUNICATIONS SYSTEMS

António Pacheco, Ph.D.

Cornell University 1994

Markov-additive processes (MAPs) are a class of Markov processes which have important applications. An MAP $(\mathbf{X}, J)$ is a Markov process whose transition probability measure is translation invariant in the *additive component* $\mathbf{X}$. We consider the case where the *Markov component* J has discrete state space. When $\mathbf{X}$ is a non-decreasing process on the nonnegative integers, $\mathbf{X}$ may be seen as a counting arrival process, and $(\mathbf{X}, J)$ as an *MAP of arrivals.* The standard example is that of different classes of arrivals into a queueing system. For MAPs of arrivals we investigate the lack of memory property, interarrival times and moments of the number of counts. We then consider transformations of the process that preserve the Markov additive property, such as linear transformations, patching of independent processes, and linear combinations. Random time transformations are also

investigated. Finally we consider secondary recordings that generate new arrival processes from the original; these include, in particular, marking, colouring and thinning. For Markov-Bernoulli recording the secondary process in each case turns out to be an MAP of arrivals. We also unify the extensive literature on univariate MAPs of arrivals with finite Markov component, which most of the time have not been connected to the theory on MAPs.

We study two storage models for communications systems for which the input and demand processes are MAPs, and the demand is satisfied if physically possible. We investigate properties of some of the functionals arising in the models. In the first model the input and demand are continuous additive functionals on a Markov chain J. Our analysis uses an embeddeding of the net input process at the epochs of transitions of J, which is a Markov random walk (MRW). The derivation of the properties of the storage level and the unsatisfied demand is based on a Wiener-Hopf factorization for this MRW. In the second storage model the input is a Markov-compound Poisson process and the demand is a Markov linear process. The study of the storage level is based on a detailed analysis of the busy period, using techniques based on infinitesimal generators. The transform of the busy period is the unique solution of a certain matrix-functional equation. Steady state results are also obtained; these are not obvious generalizations of the results for simple storage models. In particular, a generalization of the Pollaczek-Khinchin formula brings new insight.

Biographical Sketch

António Pacheco Pires was born on July 1964 in Janeiro de Baixo, a very beautiful village in the interior of Portugal, as a member of large and very friendly family. He spent his high-school years in the small city of Castelo Branco. He enjoyed very much this period during which he made very good friends. He moved to Lisboa in 1982 to start his undergraduate studies at "Faculdade de Ciências", University of Lisbon, where he graduated in 1987 with a "Licenciatura" degree in Probability and Statistics. Since then he has been a Teaching Assistant at the Departament of Mathematics of "Instituto Superior Técnico (IST)", Technical University of Lisbon, where he obtained a M.Sc. degree in Applied Probability in 1990. In 1989 he was a visiting scholar at the Department of Statistical Sciences, University College London. In 1990 he applied for several Ph.D. program in Operations Research and Statistics in the USA. After rejecting some offers, he made the wise decision to join the School of Operations Research and Industrial Engineering at Cornell University, where he completed this dissertation for his doctoral degree in 1994.

A quem eu amo.

Acknowledgements

I would like to express my sincere gratitude to the members of my special committee, Prof. N.U. Prabhu, Prof. Richard Durrett and Prof. James Renegar. Their enthusiasm in teaching and research will be a constant source of inspiration in the coming years. I feel extremely pleased, and fortunate, for having worked very closely with Uma Prabhu, my committee chairman, who made my work a very pleasant experience and has been always generous in ideas, time, stimulus and friendship. I thank Prof. Sidney Resnick and Prof. Bruce Turnbull for their assistance, and Roberto Malamut and Refik Güllü for some insightful conversations. I am grateful to my colleagues at the Departments of Mathematics of IST, who gave the best support they possible could to help me getting to and going through the Ph.D. program in Operations Research at Cornell University. In this respect, Prof. M. F. Ramalhoto, my M.Sc. thesis advisor, deserves a separate mention.

Most special thanks go to all members of my family for their support and encouragement during my life. I would like to pay tribute to my parents and my sister Adélia, who value education very much and worked very hard to give me a good one. In addition to being my first teacher and a second mother, Adélia has been a wonderful role model. For the last four years Alexandra, Cesaltina, João, and Soumodip have been particularly generous. I thank all my friends both in the School of OR&IE and outside for the very good moments we had together in the last four years. Particularly, I thank Alexandra, Carlos, Fernanda, Hassan, Helena, Inês, Jane, Juan, Marie, Maura, Refik, Roberto and Rodrigo.

Finally, thanks are also due to FLAD, INVOTAN, IST and JNICT for providing me with financial support throughout the years I spent in Ithaca.

Table of Contents

Chapter 1

Introduction

Markov-additive processes (MAPs) are a class of Markov processes which have important applications in communications systems, but in most of the literature components of models which are MAPs have not been recognized as such. An MAP $(\mathbf{X}, J)$ is a Markov process whose transition probability measure is translation invariant in the *additive component* $\mathbf{X}$. In this paper we treat the case where the *Markov component* J has discrete (finite or countable) state space E.

We study MAPs with the additive component taking values in the nonnegative (multidimensional) integers, which we denote as MAPs *of arrivals.* The standard example is that of different classes of arrivals into a queueing system. In addition, we investigate storage models for communications systems with the input and demand processes being MAPs.

In this chapter we shall introduce MAPs of arrivals and the type of storage models we consider, give a review of the related literature and an outline of the dissertation.

1.1 Markov-additive processes

We consider the vector space $\mathbb{R}^r$ with the usual componentwise addition of vectors and multiplication of a vector by a scalar, and denote the points $\mathbf{x} = (x_1, x_2, \ldots, x_r)$ of $\mathbb{R}^r$, $r \geq 1$ in boldface. We will let the index set $\mathcal{T}$ be either $\mathbb{R}_+ = [0, +\infty)$ or $\mathbb{N} = \{0, 1, 2, \ldots\}$.

Definition 1 Markov-Additive Process (MAP)

A process $(\mathbf{X}, J) = \{(\mathbf{X}, J)(t),\, t \in \mathcal{T}\}$ *on the state space* $\mathbb{R}^r \times E$ *is an MAP if*

(i) $(\mathbf{X}, J)$ *is a Markov process;*

(ii) for $s, t \in \mathcal{T}$*, the conditional distribution of* $(\mathbf{X}(s+t) - \mathbf{X}(s), J(s+t))$ *given* $(\mathbf{X}(s), J(s))$ *depends only on* $J(s)$. □

We augment E with a compactification point Δ and let $E_\Delta = E \cup \{\Delta\}$. We denote $\zeta = \inf\{t \in \mathcal{T} : J(t) = \Delta\}$ and assume that $J(t) = \Delta$ and $\mathbf{X}(t) - \mathbf{X}(\zeta) = \mathbf{0}$ almost surely (a.s.) for $t \geq \zeta$. We consider only MAPs $(\mathbf{X}, J)$ having a transition probability measure in the sense of Blumenthal ([4], Definition I.(2.1)). From Definition 1 it follows that the transition probability measure of an MAP is given by

$$
\begin{aligned}
&P\{\mathbf{X}(s+t) \in A, J(s+t) = k \mid \mathbf{X}(s) = \mathbf{x}, J(s) = j\} \\
&\qquad = P\{\mathbf{X}(s+t) - \mathbf{X}(s) \in A - \mathbf{x}, J(s+t) = k \mid J(s) = j\} \quad (1.1)
\end{aligned}
$$

for $s, t \in \mathcal{T}$, $j, k \in E_\Delta$, $\mathbf{x} \in \mathbb{R}^r$ and $A \in \mathcal{R}^r$ where $\mathcal{R}^r$ is the σ-algebra of Borel sets of $\mathbb{R}^r$. Thus an MAP is a Markov process whose transition probability measure is translation invariant in the additive component. Since $(\mathbf{X}, J)$ is Markov, it follows easily from (1.1) that J is Markov and that $\mathbf{X}$ has conditionally independent increments, given the states of J, i.e. for $0 \leq t_1 \leq \ldots \leq t_n$ $(n \geq 2)$, the increments

$$\mathbf{X}(t_1) - \mathbf{X}(0),\, \mathbf{X}(t_2) - \mathbf{X}(t_1), \ldots,\, \mathbf{X}(t_n) - \mathbf{X}(t_{n-1})$$

are conditionally independent given $J(0), J(t_1), \ldots, J(t_n)$. Since in general $\mathbf{X}$ is non-Markovian, we may therefore call J the Markov component and $\mathbf{X}$ the additive component of the MAP $(\mathbf{X}, J)$. However these properties do not lead to $(\mathbf{X}, J)$ as an MAP; a fact which has been ignored in the literature (e.g. Prabhu [31]).

We note that in the trivial case where J has only one state, the definition of an MAP says exactly that $\mathbf{X}$ is a Markov process with additive increments. Thus MAPs are an extension of processes with additive increments, but in general the additive component of an MAP does not have additive increments.

Definition 1 is in the spirit of Çinlar [7] who considers more general state space for J. A second definition of MAPs is given by Çinlar [8], who starts with the Markov component J and defines $\mathbf{X}$ as a process having properties relative to J that imply, in particular, property (ii) of Definition 1. In discrete time this amounts to viewing the additive component $\mathbf{X}$ as sums of random variables defined on a Markov chain, rather than formulating $(\mathbf{X}, J)$ as a Markov Random Walk (MRW). In this approach $\mathbf{X}$ bears a causual relationship with J, which is assumed to be more or less known at the beginning. This is the case in many applications where J represents an extraneous factor such as the environment (e.g. in some Markov-modulated queueing systems). However, in other applications, the phenomenon studied gives rise in a natural fashion to $\mathbf{X}$ and J jointly, and it is important to study the evolution of $(\mathbf{X}, J)$ as a Markov process. In such situations

Definition 1 is a natural one to use. The basic references to MAPs are still Çinlar [7, 8]. For more recent work on MAPs see Prabhu [31].

In case $\mathcal{T} = I\!N$ an MAP is said to be an MRW. Due to the discrete time feature of MRWs, their study has evolved in part independently of general MAPs. If a continuous time MAP is embedded at an appropriately selected sequence of times, the resulting process will turn out to be an MRW, which may be used to study properties of the original MAP (see chapter 4). We recall the fact that if a Markov process $(\mathbf{S}^\star, J^\star) = \{(\mathbf{S}^\star_n, J^\star_n),\ n \in I\!N\}$ on $I\!R^r \times E$ is an MRW, its transition probability measure has the property

$$P\{\mathbf{S}^\star_{m+n} \in A;\ J^\star_{m+n} = k \mid \mathbf{S}^\star_m = \mathbf{x},\ J^\star_m = j\}$$

$$= P\{\mathbf{S}^\star_{m+n} - \mathbf{S}^\star_m \in A - \mathbf{x};\ J^\star_{m+n} = k \mid J^\star_m = j\} \qquad (1.2)$$

for $m, n \in I\!N$, $j, k \in E_\Delta$, $\mathbf{x} \in I\!R^r$ and $A \in \mathcal{R}^r$. This property is the one that is commonly used to define MRWs. When the additive component $\mathbf{S}^\star$ of an MRW $(\mathbf{S}^\star, J^\star)$ takes values on $I\!R^r_+$, $(\mathbf{S}^\star, J^\star)$ is said to be a Markov renewal process (MRP). MRPs are thus the discrete versions of MAPs with the additive component taking values on $I\!R^r_+$ $(r \geq 1)$, the so called Markov subordinators.

A survey of MRWs and some new results have recently been given by Prabhu, Tang and Zhu [33]. The single major reference to MRPs is Çinlar [6]; for a recente review of the literature on MRPs see Prabhu [31]. Markov subordinators are reviewed in Prabhu and Zhu [34], where applications to queueing systems are considered.

1.2 Markov-additive processes of arrivals

A particularly important class of Markov subordinators is the class of MAPs of arrivals, i.e. MAPs $(\mathbf{X}, J)$ with the additive component taking values in the non-negative (multidimensional) integers. The increments of $\mathbf{X}$ may correspond to events, and we call $\mathbf{X}$ (the additive component of the MAP) the *arrival component*. Thus for $i = 1, 2, \ldots, r$

$$X_i(t) = \text{ total number of class } i \text{ arrivals in } (0, t].$$

In the discrete time case MAPs of arrivals are MRPs; some special cases of these have been considered in the applied literature (see Lucantoni [20] and Neuts [26] for references). We consider only continuous time-homogeneous MAPs of arrivals; the general theory of these processes follows from Neveu's [27] construction of Markov subordinators.

Since an MAP of arrivals $(\mathbf{X}, J)$ is a Markov subordinator on $\mathbb{N}^r \times E$ it is a Markov chain; moreover since $(\mathbf{X}, J)$ is time-homogeneous it is characterized by its transition rates, which are translation invariant in the arrival component $\mathbf{X}$. Thus it suffices to give for $j, k \in E$ and $\mathbf{m}, \mathbf{n} \in \mathbb{N}^r$ the transition rate from $(\mathbf{m}, j)$ to $(\mathbf{m} + \mathbf{n}, k)$, which we denote simply (since the rate does not depend on $\mathbf{m}$) by $\lambda_{jk}(\mathbf{n})$. Note that whenever the Markov component is in state j, the following three types of transitions in $(\mathbf{X}, J)$ may occur with respective rates:

(i) arrivals without change of state in J occur at rate $\lambda_{jj}(\mathbf{n})$, $\mathbf{n} > \mathbf{0}$;

(ii) change of state in J without arrivals occur at rate $\lambda_{jk}(\mathbf{0})$, $k \in E$, $k \neq j$;

(iii) arrivals with change of state in J occur at rate $\lambda_{jk}(\mathbf{n})$, $k \in E$, $k \neq j$, $\mathbf{n} > \mathbf{0}$.

We denote $\Lambda_{\mathbf{n}} = (\lambda_{jk}(\mathbf{n}))$, $\mathbf{n} \in \mathbb{N}^r$, and say that $(\mathbf{X}, J)$ is a *simple* MAP of arrivals if $\Lambda_{(n_1,\ldots,n_r)} = 0$ if $n_l > 1$ for some l. In case $\mathbf{X}(= X)$ has state space $\mathbb{N}$ we say that (X, J) is a *univariate* MAP of arrivals.

The identification of the meaning of each of the above transitions is very important for applications. In particular, arrivals are clearly identified by transitions of types (i) and (iii), so that we may talk about arrival epochs, interarrival times, and define complex operations like thinning of MAPs of arrivals. Moreover the parametrization of these processes is simple and, by being Markov chains, they have "nice" structural properties. These are specially important computationally since it is a simple task to simulate Markov chains, although some problems may arise when the transition rates do not have any special structure, which is not usually the case in practice. We do not elaborate on this point; the interested reader may refer to Neuts [26] and the references cited there.

1.2.1 Univariate MAPs of arrivals

There has not been a systematic study of MAPs of arrivals, and the literature has been focused on univariate MAPs of arrivals with finite Markov component J (but most of the time without a connection to the theory on MAPs). In this subsection we give a brief account of the history and terminology for these processes. The emphasis in on the evolution of the structure of the processes, and in the following the reference is to univariate MAPs of arrivals with finite Markov component, investigated in the applied literature.

The simplest process corresponds to the case where the Markov component has only one state. It then follows from the definition of MAPs that X is a continous time Markov process with stationary and independent increments on $I\!N$. Thus X is a compound Poisson process. In the terminology for MAPs of arrivals the process is characterized by the rates $\Lambda_n = \lambda_n$ of arrivals of batches of size n ($n \geq 1$). We denote by $\lambda = \sum_{n\geq 1} \lambda_n$ the compound arrival rate of batches. (The simple Poisson process with rate λ is the special case where $\lambda_n = 0$ for $n \neq 1$, and obviously $\lambda_1 = \lambda$). We note that the generating function of the compound Poisson process is

$$G(z;t) = \exp\left\{-t \sum_{n\geq 1} \lambda_n (1-z^n)\right\}. \tag{1.3}$$

The compound Poisson process allows only transitions of type (i) (arrivals without change of state in the Markov component) which are the only ones admissible anyway. If we leave this simple one-state case, and get into genuine examples of MAPs of arrivals, transitions of types (ii) or (iii) have to be allowed. To our knowdlege, the first defined (simple) MAP of arrivals, with transitions of types (i) and (ii), was the Markov-modulated Poisson process (MMPP), which is a Cox process with intensity rate modulated by a finite Markov chain. This process was first used in queueing models by P. Naor and U. Yechialli, followed by M.F. Neuts; a recent compilation of results and relevant references on the MMPP is given by Fischer and Meier-Hellstern [11]. The MMPP was given a definition as an MAP of arrivals by Prabhu [31].

We may also consider the MMPP where batches of arrivals are allowed, which may be constructed from m independent compound Poisson processes $X_1, \ldots, X_m$ by observing the process X_j whenever the Markov component J is in state j (this follows from Neveu's [27] characterization of Markov subordinators with finite

Markov component). This setting suggests immediately that the process has important properties, such as lack of memory, being closed under Bernoulli thinning and under superpositioning of independent processes, similar to those of the compound Poisson process; this fact has not received sufficient attention in the applied literature. This process is an MAP of arrivals with the rate matrices Λ_n $(n \geq 1)$ being diagonal matrices. Its generating function is a natural matrix extension of the generating function of the compound Poisson process, namely

$$G(z;t) = \exp\left\{ t \left[Q - \sum_{n \geq 1} \Lambda_n (1 - z^n) \right] \right\} \tag{1.4}$$

where Q is the generator matrix of the Markov component J. In fact all univariate MAPs of arrivals have this generating function, with the Λ_n not necessarily diagonal. As all Cox processes, MMPPs are more *bursty* than the Poisson process in the sense that the variance of the number of counts is greater than its mean (see Kingman [19]), whereas for the Poisson process these quantities are equal.

Allowing only transitions of types (i) and (ii) amounts to stating that only one of the components X, J changes state at a time; this is consistent with viewing X having a causual relationship with J. Type (iii) transitions are sometimes called Markov-modulated arrivals in the literature.

Transitions of types $(i) - (iii)$ were allowed for the first time by Rudemo [36] who considered a simple process. Thus, in addition to arrivals as in the MMPP, there may be arrivals at transition epochs of the Markov chain J. Explicitly, at the time a transition from state j to state k occurs in J, an arrival occurs if $(j,k) \in A$ and no arrival occurs otherwise, where $A \subseteq \{(i,l) \in E^2 : i \neq l\}$. This process is thus a simple MAP of arrivals for which Λ_1 is not necessarily diagonal. The restrictive feature of the process is that

$$\lambda_{jk}(0)\,\lambda_{jk}(1) = 0 \qquad (j, k \in E). \tag{1.5}$$

The restrictive condition (1.5) is not present in the *phase-type* (PH) renewal process introduced by Neuts (see Neuts [25]), as a generalization of the Poisson process, containing modifications of the Poisson process such as the Erlangian and hyperexponential arrival processes. Here the interarrival times have a phase-type distribution, i.e. they are identified as times till absorption in a finite state Markov chain with one absorbing state. The process may be used to model sources that are less bursty than the Poisson sources.

The first defined univariate MAP of arrivals that essentially achieves the full generality possible when the Markov component has finite state space was the versatile Markovian arrival process of M.F. Neuts, or the so called N-process. The definition of this process is a constructive one (see Neuts [25]), and in the original formulation special care was taken to include arrivals of the same type as in both the MMPP and the PH renewal process. In addition to the transitions of types $(i) - (iii)$ (in the set of transient states of a finite Markov chain), the process allows for phase-type arrivals with one absorbing state. We shall see in chapter 2 that the inclusion of phase-type arrivals does not add any additional generality to the class of N-processes, and in fact makes the model overparametrized. This lack of a proper definition of the process of arrivals with appropriate transition rates $(i) - (iii)$ has led to considerable confusion in the literature.

Machihara [22] extended Neuts's idea of phase-type arrivals by considering the case of simple phase-type arrivals with more than one absorbing state (see Example 2), and defined a process which is closely related to the phase-type Markov renewal process we discuss in Example 3. This process was extended to the case of batch arrivals by Yamada and Machihara [43].

Lucantoni, Meier-Hellstern and Neuts [21] defined a second (simple) Markovian arrival process, which was generalized to allow for batch arrivals by Lucantoni, who named it batch Markovian arrival process (BMAP). The class of BMAPs is the class of univariate MAPs of arrivals with finite Markov component, and a definition of the BMAP is given in Remark 1 below. For a history of the BMAP and its applications, with special emphasis on matrix-geometric methods and a very extensive list of references see Lucantoni [20]. The simple BMAP has been referred commonly as MAP; this is very unfortunate since this process is just a very particular case of a Markov additive process and besides, this amounts to an abuse of notation since the acronym MAP for Markov additive processes was first used by Çinlar ([7,8]).

Finaly, we note that the *point processes of arrivals* (sequence of arrival epochs along with the corresponding batch sizes) associated with the classes of univariate MAPs of arrivals with finite Markov component, BMAPs, and Yamada and Machihara's arrival processes are the same (see Proposition 2). This fact does not seem to be well understood (cf Neuts [26]).

1.2.2 Applications of MAPs of arrivals

MAPs of arrivals arise in important applications, as components of more complex systems, specially in queueing and data communications models. They have been used to model overflow from trunk groups, superpositioning of packetized voice streams, and input to ATM (Asynchronous Transfer Mode) networks, which will be used in future high-speed communications networks (see e.g. [11,20,22,25,26, 39,43] for references). They have also been used to establish queueing theoretic results and to investigate constructions on arrival streams (see e.g. [20,26] for references).

Queues with Markov additive input have also been the subject of much study (see e.g. [11,21,20,22,25,34,39,43] and references cited there). The output from queues with Markov additive input has also been considered; in general this process is not an MAP of arrivals. In particular, Olivier and Walrand [28] have shown that the ouput from an $MMPP/M/1$ queue is not an MAP of arrivals unless the input is Poisson. Takine, Sengupta and Hasegawa [39] study the extent to which the output process *conforms* to the input process generated by an MAP of arrivals for a high-speed communications network with the *leaky bucket* for *regulatory acess mechanism* (or traffic shaping mechanism).

The class of stationary univariate MAPs of arrivals with finite Markov component was shown to be dense in the family of all stationary point processes on $\mathbb{R}_+$ by Asmussen and Koole [2] (here an MAP is said to be stationary if its Markov component is stationary). This fact and the considerable tractability of MAPs of arrivals make the class of MAPs of arrivals very important for applications.

MAPs of arrivals are studied in chapter 2. We investigate properties of MAPs of arrivals, such as the lack of memory property, interarrival times and moments of the number of counts. We consider transformations of MAPs of arrivals, both deterministic and random, which preserve the Markov-additive property; a particular case of this is the superposition of independent MAPs of arrivals. In addition we study secondary recordings of MAPs of arrivals; the emphasis is on Markov-Bernoulli recordings which include important special cases of thinning and marking. For Markov-Bernoulli recordings, the secondary processes turn out to be MAPs of arrivals.

1.3 Storage models for communications systems with Markov-additive input and demand

Storage models with Markov-additive input and demand arise frequently in models for communications systems. We review the literature on this type of models. We denote by $X(t)$, $D(t)$ and $I(t)$ the input, the demand and the reject net input in $(0, t]$, respectively.

The pioniering work in fluid storage models, for which the input and demand have no jumps, was done by Anick, Mitra and Sondhi [1], who proposed the following model for a multiple source data-handling system. There are N sources of messages, which may be *on* or *off* from time to time. A switch receives these messages at a unit rate from each *on* source and transmits them at a fixed maximum rate c ($1 \leq N < \infty$, $0 < c < \infty$), storing the messages that it cannot transmit in a buffer of infinite capacity. The sources act independently, the durations of *on* and *off* times being independent random variables with exponential densities. Denoting by $J(t)$ the number of *on* sources at time t, these assumptions amount to the statement that $J = \{J(t), t \geq 0\}$ is a birth and death process on the state space $\{0, 1, \ldots, N\}$. When J is in state j there is fluid input at rate $a(j) = j$ and the demand rate is $d(j) = c$. With $Z(t)$ being the buffer content at time t, the model states that

$$\begin{aligned} Z(t) &= Z(0) + X(t) - D(t) + I(t) \\ &= Z(0) + \int_0^t a\,(J(s))\,ds - \int_0^t d(J(s))\,ds + \int_0^t l\,(Z(s), J(s))\,ds \end{aligned} \tag{1.6}$$

with the condition $Z(0) \geq 0$, where

$$l\,(x, j) = \begin{cases} 0 & x > 0 \\ (a(j) - d(j))^- & x \leq 0 \end{cases} \tag{1.7}$$

with $a^- = \max(-a, 0)$, $a \in I\!R$. Here l represents the unsatisfied demand rate, and $I(t)$ the unsatisfied demand in $(0, t]$. It should be noted that the unsatisfied demand rate depends on the input and demand rates only through their difference, i.e. the net input rate. The authors derive the steady state distribution of the process (Z, J), employing matrix-algebraic techniques.

Gaver and Lehoczky [13] investigate a model for an integrated circuit and packet switched multiplexer, in which there are two types of input – data and voice calls. The input of data occurs continuously at a constant rate c_0 (fluid input), and is transmitted at rate c_2. There are $s + u$ output channels, of which s are reserved for data transmission, while the remaining u are shared by data and voice calls, with calls having priority with preemptive discipline. Calls arrive at a Poisson rate and have independent holding times with exponential density. Calls that find all u channels that service them busy are lost. Data that cannot be transmitted are stored in a buffer of infinite capacity. Let $Z(t)$ be the buffer content and $J(t)$ the number of channels out of u not occupied by calls at time t. Then the model states that $u - J$ is identical with the number in the Erlang loss system $M/M/u/u$, while $Z(t)$ satisfies equations (1.6)-(1.7) with the input and demand given by

$$X(t) = c_0\, t, \qquad D(t) = \int_0^t c_2\,(s + J(s))\, ds. \tag{1.8}$$

Mitra [23] investigates a fluid model for a system with two groups of machines connected by a buffer. The first group of machines (producers) produces a fluid which is transferred to the buffer and is consumed by a second group of machines (consumers). These machines are *in service* or *in repair* states from time to time and the assumptions for *in service* times and *in repair* times are similar to those on Anick et al. [1] for *on* and *off* periods. The buffer has capacity $C\,(0 < C \leq \infty)$, while the production rate and consumption rate by machine, assumed to be constants, are c_1 and c_2. We let $J_1(t)$ and $J_2(t)$ denote respectively the number of *producers* and *consumers* that are *in service* at time t, from the total of N *producers* and M *consumers.* It follows that $J_1 = \{J_1(t)\, t \geq 0\}$ and $J_2 = \{J_2(t)\, t \geq 0\}$ are independent Markov processes on $\{0, 1, \ldots, N\}$ and $\{0, 1, \ldots, M\}$, respectively. Again, if we denote by $Z(t)$ the buffer content at time t and let $J = (J_1, J_2)$, the model states that $Z(t)$ satisfies equation (1.6) with the input and demand given by

$$X(t) = \int_0^t c_1\, J_1(s)\, ds, \qquad D(t) = \int_0^t c_2\, J_2(s)\, ds \tag{1.9}$$

and the following lost net input rate

$$l\,(x, j_1, j_2) = \begin{cases} (c_1\, j_1 - c_2\, j_2)^- & x \leq 0 \\ 0 & 0 < x < C \\ -(c_1\, j_1 - c_2\, j_2)^+ & x \geq C \end{cases} \tag{1.10}$$

where $a^+ = \max(a, 0)$ for $a \in I\!R$. This model is applicable in communication-integrated systems with *producers* being interpreted as sources and *consumers* as channels.

Virtamo and Norros [42] have investigated a model in which a buffer receives input of data from an $M/M/1$ queueing system at a constant rate c_0 so long as the system is busy, and transmits these data at a maximum rate c_1 ($< c_0$). Denoting by $J(t)$ the queue length, the model states that the buffer content $Z(t)$ satisfies equations (1.6)-(1.7) with the input and demand given by

$$X(t) = \int_0^t c_0 \, 1_{\{J(s)>0\}} \, ds, \qquad D(t) = c_1 \, t. \tag{1.11}$$

Note that the input during $(0, t]$ is given by $c_0 B(t)$ with $B(t)$ being the part of the time interval $(0, t]$ during which the server is busy.

The $BMAP/G/1$ queue and some of its particular cases have been studied by several authors; see in particular [11,21,20,22,25,34,39,43] and references cited there. Here customers with i.i.d. service times arrive according to an univariate MAP of arrivals (W, J) with finite Markov component, the increments on the arrival component being equal to the batch size of arriving customers. The work in the system at time t, $Z(t)$, satisfies the equation

$$Z(t) = Z(0) + X(t) - t + \int_0^t 1_{\{Z(s)\leq 0\}} \, ds \tag{1.12}$$

where $X(t) = \sum_{i=1}^{W(t)} X_i$, with $(X_1, X_2, \ldots)$ being an i.i.d. sequence of random variables with values in $[0, \infty)$. The model imples that the input (X, J) is a Markov-compound Poisson process (see Prabhu and Zhu [34]), whereas the demand is at unit rate. The techniques used to study these queueing systems are more probabilitisc in nature than the ones used to study the models we described before. For an excellent review of the work on the $BMAP/G/1$ queue see Lucantoni [20].

A model for an acess regulator was investigated by Elwalid and Mitra [10]. Here the input of data is regulated by a continuous time Markov chain J with a finite number of states, the input rate being $\lambda(j)$ when J is in state j (the so called Markov modulated fluid source). In addition to a finite data buffer receiving messages (cells), as in Mitra [23], there is a token buffer also of finite size, receiving tokens continuously at rate r. In order to get transmitted, cells have to be combined with tokens. Some cells are lost because of a regulator that controls transmission subject to a peak rate ν.

Models similar to those described here arise also in manufacturing systems. Browne and Zipkin [5] consider an (s, S) inventory system where the demand rate is controlled by an underlying Markov chain J. The model states that the demand in $(0, t]$ is given by

$$D(t) = \int_0^t d(J(s))\, ds. \tag{1.13}$$

The authors study steady state properties of the system. Sevast'yanov [38] analyzes a system with N machines in tandem with buffers in between. Each item has to be processed sucessively by machines $1, 2, \ldots, N$, and items that cannot be processed by a machine are stored in the preceding buffer. The system may be viewed as a tandem storage model modulated by an auxiliary process J indicating the number of machines in operation. With appropriate assumptions, J will turn to be an N-dimensional Markov chain, each component of which indicates whether or not the corresponding machine is in operation. This model and the one by Elwalid and

Mitra [10] are simple examples of storage networks modulated by an underlying Markov chain.

1.4 A storage model with Markov-additive input and demand

In this section we define a storage model having Markov-additive input and demand. We let $Z(t)$ be the storage level at time t, and consider the following model.

(a) *The Input.* The input $\{X(t),\, t \geq 0\}$ is such that $(X, J) = \{(X, J)(t),\, t \geq 0\}$ is a Markov-compound Poisson process (MCPP), defined on some probability space $(\Omega, \mathcal{F}, \mathcal{P})$, with infinitesimal generator given by

$$\mathcal{A}_{(X,J)}\, f(x, j) = a(j)\, \frac{\partial}{\partial x} f(x, j) + \sum_{k \in E} \sigma_{jk} \int_0^\infty [f(x+y, k) - f(x, j)] M_{jk}\{dy\}. \quad (1.14)$$

Here, E is a discrete set, a is a nonnegative function on E and $f(x, j)$ is a bounded function on $\mathbb{R}_+ \times E$ such that for each fixed j, f is continuous and has a bounded continuous derivative $\frac{\partial f}{\partial x}(x, j)$, and by $\int_a^b$ we mean $\int_{[a,b)}$. Thus, when J is in state j input arrives at a continuous rate $a(j)$, and jumps in X associated with a transition from state j to state k in J occur at rate σ_{jk} and have distribution M_{jk}. The distributions M_{jk}, $j, k \in E$ are concentrated on $[0, \infty)$.

It may be seen that

$$X(t) = X_0(t) + \int_0^t a\,(J(s))\, ds \quad (1.15)$$

with $X_0(t)$ being the part of the input in $(0, t]$ associated with jumps. We assume that (X_0, J) is a pure jump process, i.e. its sample paths are constant except for isolated jumps, and right continuous, for any initial distribution. This implies in particular that the process is càdlàg (right-continuous with left-hand limits), and that J is a pure Markov jump process (MJP). This formulation provides for the possibility of two sources of input, one slow source bringing in data in a fluid fashion and the other bringing in packets.

(b) *The Demand.* As for the demand for transmission of data we assume that it arises at a rate $d(j)$ when the current state of J is j (similar to what happens in the model considered by Mitra [23]), i.e

$$D(t) = \int_0^t d(J(s))\, ds \tag{1.16}$$

so that the infinitesimal generator of (D, J) is given by

$$\mathcal{A}_{(D,J)} f(x,j) = d(j)\frac{\partial}{\partial x} f(x,j) + \sum_{k \neq j} \sigma_{jk}[f(x,k) - f(x,j)]. \tag{1.17}$$

(c) *The Storage Policy.* The storage has infinite capacity, and the demand is satisfied if physically possible, so that the unsatisfied demand is given by

$$I(t) = \int_0^t y(J(s))^- 1_{\{Z(s) \leq 0\}}\, ds. \tag{1.18}$$

with $y = a - d$.

For this storage model we investigate in chapter 3 the properties of the actual and unsatisfied demands, along with the inverse of the actual demand. The storage equation of this model is

$$Z(t) = Z(0) + X(t) - D(t) + \int_0^t y(J(s))^- 1_{\{Z(s) \leq 0\}}\, ds \tag{1.19}$$

with $Z(0) \geq 0$ almost surely (a.s.).

We will show in section 3.1 that (1.19) has a unique solution given by

$$Z(t) = Z(0) + Y(t) + I(t) \tag{1.20}$$

where $Y(t) = X(t) - D(t)$ is the *net input* process, and

$$I(t) = [Z(0) + m(t)]^- = \left[Z(0) + \inf_{0 \le \tau \le t} Y(\tau)\right]^- . \tag{1.21}$$

We note that the unsatisfied demand depends on the input and demand only through the net input. From work currently in progress, it follows that a solution of type (1.20)-(1.21) is valid for a broad class of infinite capacity storage models with càdlàg net input, when the storage policy is to satisfy the demand if physically possible.

We note the input and the demand having a common Markov component, is not a restrictive assumption. We can formulate an input process (X, J_1) and a demand rate function $d(J_2(t))$, where $J = (J_1, J_2)$ is a two-dimensional Markov chain with possibly dependent components J_1, J_2, such that J_1, J_2 together are the Markov component of the input and demand. As an example, in Mitra [23] $J = (J_1, J_2)$ is a two-dimensional birth and death process. Also the assumption of infinite storage (buffer) capacity seems to be fairly realistic, as suggested by D. Mitra in a personal conversation. In any case, we believe that our model will serve at least as an approximation to the finite capacity case and will in fact provide a reasonable understanding of how data communications networks operate.

Our unsatisfied demand rate is time dependent and induces dependence on the storage level through the underlying Markov component. It is possible to formulate more general unsatisfied demand rates among the ones that are a function of the storage level as well the Markov component. However, considerable analytical difficulties may arise, and some desirable properties of the model may not hold.

1.4.1 The fluid model

The model with fluid input, i.e. $X_0(t) = 0$, is considered in chapter 4. Particular cases of this are the models studied by Anick, Mitra and Sondhi [1], Gaver and Lehoczky [13], Van Doorn, Jagers and de Wit [40] and Virtamo and Norros [42]. The net input $Y(t)$ in this model is then given by

$$Y(t) = A(t) - D(t) = \int_0^t y(J(s))\, ds. \tag{1.22}$$

Its sample functions are continuous a.s., and differentiable everywhere except at the transition epochs T_n $(n \geq 0)$ of the Markov chain J. If we let $J_n = J(T_n)$, then we have for $T_n \leq t \leq T_{n+1}$

$$Z(t) = Z(T_n) + y(J_n)(t - T_n) + y(J_n)^- \int_{T_n}^t 1_{\{Z(s)\leq 0\}}\, ds \tag{1.23}$$

and

$$I(t) = I(T_n) + y(J_n)^- \int_{T_n}^t 1_{\{Z(s)\leq 0\}}\, ds. \tag{1.24}$$

This shows that in order to study the process (Z, I, J) it may be of interest to first study the properties of the embedded process $(Z(T_n), I(T_n), J(T_n))$, which we do in chapter 4. A particular consequence of (1.20)-(1.21) is that $Z(T_n)$ and $I(T_n)$ may be identified as functionals on the process $(T_n, Y(T_n), J(T_n))$, which is an MRW. So the properties of this MRW are investigated, the key result being a Wiener-Hopf factorization due to Presman [35]; see Prabhu and Tang [32] and Prabhu, Tang and Zhu [33]. These properties are then used to study the properties of the storage level and the unsatisfied demand.

1.4.2 A storage model with jump input process

In chapter 5 we consider the storage model for which the input has no continuous component, i.e. $a = 0$ or $X = X_0$. Since the unsatisfied demand depends on the input and demand only through the net input, this is equivalent to the model for which the demand rate exceeds the input rate from the slow source, i.e. $d(j) \geq a(j)$ for $j \in E$. This storage model includes as a very special case the workload process of a $BMAP/G/1$ queue. Since we may assume $a = 0$, the storage equation is

$$Z(t) = Z(0) + X(t) - D(t) + \int_0^t d(J(s))\, 1_{\{Z(s) \leq 0\}}\, ds \tag{1.25}$$

with the condition $Z(0) \geq 0$, and (X, J) being a Markov-compound Poisson process with no drift.

Our analysis is based on the study of the busy period, using techniques based on infinitesimal generators. The transform of the busy period is the unique solution of a certain matrix-functional equation. We obtain transforms of the various processes of interest and investigate the steady state behaviour of the model. The results obtained are not obvious generalizations of the results for simple storage models; in particular, a generalization of the Pollaczek-Khinchin formula brings new insight.

Chapter 2

MAPs of arrivals

In this chapter we consider MAPs of arrivals. These were introduced in section 1.2, where their importance was highlighted. To study these processes it is convenient to derive some properties of general MAPs. This is done in section 2.2, and the emphasis is on transformations of MAPs that preserve the Markov additive property. MAPs of arrivals are formally introduced in section 2.3 along with some examples, and some of their elementary properties are studied in section 2.4; these include the lack of memory, interarrival times, moments of the number of counts, and a law of large numbers. In section 2.5 we consider transformations of MAPs of arrivals, both deterministic and random. In section 2.6 we define a type of secondary recording which includes important special cases of marking and thinning. We consider in particular Markov-Bernoulli recording of MAPs of arrivals, for which the secondary process turns out to be an MAP of arrivals.

2.1 Notations

In this section we introduce some further notations. For a stochastic process Y defined on a probability space $(\Omega, \mathcal{F}, \mathcal{P})$, we let $\{\mathcal{F}_t^Y\}$ be the filtration generated by Y, so that $\mathcal{F}_t^Y = \sigma\{Y(s), 0 \leq s \leq t\}$ for $t \in \mathcal{T}$, where for a family $\mathcal{G}$ of random variables $\sigma(\mathcal{G})$ is the σ-field generated by $\mathcal{G}$. As usual, for a transformation $T : \Omega_1 \to \Omega_2$ we denote $T^{-1}(A) = \{\omega_1 \in \Omega_1 : T(\omega_1) \in A\}$ for $A \subseteq \Omega_2$.

We denote by $\mathbf{e}$ a vector with unit elements, with the dimension being clear from the context. Suppose $A = (a_{jk})$ and $B = (b_{jk})$ are matrices of the same order. We define the *Schur* or *entry-wise multiplication* of A and B by $A \bullet B = (a_{jk}\, b_{jk})$. If for countable sets $F_1, \ldots, F_4$ we have $A = (a_{jk})_{F_1 \times F_2}$ and $B = (b_{il})_{F_3 \times F_4}$ then we define the *Kronecker sum* of A and B by

$$A \oplus B = (c_{(j,i)(k,l)})_{(F_1 \times F_3) \times (F_2 \times F_4)} = (a_{jk}\, \delta_{il} + \delta_{jk}\, b_{il}).$$

For a countable set F we let $l_\infty(F)$ be the Banach space of real sequences $a = (a_j)_{j \in F}$ with the norm $\| a \| = \sup_{j \in F} | a_j |$ and $B(l_\infty(F))$ be the space of bounded linear operators on $l_\infty(F)$, an element W of which may be identified by a matrix $W = (w_{jk})_{j,k \in F}$ with norm

$$\| W \| = \sup_{j \in F} \sum_{k \in F} | w_{jk} | .$$

For an MAP $(\mathbf{X}, J)$, if J is irreducible with stationary distribution $\pi = (\pi_j)$ we call the version of $(\mathbf{X}, J)$ for which $J(0)$ has distribution π the *stationary version* of the MAP $(\mathbf{X}, J)$.

We consider in $I\!R^r$ partial order relations $\leq$, $<$, $\geq$ and $>$

$$\mathbf{x} \leq \mathbf{y} \Leftrightarrow x_i \leq y_i,\ i = 1, 2, \ldots, r; \qquad \mathbf{x} < \mathbf{y} \Leftrightarrow \mathbf{x} \leq \mathbf{y} \text{ and } \mathbf{x} \neq \mathbf{y}$$

and $\mathbf{x} \geq \mathbf{y} \Leftrightarrow \mathbf{y} \leq \mathbf{x}$, $\mathbf{x} > \mathbf{y} \Leftrightarrow \mathbf{y} < \mathbf{x}$.

2.2 Some properties of MAPs

In this section we investigate some general properties of MAPs. We assume w.l.o.g. that if $(\mathbf{X}, J)$ is an MAP then $\mathbf{X}(0) = \mathbf{0}$ a.s. For clarity, we consider only time-homogeneous MAPs. By using (1.1), we have

$$\begin{aligned} P\{\mathbf{X}(s+t) \in A;\ J(s+t) = k \mid \mathbf{X}(s) = \mathbf{x}, J(s) = j)\} \\ = P\{\mathbf{X}(t) \in A - \mathbf{x};\ J(t) = k \mid J(0) = j\} \qquad (2.1) \end{aligned}$$

for $s, t \in \mathcal{T}$, $j, k \in E_\Delta$, $\mathbf{x} \in I\!R^r$ and $A \in \mathcal{R}^r$. Since this property characterizes homogeneous MAPs, it suffices to define the transition probability measure of the process,

$$P_{jk}(A;t) = P\{\mathbf{X}(t) \in A,\ J(t) = k \mid J(0) = j\} \qquad (2.2)$$

$$P_{jk}(A;t) \geq 0, \qquad P_{jk}(A;0) = \delta_{jk}\, 1_{\{\mathbf{0} \in A\}}, \qquad \sum_{k \in E_\Delta} P_{jk}(I\!R^r;t) \leq 1. \qquad (2.3)$$

When A is a singleton $\{\mathbf{x}\}$, we write $P(\mathbf{x};t)$ instead of $P(\{\mathbf{x}\};t)$. For the Chapman-Kolmogorov equations we have

$$P_{jk}(A;t+s) = \sum_{l \in E_\Delta} \int_{I\!R^r} P_{jl}(d\mathbf{x};t)\, P_{lk}(A - d\mathbf{x};s) \qquad (s, t \in \mathcal{T}) \qquad (2.4)$$

or in matrix form, with $P(A;t) = (P_{jk}(A;t))$,

$$P(A;t+s) = \int_{I\!R^r} P(d\mathbf{x};t)\, P(A - d\mathbf{x};s). \qquad (2.5)$$

From (2.2) the transition probabilities of J are found to be $\pi_{jk}(t)$, where

$$\Pi(t) = (\pi_{jk}(t)) = (P\{J(t) = k \mid J(0) = j\}) = (P_{jk}(I\!R^r;t)) \qquad (2.6)$$

with

$$\pi_{jk}(t) \geq 0, \quad \pi_{jk}(0) = \delta_{jk}, \quad \sum_{k \in E_\Delta} \pi_{jk}(t) \leq 1 \tag{2.7}$$

$$\pi_{jk}(t+s) = \sum_{l \in E_\Delta} \pi_{jl}(t)\, \pi_{lk}(s). \tag{2.8}$$

We say that $(\mathbf{X}, J)$ is a strong MAP if for any stopping time T we have

$$P\left\{\mathbf{X}(T+t) - \mathbf{X}(T) \in A,\ J(T+t) = k \mid \mathcal{F}_T^{(\mathbf{X},J)}\right\} = P_{J(T)k}(A;t) \tag{2.9}$$

for $t \in \mathcal{T}$, $k \in E_\Delta$ and $A \in \mathcal{R}^r$. The following result presents important properties of deterministic transformations of MAPs.

Theorem 1 *For $A \in \mathcal{R}^r$ and $B \in \mathcal{R}^s$ we have the following.*

(a). **Linear transformations of MAPs.**

Suppose that $(\mathbf{X}, J)$ is an MAP on $I\!R^r \times E$ with transition probability measure $P^{\mathbf{X}}$, and $T : I\!R^r \to I\!R^s$ is a linear transformation.

If $\mathbf{Y} = T(\mathbf{X})$ then $(\mathbf{Y}, J)$ is an MAP on $I\!R^s \times E$ with transition probability measure $P^{\mathbf{Y}}$ such that

$$P_{jk}^{\mathbf{Y}}(B;t) = P_{jk}^{\mathbf{X}}\left(T^{-1}(B);t\right). \tag{2.10}$$

(b). **Patching together independent MAPs.**

Suppose that $(\mathbf{X}, J_1)$ and $(\mathbf{Y}, J_2)$ are MAPs on $I\!R^r \times E_1$ and $I\!R^s \times E_2$ with transition probability measure $P^{\mathbf{X}}$ and $P^{\mathbf{Y}}$, respectively.

If $(\mathbf{X}, J_1)$ and $(\mathbf{Y}, J_2)$ are independent then $((\mathbf{X}, \mathbf{Y}), (J_1, J_2))$ is an MAP on $I\!R^{r+s} \times E_1 \times E_2$ with transition probability measure $P^{(\mathbf{X},\mathbf{Y})}$ such that

$$P_{(j_1,j_2)(k_1,k_2)}^{(\mathbf{X},\mathbf{Y})}(A \times B;t) = P_{j_1k_1}^{\mathbf{X}}(A;t)\, P_{j_2k_2}^{\mathbf{Y}}(B;t). \tag{2.11}$$

Proof: (a). Blumenthal ([3], Theorem II(1.2)) states sufficient conditions under which transformations of Markov processes are themselves Markov. These conditions do not apply necessarilly to our case, but we may prove the desired result more directly as follows.

Suppose $0 = t_0 < t_1 < \ldots < t_n < t$, $\mathbf{y}_p \in I\!\!R^s$ $(0 \le p \le n)$, $B \in \mathcal{R}^s$ and let

$$A_p = \{(\mathbf{Y}, J)(t_p) = (\mathbf{y}_p, j_p)\} \quad (0 \le p \le n).$$

From the fact that T is a measurable transformation we have

$$A_p = \{(\mathbf{X}, J)(t_p) \in T^{-1}(\{\mathbf{y}_p\}) \times \{j_p\}\} \in \sigma((\mathbf{X}, J)(t_p)) \qquad (0 \le p \le n)$$

and similarly

$$A_t = \{(\mathbf{X}, J)(t) \in T^{-1}(B) \times \{k\}\} \in \sigma((\mathbf{X}, J)(t)).$$

Since $(\mathbf{X}, J)$ is an MAP, these facts imply that

$$\begin{aligned} P\{\mathbf{Y}(t) \in B;\, J(t) = k \mid \bigcap_{0 \le p \le n} A_p\} &= P\{A_t \mid \bigcap_{0 \le p \le n} A_p\} = P\{A_t \mid A_n\} \\ &= P\{\mathbf{Y}(t) \in B;\, J(t) = k \mid \mathbf{Y}(t_n) = \mathbf{y}_n,\, J(t_n) = j_n\} \\ &= P\{\mathbf{Y}(t) - \mathbf{Y}(t_n) \in B - \mathbf{y}_n;\, J(t) = k \mid A_n\} \\ &= P\{\mathbf{X}(t) - \mathbf{X}(t_n) \in T^{-1}(B - \mathbf{y}_n);\, J(t) = k \mid A_n\} \\ &= P\{\mathbf{X}(t) - \mathbf{X}(t_n) \in T^{-1}(B - \mathbf{y}_n);\, J(t) = k \mid J(t_n) = j_n\} \\ &= P\{\mathbf{X}(t - t_n) \in T^{-1}(B - \mathbf{y}_n);\, J(t - t_n) = k \mid J(0) = j_n\} \\ &= P\{\mathbf{Y}(t - t_n) \in B - \mathbf{y}_n;\, J(t - t_n) = k \mid J(0) = j_n\}. \end{aligned}$$

This shows that $(\mathbf{Y}, J)$ is an MAP with the given transition probability measure.

(b). Since $(\mathbf{X}, J_1)$ and $(\mathbf{Y}, J_2)$ are independent time-homogeneous Markov processes it is a well known result that $(\mathbf{X}, \mathbf{Y}, J_1, J_2)$ is a time-homogeneous Markov process. Property (ii) of MAPs for $((\mathbf{X}, \mathbf{Y}), (J_1, J_2))$ follows using arguments similar to the ones needed to prove that $(\mathbf{X}, \mathbf{Y}, J_1, J_2)$ is Markov, but its proof is given for completeness.

Suppose $s, t \in \mathcal{T}$. Since $(\mathbf{X}, J_1)$ and $(\mathbf{Y}, J_2)$ are MAPS and are independent, it is easy to see that

(1) the conditional distribution of $(\mathbf{X}(s+t) - \mathbf{X}(s), J_1(s+t))$ given $(\mathbf{X}, \mathbf{Y}, J_1, J_2)(s)$ depends only on $J_1(s)$.

(2) the conditional distribution of $(\mathbf{Y}(s+t) - \mathbf{Y}(s), J_2(s+t))$ given $(\mathbf{X}, \mathbf{Y}, J_1, J_2)(s)$ depends only on $J_2(s)$.

From (1)-(2) and the independence of $(\mathbf{X}, J_1)$ and $(\mathbf{Y}, J_2)$ we conclude that

(3) $(\mathbf{X}(s+t) - \mathbf{X}(s), J_1(s+t))$ and $(\mathbf{Y}(s+t) - \mathbf{Y}(s), J_2(s+t))$ are conditionally independent given $(\mathbf{X}, \mathbf{Y}, J_1, J_2)(s)$.

From (1)-(3), the conditional distribution of

$$((\mathbf{X}, \mathbf{Y})(s+t) - (\mathbf{X}, \mathbf{Y})(s), (J_1, J_2)(s+t))$$

given $(\mathbf{X}, \mathbf{Y}, J_1, J_2)(s)$ depends only on $(J_1, J_2)(s)$. From (3) and (1)-(2), it follows that the probability transition measure of $((\mathbf{X}, \mathbf{Y}), (J_1, J_2))$ is as stated. □

We note that in Theorem 1, if the original processes are strong MAPs so are the transformed processes. In addition, if the original processes are Markov subordinators so are the resulting processes obtained from patching together independent processes or by taking linear transformations with nonnegative coefficients. Similar remarks apply to Corollary 1.

Corollary 1 *For $\alpha, \beta \in I\!R$, $A \in \mathcal{R}^r$ and $B \in \mathcal{R}^s$ we have the following.*

(a). **Marginals of MAPs.**

If $(\mathbf{X}, J)$ is an MAP on $I\!R^r \times E$ with transition probability measure $P^{\mathbf{X}}$ and $\mathbf{Z} = (X_{i_1}, X_{i_2}, \ldots, X_{i_s})$ with $1 \leq i_1 < \ldots < i_s \leq r$ then $(\mathbf{Z}, J)$ is an MAP on $I\!R^s \times E$.

If w.l.o.g. we assume that $\mathbf{Z} = (X_1, X_2, \ldots, X_s)$ and let $P^{\mathbf{Z}}$ be the transition probability measure of $(\mathbf{Z}, J)$ then

$$P^{\mathbf{Z}}_{jk}(B) = P^{\mathbf{X}}_{jk}(B \times I\!R^{r-s}). \tag{2.12}$$

(b). **Linear combinations of dependent MAPs.**

If $\mathbf{X} = (X_1, \ldots, X_r)$, $\mathbf{Y} = (Y_1, \ldots, Y_r)$ and $((\mathbf{X}, \mathbf{Y}), J)$ is an MAP on $I\!R^{2r} \times E$ with transition probability measure $P^{(\mathbf{X},\mathbf{Y})}$, then the process $(\alpha\mathbf{X} + \beta\mathbf{Y}, J)$ is an MAP on $I\!R^r \times E$ with transition probability measure

$$P_{jk}(A; t) = P^{(\mathbf{X},\mathbf{Y})}_{jk}\left(\{\mathbf{z} = (\mathbf{x}, \mathbf{y}) \ \ (\mathbf{x}, \mathbf{y} \in I\!R^r) \ : \ \alpha\mathbf{x} + \beta\mathbf{y} \in A\}\,; t\right). \tag{2.13}$$

(c). **Linear combinations of independent MAPs.**

If $(\mathbf{X}, J_1)$ and $(\mathbf{Y}, J_2)$ are independent MAPs on $I\!R^r \times E_1$ and $I\!R^r \times E_2$ with transition probability measure $P^{\mathbf{X}}$ and $P^{\mathbf{Y}}$, respectively, then $(\alpha\mathbf{X} + \beta\mathbf{Y}, (J_1, J_2))$ is an MAP on $I\!R^r \times E_1 \times E_2$ with transition probability measure

$$P_{(j_1,j_2)(k_1,k_2)}(A; t) = \int_{\{\mathbf{z}=(\mathbf{x},\mathbf{y}) \ (\mathbf{x},\mathbf{y} \in I\!R^r) \,:\, \alpha\mathbf{x}+\beta\mathbf{y} \in A\}} P^{\mathbf{X}}_{j_1k_1}(d\mathbf{x}; t)\, P^{\mathbf{Y}}_{j_2k_2}(d\mathbf{y}; t). \tag{2.14}$$

Proof: (a) and (b). The statements follow from Theorem 1 (a) since marginals and linear combinations are special cases of linear transformations.

(c). To prove the statement we first patch together $(\mathbf{X}, J_1)$ and $(\mathbf{Y}, J_2)$ to get in their product probability space the process $((\mathbf{X}, \mathbf{Y}), (J_1, J_2))$, which by Theorem 1 (b) is an MAP with probability transition measure given by (2.11). Using (b), it then follows that $(\alpha\mathbf{X} + \beta\mathbf{Y}, (J_1, J_2))$ is an MAP of arrivals with transition probability measure as given. □

If $(\mathbf{X}, J)$ is an MAP on $\mathbb{R}^r \times E$ and $0 = t_0 \le t_1 \le t_2 \le \ldots$ is a deterministic sequence, then the process $(S^\star, J^\star) = \{(S_n^\star, J_n^\star) = (\mathbf{X}(t_n), J(t_n)), n \in \mathbb{N}\}$ is an MRW on $\mathbb{R}^r \times E$ (which is time-homogeneous in case $t_n = nh\, (n \in \mathbb{N})$ for some $h > 0$). In order for the embedded process to be an MRW, the embedded sequence does not need to be deterministic, and more interesting embeddings may be obtained by using a random sequence of embedding times.

As for MAPs, we will consider the time-homogeneous case of MRWs, in which the second probability in (1.2) does not depend on m. Since $(\mathbf{S}^\star, J^\star)$ is an MAP, it suffices to define the one-step transition probability measure

$$V_{jk}(A) = P\{\mathbf{S}_1^\star \in A;\, J_1^\star = k \mid J_0^\star = j\} \tag{2.15}$$

along with the initial distribution of $J^\star$.

Theorem 2 *Suppose* $(\mathbf{X}, J)$ *is an MAP on* $\mathbb{R}^r \times E$, $T_n^\star\, (n \ge 0)$ *are stopping times such that* $0 = T_0^\star \le T_1^\star \le \ldots < \infty$ *a.s. Denote*

$$J_n^\star = J(T_n^\star), \qquad \mathbf{S}_n^\star = \mathbf{X}(T_n^\star) \qquad (n \in \mathbb{N}).$$

If $(\mathbf{X}, J)$ *is a strong MAP then* $(T^\star, \mathbf{S}^\star, J^\star) = \{(T_n^\star, \mathbf{S}_n^\star, J_n^\star), n \in \mathbb{N}\}$ *is an*

MRW on $I\!R_+ \times I\!R^r \times E$*. Moreover, in case the conditional distribution of* $(T^\star_{n+1} - T^\star_n, S^\star_{n+1} - S^\star_n, J^\star_{n+1})$ *given* $J^\star_n$ *does not depend on* n*, then* $(T^\star, \mathbf{S}^\star, J^\star)$ *is an homogeneous MRW with one-step transition probability measure*

$$V_{jk}(A \times B) = P\{T^\star_1 \in A;\ \mathbf{X}(T^\star_1) \in B;\ J(T^\star_1) = k \mid J(0) = j\}$$

for $j, k \in E_\Delta$*,* $A \in \mathcal{R}_+$ *and* $B \in \mathcal{R}^r$*.* □

In applications, the most common use of Theorem 2 is for the case where the embedding points are the sucessive transition epochs in J.

2.3 MAPs of arrivals

In this and the following sections we consider time-homogeneous MAPs of arrivals. Transitions of an MAP of arrivals $(\mathbf{X}, J)$ on $I\!N^r \times E$ may be identified by the states of J immediately before (j) and after (k) the transition occurs along with the corresponding observed increment in $\mathbf{X}$ $(\mathbf{n})$. Thus we characterize each transition by one element of the set $S_r(E)$

$$S_r(E) = \{(j, k, \mathbf{n}) \in E^2 \times I\!N^r : (\mathbf{n}, k) \neq (\mathbf{0}, j)\}. \tag{2.16}$$

For a set $C \subseteq S_r(E)$ and transition rates $\lambda_{jk}(\mathbf{n})$ we let $\lambda(C) = \{\lambda_{jk}(\mathbf{n}) : (j, k, \mathbf{n}) \in C\}$. The process $(\mathbf{X}, J)$ is parametrized by the set of transition rates $\lambda(S_r(E))$. In general some of the rates will be zero, thus it suffices to give the positive rates $\lambda(S^+_r(E))$ where

$$S^+_r(E) = \{(j, k, \mathbf{n}) \in S_r(E) : \lambda_{jk}(n) > 0\}. \tag{2.17}$$

In the following we always assume that $\lambda_{jk}(\mathbf{n}) = 0$ for $(j, k, \mathbf{n}) \notin S_r(E)$. For the transition probability measure of $(\mathbf{X}, J)$ we have

$$P\left(\mathbf{X}(h) = \mathbf{n},\, J(h) = k \mid J(0) = j\right) = \begin{cases} \lambda_{jk}(\mathbf{n})\, h + o(h) & (\mathbf{n}, k) \neq (\mathbf{0}, j) \\ 1 - \gamma_j\, h + o(h) & (\mathbf{n}, k) = (\mathbf{0}, j) \end{cases} \tag{2.18}$$

with

$$\gamma_j = \sum_{\{(k,\mathbf{n}):(j,k,\mathbf{n})\in S_r^+(E)\}} \lambda_{jk}(\mathbf{n}) < \infty. \tag{2.19}$$

A process with (2.19) is said to be *stable*. The infinitesimal generator of $(\mathbf{X}, J)$ is given by

$$\mathcal{A}f(\mathbf{m}, j) = \sum_{\{(k,\mathbf{n}):(j,k,\mathbf{n})\in S_r^+(E)\}} \lambda_{jk}(\mathbf{n})\, [f(\mathbf{m} + \mathbf{n}, k) - f(\mathbf{m}, j)] \tag{2.20}$$

with f being a bounded real function on $\mathbb{N}^r \times E$. This shows that the process is determined by the matrices $\Lambda_{\mathbf{n}} = (\lambda_{jk}(\mathbf{n}))$, with $\mathbf{n} \in \mathbb{N}^r$. Let

$$\Gamma = (\gamma_j\, \delta_{jk}), \qquad \Lambda = (\lambda_{jk}) = \sum_{\mathbf{n}>\mathbf{0}} \Lambda_{\mathbf{n}} \qquad \Sigma = (\sigma_{jk}) = \sum_{\mathbf{n}\in\mathbb{N}^r} \Lambda_{\mathbf{n}} \tag{2.21}$$

$$Q = (q_{jk}) = \Sigma - \Gamma, \qquad Q^{\mathbf{0}} = (q^{\mathbf{0}}_{jk}) = \Lambda_{\mathbf{0}} - \Gamma. \tag{2.22}$$

It is easy to see that Λ (Σ) is the matrix of transition rates in J associated with arrivals (either arrivals or non-arrivals). Note that Q and $Q^{\mathbf{0}}$ are matrices with non-positive diagonal entries and nonnegative off-diagonal entries. Moreover the row sums of Q $(Q^{\mathbf{0}})$ are null (non-positive). This implies in particular that Q is the generator matrix of some Markov chain. We will see that Q is the generator matrix of the Markov component J and $Q^{\mathbf{0}}$ is the generator matrix associated with non-arrival transitions in $(\mathbf{X}, J)$. These facts may be anticipated by inpection of the transition rates in the non-diagonal entries of these matrices.

The process $(\mathbf{X}, J)$ is thus characterized by $(Q, \{\Lambda_\mathbf{n}\}_{\mathbf{n}>\mathbf{0}})$, and we say that $(\mathbf{X}, J)$ is an MAP of arrivals with $(Q, \{\Lambda_\mathbf{n}\}_{\mathbf{n}>\mathbf{0}})$-*source*. This term is a natural extension of the (Q, Λ)-source term commonly used in the literature on MMPPs, where $\Lambda = \Lambda_1$. The stability condition (2.19) has the following implication.

Lemma 1 *If $(\mathbf{X}, J)$ is a stable MAP of arrivals then $(\mathbf{X}, J)$ is explosive if and only if J is explosive.*

Proof: Since J makes no more changes of state than $(\mathbf{X}, J)$, it follows trivially that if J is explosive then $(\mathbf{X}, J)$ is also explosive. Conversely, if J is non-explosive then in a finite time interval J changes state only a finite number of times a.s., and during a subinterval in which J is the same state $\mathbf{X}$ changes state a finite number of times a.s., since $(\mathbf{X}, J)$ is stable. This leads to the conclusion that in finite time $(\mathbf{X}, J)$ changes state a finite number of times a.s., i.e. if J is non-explosive then $(\mathbf{X}, J)$ is also non-explosive. This completes the proof of the statement. □

In applications it makes sense to assume that J is non-explosive, and by Lemma 1 this implies that $(\mathbf{X}, J)$ is non-explosive. Thus in the rest of the paper we consider non-explosive MAPs of arrivals, which in particular are strong Markov, so that the number of arrivals in finite time intervals is a.s. finite, which is the usual condition imposed on arrival processes. For simplicity, we impose a slightly stronger condition for $(\mathbf{X}, J)$, namely

$$\gamma = \sup_{j \in E} \gamma_j = \sup_{j \in E} \sum_{\{(k,\mathbf{n}):(j,k,\mathbf{n}) \in S_r^+(E)\}} \lambda_{jk}(\mathbf{n}) < \infty. \tag{2.23}$$

However we note that (2.23) is not needed for some of the results we give, and that (2.23) holds trivially when E is finite. For $s, t \geq 0$, $j, k \in E$, $\mathbf{n} \in I\!N^r$ the transition probability measure of $(\mathbf{X}, J)$ is such that

$$P_{jk}(\mathbf{n}; t) = P\{\mathbf{X}(t) = \mathbf{n},\ J(t) = k \mid J(0) = j\} \tag{2.24}$$

$$P_{jk}(\mathbf{n}; t) \geq 0, \quad P_{jk}(\mathbf{n}; 0) = \delta_{(\mathbf{0},j)(\mathbf{n},k)}, \quad \sum_{k \in E} P_{jk}(I\!N^r; t) = 1 \tag{2.25}$$

The Chapman-Kolmogorov equations are

$$P_{jk}(\mathbf{n}; t+s) = \sum_{l \in E} \sum_{\mathbf{0} \leq \mathbf{m} \leq \mathbf{n}} P_{jl}(\mathbf{m}; t)\, P_{lk}(\mathbf{n} - \mathbf{m}; s) \tag{2.26}$$

and obviously $\Pi(t) = (\pi_{jk}(t)) = (P_{jk}(I\!N^r; t))$. From (2.18) we have for $h > 0$ and $j \neq k$

$$\pi_{jk}(h) = P_{jk}(I\!N^r; h) = \sum_{\mathbf{n} \in I\!N^r} [\lambda_{jk}(\mathbf{n})h + o(h)] = \sigma_{jk}h + o(h)$$

which shows that Q is the generator matrix of the Markov component J of an MAP of arrivals with $(Q, \{\Lambda_{\mathbf{n}}\}_{\mathbf{n}>\mathbf{0}})$-source. In the rest of the section we give some examples of MAPs of arrivals.

Example 1 *Arrivals, departures and overflow from a Markovian network.*

Suppose we have a Markovian network with r nodes and let $J_i(t)$ $(1 \leq i \leq r)$ be the number of units at node i at time t, and $J = (J_1, J_2, \ldots, J_r)$; suppose also that nodes $j_1, j_2, \ldots, j_s$ $(s \leq r)$ have finite capacity while the rest have infinite capacity. Let Y_i be the number of external units which entered node i, Z_l be the number of units which left the system from node l, and W_p the overflow at node j_p by time t, and denote $\mathbf{X} = (\mathbf{Y}, \mathbf{Z}, \mathbf{W})$ with

$$\mathbf{Y} = (Y_1, Y_2, \ldots, Y_r), \qquad \mathbf{Z} = (Z_1, Z_2, \ldots, Z_r), \qquad \mathbf{W} = (W_1, W_2, \ldots, W_s).$$

The process $(\mathbf{X}, J)$ is an MAP of arrivals on a subset of $I\!N^{3r+s}$. This holds even with batch input and state dependent input, output or routing rates. We could also include arrival components counting the number of units going from one set of nodes to another.

The fact that $(\mathbf{X}, J)$ is an MAP holds in particular for the queueing networks with dependent nodes and concurrent movements studied by Serfozo [37]. Networks are systems with inherent dependencies, which may be either outside the control of the manager of the system or introduced by the manager. Dependencies may be introduced to avoid congestion, balance the worload at nodes, increase the throughput, etc.; thus the study of queueing networks with dependencies is very important for applications. □

Example 2 *Compound phase-type (CPH) arrival processes.*

Consider a continuous time Markov chain $J^\star$ on $\{1, 2, \ldots, m, m+1, \ldots, m+r\}$ with stable and conservative infinitesimal generator matrix

$$Q^\star = \begin{bmatrix} Q^0 & \Lambda^\star \\ 0 & 0 \end{bmatrix} \tag{2.27}$$

with Q^0 being a $m \times m$ matrix, so that states $m+1, \ldots, m+r$ are absorbing. After absorption into state $m+l$ the Markov chain is instantaneously restarted (independently of previous restartings) into transient state k with probability α_{lk}; moreover if absorption is from state j then associate with it an arrival of a batch of size $n > 0$ (independently of the size of other batches) with probability $p_{jl}(n)$.

This model was used by Machihara [22] to model service interruptions in a queueing system, where interruptions are initiated at absorption epochs. Machihara viewed these interruption epochs as the arrival epochs of phase-type Markov renewal customers. If the service interruptions are due to physical failure with the absorption state indicating the type of failure, then it becomes important to record the type of failure (absorption state) which occurs at each failure time (absorption epoch). Accordingly, we let $X_l(t)\,(1 \leq l \leq r)$ be the number of arrivals associated with absorptions into state $m+l$ in $(0,t]$, and let J be the Markov chain on states $E = \{1, 2, \ldots, m\}$ obtained by carrying out the described instantaneous restartings after absorptions and requiring the sample functions of the resulting process to be right-continuous. If we let $\alpha = (\alpha_{lk})$ and for $n \geq 1$

$$P_n = (p_{jl}(n)), \qquad \Psi_n = (\psi_{jl}(n)) = \Lambda^\star \bullet P_n, \qquad Q = Q^0 + \sum_{m \geq 1} \Psi_m \alpha \quad (2.28)$$

then it is easy to see that J has stable and conservative infinitesimal generator matrix Q. Moreover, if we let $\mathbf{X} = (X_1, X_2, \ldots, X_r)$ then $(\mathbf{X}, J)$ is an MAP of arrivals on $I\!N^r \times E$ with $(Q, \{\Lambda_{\mathbf{n}}\}_{\mathbf{n}>\mathbf{0}})$-source where for $\mathbf{n} = (n_1, n_2, \ldots, n_r) > \mathbf{0}$

$$\lambda_{jk}(\mathbf{n}) = \begin{cases} \psi_{jl}(n_l)\,\alpha_{lk} & \sum_{p=1}^r n_p = n_l \\ 0 & \text{otherwise} \end{cases}.$$

Note that arrivals occur only in one of the coordinates of the arrival component at each time. Similarly, if we let $X(t)$ be the total number of arrivals in $(0,t]$, then the process (X, J) is an univariate MAP of arrivals with $(Q, \{\Psi_n \alpha\}_{n>0})$-source. We call (X, J) a CPH arrival process with representation $(Q^0, \Lambda^\star, \{P_n\}_{n>0}, \alpha)$.

For control of the system it is important to know which types of failures occur most frequently in order to minimize the loss due to service interruptions, for which we need to consider the process $(\mathbf{X}, J)$ instead of the CPH arrival process (X, J). This shows that multivariate arrival components may be needed for an appropriate study of some systems. The use of (X, J) is justified only if different failures produce similar effects and have approximately equal costs. □

Remark 1 If (Y, J) is an MAP of arrivals on $I\!N \times \{1, 2, \ldots, m\}$ with $(Q, \{\Lambda_n\}_{n>0})$-source, and we use the lexicographic ordering of the states of (Y, J), then (Y, J) has infinitesimal generator matrix with upper triangular block structure

$$\begin{bmatrix} Q^0 & \Lambda_1 & \Lambda_2 & \Lambda_3 & \ldots \\ 0 & Q^0 & \Lambda_1 & \Lambda_2 & \ldots \\ 0 & 0 & Q^0 & \Lambda_1 & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \end{bmatrix}. \tag{2.29}$$

A BMAP with representation $\{D_k,\ k \geq 0\}$ has generator matrix of the form (2.29) with $D_0 = Q^0$ and $D_n = \Lambda_n\,(n \geq 1)$. For BMAPs it is usually assumed that Q^0 is invertible and Q is irreducible, but we do not view these conditions as essential; moreover even the authors of the BMAP sometimes consider the process without these conditions (e.g. Neuts [26] considers terminating BMAPs, which do not have Q^0 invertible). □

Proposition 1 *The class of CPH arrival processes is equivalent to the class of univariate MAPs of arrivals with finite Markov component (i.e. the class of BMAPs).*

Proof: A CPH arrival process (X, J) with representation $(Q^0, \Lambda^\star, \{P_n\}_{n>0}, \alpha)$ is a univariate MAP of arrivals with $(Q, \{(\Lambda^\star \bullet P_n)\alpha\}_{n>0})$-source, as shown in Example 2, and has finite Markov component.

Conversely, suppose we are given an MAP of arrivals (Y, J) on $I\!N \times \{1, 2, \ldots, m\}$ with $(Q, \{\Lambda_n\}_{n>0})$-source. For $n > 0$ we can obtain matrices $P_n = (p_{jk}(n))$ such that $\Lambda_n = \Lambda \bullet P_n$, with $\{p_{jk}(n),\ n > 0\}$ being a probability function. The process (Y, J) is then indistinguishable from a CPH arrival process with representation $(Q^0, \Lambda, \{P_n\}_{n>0}, I)$. □

Although the use of phase-type arrivals may be important from an operational point of view, in the context of univariate MAPs of arrivals with finite Markov component everything that may be accomplished with phase-type arrivals may be achieved without phase-type arrivals, and vice-versa (as Proposition 1 shows). This implies that the difference between phase-type and non-phase-type arrivals may be merely conceptual, at least if nothing is done to distinguish between these two classes of arrivals. It is also clear that in the class of MAPs of arrivals with multivariate arrival component the individualization of these two classes may be done easily. This suggests some care in the definition of univariate arrival processes which have (conceptually) both phase-type and non-phase-type arrivals. A very good example of the need for a careful definition of these processes is given by the N-process. Since the N-process allows for changes of state of types (i) – (iii) (not associated with phase-type arrivals) in their full generality in case the Markov component is finite, the addition of phase-type arrivals makes the model overparametrized.

2.4 Some properties of MAPs of arrivals

For MAPs of arrivals we now investigate the partial lack of memory property, interarrival times, moments of the number of counts, and a strong law of large numbers.

We consider first the partial lack of memory property. Since the arrival component $\mathbf{X}$ has non-decreasing sample functions (i.e. $\mathbf{X}(t) \geq \mathbf{X}(s)$ a.s. for $0 \leq s \leq t$), the epoch of first arrival becomes

$$T^\circ = \inf\{t : \mathbf{X}(t) > \mathbf{0}\}. \tag{2.30}$$

Theorem 3 *If* $(\mathbf{X}, J)$ *is an MAP of arrivals on* $\mathbb{N}^r \times E$ *with* $(Q, \{\Lambda_\mathbf{n}\}_{\mathbf{n}>\mathbf{0}})$*-source, we have the following.*

(a). $T^\circ > t \Leftrightarrow \mathbf{X}(t) = \mathbf{0}$ *a.s.*

(b). T° *is a stopping time.*

(c). **Partial lack of memory property.**

For $t \in \mathbb{R}_+$ *and* $j, k \in E$ *denote*

$$U^{\mathbf{0}}_{jk}(t) = P\{T^\circ > t,\, J(t) = k \mid J(0) = j\} \tag{2.31}$$

and $U^{\mathbf{0}}(t) = (U^{\mathbf{0}}_{jk}(t))$. *Then the family of matrices* $\{U^{\mathbf{0}}(t),\, t \in \mathbb{R}_+\}$ *forms a semigroup, i.e.*

$$U^{\mathbf{0}}(t+s) = U^{\mathbf{0}}(t)\, U^{\mathbf{0}}(s) \qquad (t, s \in \mathbb{R}_+). \tag{2.32}$$

(d). Moreover $U^{\mathbf{0}}(t) = e^{tQ^{\mathbf{0}}}$, *so that (2.32) reads*

$$e^{(t+s)Q^{\mathbf{0}}} = e^{tQ^{\mathbf{0}}} e^{sQ^{\mathbf{0}}} \qquad (t, s \in \mathbb{R}_+). \tag{2.33}$$

Proof: *(a).* The statement folows easily from the fact that $(\mathbf{X}, J)$ is a Markov subordinator.

(b). The statement is an immediate consequence of *(a)*.

(c). From *(a)*, $U^{\mathbf{0}}_{jk}(u) = P_{jk}(\mathbf{0}; u)$ for $j, k \in E$ and $u \in \mathbb{R}_+$. From the Chapman-Kolmogorov equations (2.26) we find that for $s, t \in \mathbb{R}_+$ and $j, k \in E$

$$U^{\mathbf{0}}_{jk}(t+s) = \sum_{l \in E} U^{\mathbf{0}}_{jl}(t)\, U^{\mathbf{0}}_{lk}(s)$$

which is the semigroup property (2.32).

(d). From (2.31) and (2.18) we have for $j, k \in E$

$$U^{\mathbf{0}}_{jk}(h) = \begin{cases} \lambda_{jk}(\mathbf{0})\, h + o(h) & j \neq k \\ 1 - \gamma_j\, h + o(h) & j = k \end{cases} \tag{2.34}$$

so that $\{U^{\mathbf{0}}(t),\, t \geq 0\}$ is continuous semigroup with associated infinitesimal generator matrix $Q^{\mathbf{0}}$. We note that $U^{\mathbf{0}}(0) = I$, $[U^{\mathbf{0}}(t)]' = U^{\mathbf{0}}(t)\, Q^{\mathbf{0}}$, and

$$\| Q^{\mathbf{0}} \| = \| \Lambda_{\mathbf{0}} - \Gamma \| \leq \| \Lambda_{\mathbf{0}} \| + \| \Gamma \| \leq 2\gamma < \infty$$

which implies that $U^{\mathbf{0}}(t) = e^{tQ^{\mathbf{0}}}$, in view of Lemma 2, which is Proposition 2.5 (Chapter 1) of Golstein [15]. □

Lemma 2 *Let W be a Banach space and $B(W)$ be the space of all bounded linear operators from W to W. If $\eta \in B(W)$, then*

$$\alpha = \left\{ \alpha(t) = e^{t\eta} \doteq \sum_{n=0}^{\infty} \frac{(t\eta)^n}{n!};\ t \in \mathbb{R}_+ \right\} \tag{2.35}$$

is a semigroup satisfying:

$$\| \alpha(t) - I \| \to 0, \qquad \text{as } t \to 0+. \tag{2.36}$$

Moreover, η is the generator of α. Conversely, if α is a semigroup satisfying (2.36), then the generator η of α belongs to $B(W)$ and $\alpha(t) = e^{t\eta}$. □

The partial lack of memory property has been investigated for the MMPP by Prabhu [31] as a natural extension of the well kown lack of memory property of the Poisson process.

From (2.31), we see that $Q^{\mathbf{0}}$ is the generator matrix of the transitions in J not associated with arrivals in $(\mathbf{X}, J)$. In general $U^{\mathbf{0}}$ is dishonest, i.e. $\sum_{k\in E} U^{\mathbf{0}}_{jk}(t) < 1$ for some $j \in E$ and $t > 0$, and

$$1 - \sum_{k\in E} U^{\mathbf{0}}_{jk}(t) = P\{T^{\circ} \leq t \mid J(0) = j\}. \tag{2.37}$$

We note that Q^0 was used as a building block of the BMAP, the CPH arrival process and Yamada and Machihara's arrival process (which is discussed in Example 3).

We define the sucessive arrival epochs of the MAP of arrivals $(\mathbf{X}, J)$

$$T^{\circ}_p = \inf\{t : \mathbf{X}(t) > \mathbf{X}(T^{\circ}_{p-1})\} \qquad (p \geq 1) \tag{2.38}$$

where $T^{\circ}_0 = 0$, so that $T^{\circ}_1 = T^{\circ}$. Owing to the presence of the Markov process J we expect the interarrival times $T^{\circ}_p - T^{\circ}_{p-1}$ to be Markov-dependent.

Theorem 4 Interarrival times.

If we denote $\mathbf{X}^{\circ}_p = \mathbf{X}(T^{\circ}_p)$ *and* $J^{\circ}_p = J(T^{\circ}_p)$, *then* $\{(T^{\circ}_p, \mathbf{X}^{\circ}_p, J^{\circ}_p),\ p \geq 0\}$ *is an MRP whose one-step transition probability density is given by*

$$V(t, \mathbf{n}) = (v_{jk}(t, \mathbf{n})) = e^{tQ^{\mathbf{0}}}\,\Lambda_{\mathbf{n}}. \tag{2.39}$$

Proof: Since $(\mathbf{X}, J)$ is non-explosive it is a strong Markov process. Now since

$$T^{\circ}_{p+1} - T^{\circ}_p = \inf\{t - T^{\circ}_p : \mathbf{X}(t) > \mathbf{X}(T^{\circ}_p)\}$$

$$\mathbf{X}^{\circ}_{p+1} - \mathbf{X}^{\circ}_p = \mathbf{X}(T^{\circ}_{p+1}) - \mathbf{X}(T^{\circ}_p)$$

we see that given $J^{\circ}_0, (T^{\circ}_1, \mathbf{X}^{\circ}_1, J^{\circ}_1), \ldots, (T^{\circ}_p, \mathbf{X}^{\circ}_p, J^{\circ}_p)$, the distribution of $(T^{\circ}_{p+1} - T^{\circ}_p, \mathbf{X}^{\circ}_{p+1} - \mathbf{X}^{\circ}_p, J^{\circ}_{p+1})$ depends only on J°_p and is the same as that of $(T^{\circ}_1, \mathbf{X}^{\circ}_1, J^{\circ}_1)$

given J_0°. This implies that $\{(T_p^\circ, \mathbf{X}_p^\circ, J_p^\circ),\ p \geq 0\}$ is an MRP. Moreover, since the probability of two or more transitions in $(\mathbf{X}, J)$ in a time interval of length h is of order $\circ(h)$ we have for the one-step transition probability density of the MRP

$$\Big(P\{T_{p+1}^\circ - T_p^\circ \in (t, t+dt],\ \mathbf{X}_{p+1}^\circ - \mathbf{X}_p^\circ = \mathbf{n},\ J_{p+1}^\circ = k \mid J_p^\circ = j\}\Big)$$
$$= \Big(\textstyle\sum_{l\in E} P_{jl}(\mathbf{0}; t) P_{lk}(\mathbf{n}; dt)\Big) + \circ(dt) = U^{\mathbf{0}}(t)\Lambda_{\mathbf{n}} dt + \circ(dt)$$

which leads to (2.39), in view of Theorem 3 (d). □

Interarrival times have received much consideration in the applied literature. Some authors have used the MRP $\{(T_p^\circ, \mathbf{X}_p^\circ, J_p^\circ)\}$ as defined in Theorem 4 to characterize special cases of MAPs of arrivals (e.g. Neuts [26]). From Theorem 4 we can define a semi-Markov process $(\mathbf{X}^\star, J^\star)$ by letting

$$(\mathbf{X}^\star, J^\star)(t) = (\mathbf{X}_p^\circ, J_p^\circ) \qquad\qquad (T_p^\circ \leq t < T_{p+1}^\circ).$$

For some special cases of MAPs of arrivals authors have identified the arrival process as being the semi-Markov process $(\mathbf{X}^\star, J^\star)$ (e.g. Lucantoni et al. [21]); this identification is wrong, and may in part be due to the fact that the authors do not state precisely if the process of interest is $(\mathbf{X}, J)$ or just the point process of arrivals. Moreover, if in the modelling stage some authors give a lot of (perhaps too much) importance to the Markov component J they do not seem to give it due importance later in the analysis.

Example 3 *The compound phase-type Markov renewal process (CPH-MRP).*

Consider a CPH arrival process (X, J) (as described in Example 2) with representation $(Q^0, \Lambda^\star, \{P_n\}_{n>0}, \alpha)$. We let $\{T_p^\circ\}$ be the epochs of increments in

X and define $J_p^\circ = J(T_p^\circ)$, $X_p^\circ = X(T_p^\circ)$. Then by Theorem 4 the process $\{(T_p^\circ, X_p^\circ, J_p^\circ),\ p \in I\!N\}$ is an MRP with transition probability density

$$V(t,n) = (v_{jk}(t,n)) = e^{tQ^0}(\Lambda^\star \bullet P_n)\,\alpha.$$

We shall call this process CPH-MRP with representation $(Q^0, \Lambda^\star, \{P_n\}_{n>0}, \alpha)$. Its associated point process of arrivals is $\{(T_p^\circ, X_p^\circ - X_{p-1}^\circ),\ p \geq 1\}$; this process has information about the arrival epochs and the batch sizes of arrivals, but not of the Markov component of the MRP.

Closely related to this is the process defined for simple arrivals by Machihara [22] and for the general case by Yamada and Machihara [43]; in fact in the context of Example 2 the components T_p°, X_p° of the two processes are the same, while the Markov component is $J^\star$, where $J_p^\star$ represents the state the p-th absorption occurs into. This MRP $\{(T_p^\circ, X_p^\circ, J_p^\star)\}$ has transition probability density

$$V^\star(t,n) = \alpha e^{tQ^0}(\Lambda^\star \bullet P_n)$$

and the authors state that it has the representation $(\alpha, Q^0, \Lambda^\star, \{P_n\}_{n>0})$. Yamada an Machihara assume that Q^0 is invertible, but (as for the BMAP in Remark 1) we view their process without this constraint. We note that unless the initial distribution π° of J° and $\pi^\star$ of $J^\star$ are chosen so that $\pi^\circ = \pi^\star\alpha$ the point processes of arrivals associated with $\{(T_p^\circ, X_p^\circ, J_p^\circ)\}$ and $\{(T_p^\circ, X_p^\circ, J_p^\star)\}$ need not be stochastically equivalent.

Our definition of CPH-MRPs does not coincide with the definitions in [22,26, 43]. In Neuts [26] PH-MRPs are MRPs with interarrival times with a phase-type distribution. □

Proposition 2 *The classes of point processes of arrivals associated with CPH-MRPs, Yamada and Machihara's processes, univariate MAPs of arrivals with finite Markov component, and BMAPs are equal.*

Proof: From Proposition 1 it is clear that the classes of point processes of arrivals associated with CPH-MRPs, univariate MAPs of arrivals with finite Markov component, and BMAPs are equal. We now show that the class of point processes of arrivals associated with Yamada and Machihara's processes and univariate MAPs of arrivals with finite Markov component are the same. Suppose we are given an Yamada and Machihara's process with representation $(\alpha, Q^0, \Lambda^\star, \{P_n\}_{n>0})$ and initial distribution $\pi^\star$ for the Markov component. From Examples (2) and (3), it is easy to see that this process has the same associated point process of arrivals as an univariate MAP of arrivals (X, J) with $(Q^0 + \sum_{n\geq 1}(\Lambda^\star \bullet P_n)\alpha, \{(\Lambda^\star \bullet P_n)\alpha\}_{n>0})$-source and initial distribution $\pi^\star\alpha$ for J. Conversely, an univariate MAP of arrivals with finite Markov component, $(Q, \{\Lambda_n\}_{n>0})$-source and initial distribution π for J has the same associated point process of arrivals as an Yamada and Machihara's process with representation $(I, Q^0, \Lambda, \{P_n\}_{n>0})$ and initial distribution π for the Markov component, where P_n is chosen so that $\Lambda_n = \Lambda \bullet P_n$. $\Box$

For $\mathbf{n} \in I\!N^r$ and $\mathbf{z} \in I\!R^r_+$ such that $\mathbf{0} \leq \mathbf{z} \leq \mathbf{e}$, and with $\mathbf{z}^{\mathbf{n}} = \prod_{i=1}^r z_i^{n_i}$, we denote

$$\Phi(\mathbf{z}) = \sum_{\mathbf{n}\in I\!N^r} \Lambda_{\mathbf{n}}\, \mathbf{z}^{\mathbf{n}} \tag{2.40}$$

and

$$G(\mathbf{z};t) = \Big(G_{jk}(\mathbf{z};t)\Big) = \left(\sum_{\mathbf{n}\in I\!N^r} P_{jk}(\mathbf{n};t)\, \mathbf{z}^{\mathbf{n}}\right). \tag{2.41}$$

By considering the process over the time intervals $(0, t]$ and $(t, t + dt]$, we are lead to the following matrix differential equation:

$$\frac{d}{dt} P(\mathbf{n}; t) = -P(\mathbf{n}; t)\,\Gamma + \sum_{\mathbf{0} \le \mathbf{m} \le \mathbf{n}} P(\mathbf{m}; t)\,\Lambda_{\mathbf{n}-\mathbf{m}}. \tag{2.42}$$

In terms of the generating function matrix $G(\mathbf{z}; t)$, this gives

$$\frac{d}{dt} G(\mathbf{z}; t) = G(\mathbf{z}; t)\,[\Phi(\mathbf{z}) - \Gamma]\,. \tag{2.43}$$

Theorem 5 *An MAP of arrivals with $(Q, \{\Lambda_{\mathbf{n}}\}_{\mathbf{n}>\mathbf{0}})$-source has generating function matrix*

$$G(\mathbf{z}; t) = \exp\{t[\Phi(\mathbf{z}) - \Gamma]\} = \exp\left\{ t \left[Q - \sum_{\mathbf{n}>\mathbf{0}} \Lambda_{\mathbf{n}}(1 - \mathbf{z}^{\mathbf{n}}) \right] \right\} \tag{2.44}$$

and $\Pi(t) = e^{tQ}$.

Proof: Since $\Phi(\mathbf{z}) - \Gamma = Q - \sum_{\mathbf{n}>\mathbf{0}} \Lambda_{\mathbf{n}}(1 - \mathbf{z}^{\mathbf{n}})$ and in view of (2.43) and Lemma 2, to prove (2.44) it suffices to show that $\| \Phi(\mathbf{z}) - \Gamma \| < \infty$. But since $\| \Gamma \| = \| \Sigma \| = \gamma$ we have

$$\begin{aligned} \| \Phi(\mathbf{z}) - \Gamma \| &\le \| \Phi(\mathbf{z}) \| + \| \Gamma \| = \| \sum_{\mathbf{n} \in \mathbb{N}^r} \Lambda_{\mathbf{n}} \mathbf{z}^{\mathbf{n}} \| + \gamma \\ &\le \| \sum_{\mathbf{n} \in \mathbb{N}^r} \Lambda_{\mathbf{n}} \| + \gamma = \| \Sigma \| + \gamma = 2\gamma < \infty. \end{aligned}$$

This proves (2.44). The fact that $\Pi(t) = e^{tQ}$ follows from (2.44) by letting $\mathbf{z} \uparrow \mathbf{e}$.

□

Next we give results for the moments of the number of counts of MAPs of arrivals. These have been considered for some particular cases of MAPs of arrivals by different authors; see in particular [11,24,25,31]. For the derivations we use the generating function matrix of MAPs, as given in Theorem 5. For a linear combination with nonnegative integer coefficients $Y = \alpha\mathbf{X}$ of $\mathbf{X}$ and $p \in I\!N$, we let

$$I\!E_p^Y(t) = (E[Y^p(t);\ J(t) = k \mid J(0) = j]) = \left(\sum_{\mathbf{n} \in I\!N^r} (\alpha \mathbf{n})^p \, P_{jk}(\mathbf{n}; t) \right) \tag{2.45}$$

$$\overline{\Sigma}_p^Y = \sum_{\mathbf{n} \in I\!N^r} (\alpha \mathbf{n})^p \Lambda_{\mathbf{n}}. \tag{2.46}$$

Theorem 6 *For* $1 \leq i \leq r$, *if* $\| \overline{\Sigma}_1^{X_i} \| < \infty$, *then*

$$I\!E_1^{X_i}(t) = \int_0^t \Pi(s) \, \overline{\Sigma}_1^{X_i} \, \Pi(t-s) \, ds \tag{2.47}$$

whereas if $\| \overline{\Sigma}_2^{X_i} \| < \infty$, *then*

$$I\!E_2^{X_i}(t) = \int_0^t \left[2 I\!E_1^{X_i}(s) \, \overline{\Sigma}_1^{X_i} + \Pi(s) \, \overline{\Sigma}_2^{X_i} \right] \Pi(t-s) \, ds\,. \tag{2.48}$$

Proof: Since $G(\mathbf{e}; t) = \Pi(t) = e^{tQ}$ and $\left[\frac{d}{dz_i} G(\mathbf{z}; t) \right]_{\mathbf{z}=\mathbf{e}} = I\!E_1^{X_i}(t)$, using (2.43), we find that $I\!E_1^{X_i}(t)$ satisfies the differential equation

$$\frac{d}{dt} \, I\!E_1^{X_i}(t) - I\!E_1^{X_i}(t) \, Q = e^{tQ} \, \overline{\Sigma}_1^{X_i}\,. \tag{2.49}$$

Postmultiplying both members of this equation by e^{-tQ} we get

$$\frac{d}{dt} \left[I\!E_1^{X_i}(t) \, e^{-tQ} \right] = e^{tQ} \, \overline{\Sigma}_1^{X_i} \, e^{-tQ}.$$

In case $\| \overline{\Sigma}_1^{X_i} \| < \infty$ this implies (2.47), by Lemma 2 and the facts that $I\!E_1^{X_i}(0) = (0)$ and $\Pi(t) = e^{tQ}$, $\forall t \geq 0$.

Since $\left[\frac{d^2}{dz_i^2} G(\mathbf{z}; t) \right]_{\mathbf{z}=\mathbf{e}} = I\!E_2^{X_i}(t) - I\!E_1^{X_i}(t)$, again using (2.43), we find that $I\!E_2^{X_i}(t) - I\!E_1^{X_i}(t)$ satisfies the differential equation

$$\frac{d}{dt} \, [I\!E_2^{X_i}(t) - I\!E_1^{X_i}(t)] - [I\!E_2^{X_i}(t) - I\!E_1^{X_i}(t)] \, Q = 2 I\!E_1^{X_i}(t) \, \overline{\Sigma}_1^{X_i} + e^{tQ} \, [\overline{\Sigma}_2^{X_i} - \overline{\Sigma}_1^{X_i}].$$

In view of (2.49) it follows that

$$\frac{d}{dt}\,\mathbb{E}_2^{X_i}(t) - \mathbb{E}_2^{X_i}(t)\,Q = 2\mathbb{E}_1^{X_i}(t)\,\overline{\Sigma}_1^{X_i} + e^{tQ}\,\overline{\Sigma}_2^{X_i}.$$

Proceeding as before with (2.49), and using the fact that $\mathbb{E}_2^{X_i}(0) = (0)$, we obtain (2.48) in case $\| \overline{\Sigma}_2^{X_i} \| < \infty$. □

Corollary 2 *Suppose J is irreducible with stationary distribution $\pi = (\pi_j)$. For the stationary version of $(\mathbf{X}, J)$, and with $1 \le i \le r$, we have the following.*

(a). If $\| \overline{\Sigma}_1^{X_i} \| < \infty$, then

$$(E[X_i(t); J(t) = k]) = t\,\pi\,\overline{\Sigma}_1^{X_i}\,\mathbf{e}\,\pi + \pi\,\overline{\Sigma}_1^{X_i}\left(I - e^{tQ}\right)(\mathbf{e}\,\pi - Q)^{-1} \tag{2.50}$$

and in particular

$$E(X_i(t)) = t\,\pi\,\overline{\Sigma}_1^{X_i}\,\mathbf{e}. \tag{2.51}$$

(b). If $\| \overline{\Sigma}_2^{X_i} \| < \infty$ then

$$\begin{aligned} E(X_i^2(t)) &= \left[t\,\pi\,\overline{\Sigma}_1^{X_i}\,\mathbf{e}\right]^2 + t\,\pi\,\overline{\Sigma}_2^{X_i}\,\mathbf{e} + 2\,t\,\pi\,\overline{\Sigma}_1^{X_i}\,C\,\overline{\Sigma}_1^{X_i}\,\mathbf{e} \\ &\quad -2\,\pi\,\overline{\Sigma}_1^{X_i}\left(I - e^{tQ}\right)(\mathbf{e}\,\pi - Q)^{-2}\,\overline{\Sigma}_1^{X_i}\,\mathbf{e} \end{aligned} \tag{2.52}$$

with $C = (I - \mathbf{e}\,\pi)(\mathbf{e}\,\pi - Q)^{-1}$ and, in particular,

$$\frac{Var(X_i(t))}{t} \to \pi\left[\overline{\Sigma}_2^{X_i} + 2\,\overline{\Sigma}_1^{X_i}\,C\,\overline{\Sigma}_1^{X_i}\right]\mathbf{e}. \tag{2.53}$$

(c). If $\| \overline{\Sigma}_2^{X_i} \| < \infty\,(1 \le i \le r)$, then for $1 \le i, l \le r$,

$$\frac{Cov(X_i(t), X_l(t))}{t} \to \pi\left[\sum_{\mathbf{n}\in\mathbb{N}^r} n_i n_l \Lambda_{\mathbf{n}} + \overline{\Sigma}_1^{X_i}\,C\,\overline{\Sigma}_1^{X_l} + \overline{\Sigma}_1^{X_l}\,C\,\overline{\Sigma}_1^{X_i}\right]\mathbf{e}. \tag{2.54}$$

Proof: (a) and (b). The following identities can be proved easily:

$$\int_0^t \left(e^{sQ} - \mathbf{e}\pi\right) ds = \left(I - e^{tQ}\right) (\mathbf{e}\pi - Q)^{-1} \tag{2.55}$$

$$\int_0^t \int_0^s e^{uQ}\, du\, ds = \frac{1}{2}\mathbf{e}\pi\, t^2 + \left(tI - \int_0^t e^{sQ}\, ds\right) (\mathbf{e}\pi - Q)^{-1}. \tag{2.56}$$

The statements follow by using these identities in Theorem 6, and the fact that

$$\pi\Pi(t) = \pi,\ \forall\, t \geq 0, \qquad \text{and} \qquad \Pi(t) \to \mathbf{e}\,\pi,\ \text{ as } t \to \infty.$$

(c). Proceeding as in Theorem 6 with $z_i = z_l = z$ and taking derivatives with respect to z, we would conclude that

$$I\!\!E_2^{X_i+X_l}(t) = \int_0^t \left[2I\!\!E_1^{X_i+X_l}(s)\,\overline{\Sigma}_1^{X_i+X_l} + \Pi(s)\,\overline{\Sigma}_2^{X_i+X_l}\right] \Pi(t-s)\, ds.$$

By the arguments used to prove (a) and (b) we may then conclude that

$$\frac{Var(X_i(t) + X_l(t))}{t} \to \pi \left[\overline{\Sigma}_2^{X_i+X_l} + 2\left(\overline{\Sigma}_1^{X_i} + \overline{\Sigma}_1^{X_l}\right) C \left(\overline{\Sigma}_1^{X_i} + \overline{\Sigma}_1^{X_l}\right)\right] \mathbf{e}.$$

The statement follows easily from this and (2.53), by using the fact that

$$Cov(X_i(t), X_l(t)) = \frac{1}{2}\left[Var(X_i + X_l(t)) - Var(X_i(t)) - Var(X_l(t))\right]. \qquad \square$$

In the case J has a stationary distribution $\pi = (\pi_j)$ we let

$$\lambda = (\lambda_1, \ldots, \lambda_r), \qquad \lambda_i = \pi\overline{\Sigma}_1^{X_i}\mathbf{e} \quad (1 \leq i \leq r). \tag{2.57}$$

Theorem 7 Law of Large Numbers.

Suppose J is irreducible and has stationary distribution π, $\lambda_i < \infty, \forall i$, then for all initial distributions

$$\frac{\mathbf{X}(t)}{t} \to \lambda \quad a.s. \tag{2.58}$$

Proof: Consider the sequence $\{T_p\}_{p\geq 0}$ of sucessive transition epochs in $(\mathbf{X}, J)$, with $T_0 = 0$. We assume w.l.o.g. that $T_1 < \infty$ a.s., since otherwise $\mathbf{X}(t) \equiv \mathbf{0}$ a.s., $\Lambda_{\mathbf{n}} = 0\,(\mathbf{n} \in I\!N^r)$ and everything is satisfied trivially.

For fixed $i\,(1 \leq i \leq r)$ we define the MRP $\{T_p, Y_p, J_p\}$ with $J_p = J(T_p)$ and $Y_p = X_i(T_p)$ $(p \geq 0)$, then

$$(E_j(Y_1)) = (E(Y_1 \mid J_0 = j)) = \left(\frac{1}{\gamma_j}\sum_{k\in E}\sum_{\mathbf{n}\in I\!N^r} n_i \lambda_{jk}(\mathbf{n})\right) = \Gamma^{-1}\overline{\Sigma}_1^{X_i}\mathbf{e}$$

and $(E_j(T_1)) = (E(T_1 \mid J_0 = j)) = (1/\gamma_j) = \Gamma^{-1}\mathbf{e}$. Moreover $\{J_p\}$ is an irreducible discrete time Markov chain with stationary distribution $\pi^\star = \beta^{-1}\pi\Gamma$, where $\beta = \sum \pi_k \gamma_k$. This implies, by Theorem 12 in Prabhu, Tang and Zhu [33], that for any initial distribution we have

$$\frac{Y_n}{n} \rightarrow \sum_{j\in E} \pi_j^\star E_j(Y_1) = \beta^{-1}\pi\overline{\Sigma}_1^{X_i}\mathbf{e} \quad \text{a.s.} \tag{2.59}$$

$$\frac{T_n}{n} \rightarrow \sum_{j\in E} \pi_j^\star E_j(T_1) = \beta^{-1} \quad \text{a.s..} \tag{2.60}$$

For $t \in I\!R_+$ we let $n(t) = max\{p \in I\!N : T_p \leq t\}$; by (2.23) $n(t) < \infty$ a.s. Moreover $n(t) \uparrow \infty$ a.s. as $t \uparrow \infty$. Now since

$$\frac{Y_{n(t)}}{n(t)}\,\frac{n(t)}{n(t)+1}\,\frac{n(t)+1}{T_{n(t)+1}} = \frac{Y_{n(t)}}{T_{n(t)+1}} \leq \frac{X_i(t)}{t} \leq \frac{Y_{n(t)+1}}{T_{n(t)}} = \frac{Y_{n(t)+1}}{n(t)+1}\,\frac{n(t)+1}{n(t)}\,\frac{n(t)}{T_{n(t)}}$$

using (2.59)-(2.60), we conclude that

$$\frac{X_i(t)}{t} \rightarrow \pi\,\overline{\Sigma}_1^{X_i}\,\mathbf{e} \quad \text{a.s.} \tag{2.61}$$

and (2.58) follows. □

We note that for the stationary version of $(\mathbf{X}, J)$, (2.61) may be obtained from (2.51) and (2.53) by using Tchebychev's inequality and the Borel-Cantelli lemma, as in the proof of the law of large numbers for the Poisson process in Kingman [19].

2.5 Transformations of MAPs of arrivals

In this section we consider transformations of MAPs of arrivals. We study first deterministic transformations and later random transformations.

Considering transformations of MAPs, and for $\mathbf{Z} = D\mathbf{X}$ with $D_{p\times r}$ being a matrix of constants with values in $I\!N$, it is useful to define for $\mathbf{m} > \mathbf{0}$

$$\Lambda_{\mathbf{m}}^{\mathbf{Z}} = \sum_{\{\mathbf{n}\in I\!N^r : D\mathbf{n}=\mathbf{m}\}} \Lambda_{\mathbf{n}}. \tag{2.62}$$

Theorem 8 (a). **Linear transformations of MAPs of arrivals.**

If $(\mathbf{X}, J)$ *is an MAP of arrivals on* $I\!N^r \times E$ *with* $(Q, \{\Lambda_{\mathbf{n}}\}_{\mathbf{n}>\mathbf{0}})$*-source and* $\mathbf{Z} = D\mathbf{X}$ *where* $D_{p\times r}$ *is a matrix of constants with values in* $I\!N$*, then* $(\mathbf{Z}, J)$ *is an MAP of arrivals on* $I\!N^p \times E$ *with* $(Q, \{\Lambda_{\mathbf{m}}^{\mathbf{Z}}\}_{\mathbf{m}>\mathbf{0}})$*-source.*

(b). **Patching together independent MAPs of arrivals.**

Suppose $(\mathbf{X}, J_1)$ *and* $(\mathbf{Y}, J_2)$ *are MAPs of arrivals on* $I\!N^r \times E_1$ *and* $I\!N^s \times E_2$ *with sources* $(Q^1, \{\Lambda_{\mathbf{n}}^1\}_{\mathbf{n}>\mathbf{0}})$ *and* $(Q^2, \{\Lambda_{\mathbf{m}}^2\}_{\mathbf{m}>\mathbf{0}})$*, respectively.*

If $(\mathbf{X}, J_1)$ *and* $(\mathbf{Y}, J_2)$ *are independent then* $((\mathbf{X}, \mathbf{Y}), (J_1, J_2))$ *is an MAP of arrivals on* $I\!N^{r+s} \times E_1 \times E_2$ *with* $(Q^1 \oplus Q^2, \{\Lambda_{(\mathbf{n},\mathbf{m})}\}_{(\mathbf{n},\mathbf{m})>\mathbf{0}})$*-source, where for* $(\mathbf{n}, \mathbf{m}) > \mathbf{0}$

$$\Lambda_{(\mathbf{n},\mathbf{m})} = [\Lambda_{\mathbf{n}}^1\, 1_{\{\mathbf{m}=\mathbf{0}\}}] \oplus [1_{\{\mathbf{n}=\mathbf{0}\}}\, \Lambda_{\mathbf{m}}^2]. \tag{2.63}$$

Proof: (a). The fact that $(\mathbf{Z}, J)$ is an MAP follows from Theorem 1 (a), and since D has entries with values in $I\!N$ it follows that $\mathbf{Z}$ takes values in $I\!N^p$. Thus $(\mathbf{Z}, J)$ is an MAP of arrivals on $I\!N^p \times E$; moreover, using (2.10) and (2.18), if follows that $P(\mathbf{Z}(h) = \mathbf{m},\, J(h) = k \mid J(0) = j)$ is equal to

$$\sum_{\mathbf{n}:D\mathbf{n}=\mathbf{m}} P(\mathbf{X}(h) = \mathbf{n},\, J(h) = k \mid J(0) = j) = \left[\sum_{\mathbf{n}:D\mathbf{n}=\mathbf{m}} \lambda_{jk}(\mathbf{n})\right] h + \circ(h)$$

for $(j,k,\mathbf{m}) \in S_p(E)$. This implies that $(\mathbf{Z},J)$ has $(Q,\{\Lambda^{\mathbf{Z}}_{\mathbf{m}}\}_{\mathbf{m}>\mathbf{0}})$-source.

(b). Using Theorem 1 (b), it follows easily that $((\mathbf{X},\mathbf{Y}),(J_1,J_2))$ is an MAP of arrivals on $I\!N^{r+s} \times E_1 \times E_2$. From the independence of J_1 and J_2, it is well known (and easy to check) that (J_1,J_2) has generator matrix $Q^1 \oplus Q^2$. We let P^1 and P^2 be the transition probability measures of $(\mathbf{X},J_1)$ and $(\mathbf{Y},J_2)$, respectively. For $((j_1,j_2),(k_1,k_2),(\mathbf{n},\mathbf{m})) \in S_{r+s}(E_1 \times E_2)$ the transition probability measure of $((\mathbf{X},\mathbf{Y}),(J_1,J_2))$ is such that

$$\begin{aligned}\left(P_{(j_1,j_2)(k_1,k_2)}((\mathbf{n},\mathbf{m});h)\right) &= \left(P^1_{j_1k_1}(\mathbf{n};h)\,P^2_{j_2k_2}(\mathbf{m};h)\right)\\ &= \left(\lambda^1_{j_1k_1}(\mathbf{n})\,1_{\{\mathbf{m}=\mathbf{0}\}}\,\delta_{j_2k_2} + \delta_{j_1k_1}\,1_{\{\mathbf{n}=\mathbf{0}\}}\lambda^2_{j_2k_2}(\mathbf{m})\right)\,h + o(h)\\ &= \left([\Lambda^1_{\mathbf{n}}\,1_{\{\mathbf{m}=\mathbf{0}\}}] \oplus [1_{\{\mathbf{n}=\mathbf{0}\}}\,\Lambda^2_{\mathbf{m}}]\right)\,h + o(h)\end{aligned}$$

which proves the statement. □

Corollary 3 *For $\alpha,\beta \in I\!N$ we have the following.*

(a). **Marginals of MAPs of arrivals.**

If $(\mathbf{X},J)$ is an MAP of arrivals on $I\!N^r \times E$ with $(Q,\{\Lambda_{\mathbf{n}}\}_{\mathbf{n}>\mathbf{0}})$-source and $\mathbf{Z} = (X_{i_1},\ldots,X_{i_p})$ with $1 \le i_1 < \ldots < i_p \le r$, then $(\mathbf{Z},J)$ is an MAP of arrivals on $I\!N^p \times E$ with $(Q,\{\Lambda^{\mathbf{Z}}_{\mathbf{m}}\}_{\mathbf{m}>\mathbf{0}})$-source, where for $\mathbf{m} > \mathbf{0}$

$$\Lambda^{\mathbf{Z}}_{\mathbf{m}} = \sum_{\{\mathbf{n}:(n_{i_1},n_{i_2},\ldots,n_{i_p})=\mathbf{m}\}} \Lambda_{\mathbf{n}}. \tag{2.64}$$

(b). **Linear combinations of dependent MAPs of arrivals.**

If $\mathbf{X} = (X_1,\ldots,X_r)$, $\mathbf{Y} = (Y_1,\ldots,Y_r)$ and $((\mathbf{X},\mathbf{Y}),J)$ is an MAP of arrivals on $I\!N^{2r} \times E$ with $(Q,\{\Lambda_{(\mathbf{n},\mathbf{m})}\}_{(\mathbf{n},\mathbf{m})>\mathbf{0}})$-source, then the process $(\mathbf{Z},J) = (\alpha\mathbf{X} + \beta\mathbf{Y},J)$ is an MAP of arrivals on $I\!N^r \times E$ with $(Q,\{\Lambda^{\mathbf{Z}}_{\mathbf{a}}\}_{\mathbf{a}>\mathbf{0}})$-source, where for

$\mathbf{a} > \mathbf{0}$ *we have*

$$\Lambda^{\mathbf{Z}}_{\mathbf{a}} = \sum_{\{\mathbf{n},\mathbf{m}\in \mathbb{N}^r : \alpha\mathbf{n}+\beta\mathbf{m}=\mathbf{a}\}} \Lambda_{(\mathbf{n},\mathbf{m})}. \tag{2.65}$$

(*c*). **Linear combinations of independent MAPs of arrivals.**

Suppose $(\mathbf{X}, J_1)$ *and* $(\mathbf{Y}, J_2)$ *are MAPs of arrivals on* $\mathbb{N}^r \times E_1$ *and* $\mathbb{N}^r \times E_2$ *with sources* $(Q^1, \{\Lambda^1_{\mathbf{n}}\}_{\mathbf{n}>\mathbf{0}})$ *and* $(Q^2, \{\Lambda^2_{\mathbf{m}}\}_{\mathbf{m}>\mathbf{0}})$*, respectively.*

If we let $\mathbf{Z} = \alpha\mathbf{X} + \beta\mathbf{Y}$ *and* $J = (J_1, J_2)$*, then* $(\mathbf{Z}, J)$ *is an MAP of arrivals on* $\mathbb{N}^r \times E_1 \times E_2$ *with source* $(Q^1 \oplus Q^2, \{\Lambda^{\mathbf{Z}}_{\mathbf{a}}\}_{\mathbf{a}>\mathbf{0}})$*-source. Here*

$$\Lambda^{\mathbf{Z}}_{\mathbf{a}} = \sum_{\{\mathbf{n},\mathbf{m}\in \mathbb{N}^r : \alpha\mathbf{n}+\beta\mathbf{m}=\mathbf{a}\}} \Lambda_{(\mathbf{n},\mathbf{m})} \tag{2.66}$$

for $\mathbf{a} > \mathbf{0}$*, where* $\Lambda_{(\mathbf{n},\mathbf{m})}$ *is given by (2.63).*

Proof: (*a*) and (*b*). The statements are consequence of Theorem 8 (*a*), since marginals and linear combinations are linear transformations.

(*c*). We first patch together the processes $(\mathbf{X}, J_1)$ and $(\mathbf{Y}, J_2)$. The resulting process $((\mathbf{X}, \mathbf{Y}), (J_1, J_2))$, which is defined on the product probability space of the original processes, is by Theorem 8 (*b*) an MAP of arrivals on $\mathbb{N}^{2r} \times E_1 \times E_2$ with $(Q^1 \oplus Q^2, \{\Lambda_{(\mathbf{n},\mathbf{m})}\}_{(\mathbf{n},\mathbf{m})>\mathbf{0}})$-source, with $\Lambda_{(\mathbf{n},\mathbf{m})}$ given by (2.63). Using this, we then obtain $(\alpha\mathbf{X} + \beta\mathbf{Y}, (J_1, J_2))$, which by (*b*) has $(Q^1 \oplus Q^2, \{\Lambda^{\mathbf{Z}}_{\mathbf{a}}\}_{\mathbf{a}>\mathbf{0}})$-source, with $\Lambda^{\mathbf{Z}}_{\mathbf{a}}$ given by (2.66). □

Theorem 8 and Corollary 3 show that the class of MAPs of arrivals is closed under important transformations. Moreover the transition rates of the transformed processes are easily obtained. We note that Corollary 3 (*b*) and (*c*) could have been stated for linear combinations with a finite number of terms. Corollary 3 (*c*) leads to the following result which has been used to establish certain asymptotic results and to study the effect of multiplexing bursty traffic streams in an ATM network (see [11,20] for references).

Corollary 4 Finite sums of independent MAPs of arrivals.

Suppose $(\mathbf{X}_1, J_1), \ldots, (\mathbf{X}_K, J_K)$ *are independent MAPs of arrivals on* $\mathbb{N}^r \times E_1, \ldots, \mathbb{N}^r \times E_K$ *with sources* $(Q^1, \{\Lambda^1_{\mathbf{n}}\}_{\mathbf{n}>\mathbf{0}}), \ldots, (Q^K, \{\Lambda^K_{\mathbf{n}}\}_{\mathbf{n}>\mathbf{0}})$*, then*

$$(\mathbf{X}_1 + \ldots + \mathbf{X}_K, (J_1, J_2, \ldots, J_K))$$

is an MAP of arrivals on $\mathbb{N}^r \times E_1 \times E_2 \times \ldots E_K$ *with source*

$$\left(Q^1 \oplus Q^2 \oplus \ldots \oplus Q^K, \left\{\Lambda^1_{\mathbf{n}} \oplus \Lambda^2_{\mathbf{n}} \oplus \ldots \oplus \Lambda^K_{\mathbf{n}}\right\}_{\mathbf{n}>\mathbf{0}}\right). \tag{2.67}$$

Proof: It suffices to prove the statement for $K = 2$. Using Corollary 3 (c) we conclude that $(\mathbf{X}_1 + \mathbf{X}_2, (J_1, J_2))$ is an MAP of arrivals on $\mathbb{N}^r \times E_1 \times E_2$ with $(Q_1 \oplus Q_2, \{\Lambda^\star_{\mathbf{n}}\}_{\mathbf{n}>\mathbf{0}})$ where for $\mathbf{n} > \mathbf{0}$

$$\Lambda^\star_{\mathbf{n}} = \Lambda^1_{\mathbf{n}} \oplus 0 + 0 \oplus \Lambda^2_{\mathbf{n}} = \Lambda^1_{\mathbf{n}} \oplus \Lambda^2_{\mathbf{n}}. \quad \square$$

Corollary 4 shows that the class of MAPs of arrivals is closed under finite superpositions of independent processes, which is a generalization of the similar result for Poisson processes and for MMPPs. For some other univariate MAPs of arrivals with finite Markov component the result has been mentioned by some authors without a proof. The proof for the MMPP on Neuts [25] is based essentilly in properties of the Poisson process (this is because the MMPP process may be viewed as constructed from a series of independent Poisson processes such that the i-th Poisson process is observed only when the Markov component is in state i, as remarked in section 1.2). The correspondent proof for more general MAPs of arrivals needs a more sophisticated reasoning. By contrast, our proof is simple, being based more directly on the Markov property of the arrival processes, which is also the basic property of the Poisson process leading to the additive property. In fact the derivation of the properties of MAPs of arrivals may be done without using (or knowing) any properties of Poisson processes.

In computational terms, the rapid (geometric) increase in the dimensionality of the state space of the Markov component of $(\mathbf{X}_1+\mathbf{X}_2+\ldots+\mathbf{X}_K,(J_1,J_2,\ldots,J_K))$ as the number K of added independent MAPs of arrivals increases, limits the utility of Corollary 4. We show in Example 4 that this unpleasant situation may be avoided in the case where the processes added have identical parameters.

Example 4 *Superposition of independent and identical MAPs of arrivals.*

Suppose that the processes $(\mathbf{X}_i, J_i)$ $(1 \leq i \leq K)$ are independent MAPs of arrivals on $\mathbb{N}^r \times \{0,1,\ldots,m\}$ with common rate matrices $\Lambda_{\mathbf{n}}$. If we define

$$J_p^\star(t) = \#\{1 \leq i \leq K : J_i(t) = p\} \qquad (1 \leq p \leq m)$$

then $(J_1^\star, J_2^\star, \ldots, J_m^\star)$ is a Markov process on the state space

$$E = \{(i_1, i_2, \ldots, i_m) \in \mathbb{N}^m : 0 \leq i_1 + i_2 + \ldots + i_m \leq K\}.$$

Moreover, it may be seen that the process, $(\mathbf{X}_1+\mathbf{X}_2+\ldots+\mathbf{X}_K,(J_1^\star,J_2^\star,\ldots,J_m^\star))$ is an MAP of arrivals on $\mathbb{N}^r \times E$ whose transition rates $\lambda^\star_{\mathbf{ij}}(\mathbf{n})$ are such that for $(\mathbf{i},\mathbf{j},\mathbf{n}) \in S_r(E)$

$$\lambda^\star_{\mathbf{ij}}(\mathbf{n}) = \begin{cases} 0 & \#d(\mathbf{i},\mathbf{j}) > 2 \\ i_l \lambda_{lp}(\mathbf{n}) & i_l - j_l = j_p - i_p = 1,\ d(\mathbf{i},\mathbf{j}) = \{l,p\} \\ \sum_{l=0}^m i_l \lambda_{ll}(\mathbf{n}) & \mathbf{i} = \mathbf{j} \end{cases}$$

where

$$i_0 = K - \sum_{l=1}^{m} i_l, \qquad j_0 = K - \sum_{l=1}^{m} j_l, \qquad d(\mathbf{i},\mathbf{j}) = \{l : 0 \le l \le m,\, i_l \ne j_l\}$$

Very particular cases of this example have been considered in a non-comprehensive way by some authors (see e.g. Fischer and Meier-Hellstern [11] and Neuts [25] for superpositioning of two-state MMPPs, which are also called *switched Poisson processes*). □

The reduction in the number of states from $(J_1, \ldots, J_K)$ to $(J_1^\star, \ldots, J_m^\star)$ in Example 4 increases with the number of added processes K and is specially significant when K is large compared with m. In the special case where the individual MAPs of arrivals have a two-state Markov component ($m = 1$) the state space is reduced from 2^K states to $K + 1$ states. This result is important in applications where MAPs of arrivals with a two-state Markov component have been used to model the input from bursty sources in communications systems (see [11,20] for references).

Theorem 9 Random time transformations of MAPs of arrivals.

Suppose that $(\mathbf{X}, J)$ *is an MAP of arrivals on* $\mathbb{N}^r \times E$ *with* $(Q, \{\Lambda_{\mathbf{n}}\}_{\mathbf{n}>\mathbf{0}})$*-source,* f *is a nonnegative function on* E, $A = (a_{jk}) = (f(j)\delta_{jk})$, *and*

$$A(t) = \int_0^t f(J(s))\, ds. \tag{2.68}$$

(a). If we define $(\mathbf{X}^A(t), J^A(t)) = (\mathbf{X}(A(t)), J(A(t)))$ *then the process* $(\mathbf{X}^A, J^A)$ *is an MAP of arrivals on* $\mathbb{N}^r \times E$ *with* $(AQ, \{A\Lambda_{\mathbf{n}}\}_{\mathbf{n}>\mathbf{0}})$*-source.*

(b) Suppose $\inf_{j\in E} f(j) > 0$*. If we define for* $t \in \mathbb{R}_+$

$$B(t) = \inf\{x > 0 : A(x) \ge t\} \tag{2.69}$$

and let $(\mathbf{X}^B(t), J^B(t)) = (\mathbf{X}(B(t)), J(B(t)))$ *then the process* $(\mathbf{X}^B, J^B)$ *is an MAP of arrivals on* $I\!N^r \times E$ *with* $(A^{-1}Q, \{A^{-1}\Lambda_{\mathbf{n}}\}_{\mathbf{n}>\mathbf{0}})$*-source.*

Proof: We note that the probability of two or more transitions in $(\mathbf{X}, J)$ in a time interval with lenght h is $\mathrm{o}(h)$ and that given $J(0) = j$ the probability of no transitions in $(\mathbf{X}, J)$ in $(0, t]$ is equal to $e^{-\gamma_j t}$.

(a). Since $A(t) \in \mathcal{F}_t^J$ and (A, J) is a time-homogeneous Markov subordinator, properties (i) and (ii) of MAPs for $(\mathbf{X}^A, J^A)$ follow directly from the correspondent properties for the MAP $(\mathbf{X}, J)$, and moreover $(\mathbf{X}^A, J^A)$ is time-homogeneous. Thus $(\mathbf{X}^A, J^A)$ is a time-homogeneous MAP with the same state space as $(\mathbf{X}, J)$. We denote by P^A the transition probability measure of $(\mathbf{X}^A, J^A)$. For $(j, k, \mathbf{n}) \in S_r(E)$ we have

$$\begin{aligned} P_{jk}^A(\mathbf{n}; h) &= \int_0^{f(j)h} e^{-\gamma_j u} \lambda_{jk}(\mathbf{n}) e^{-\gamma_k f(k)(h-u)}\, du + \mathrm{o}(h) \\ &= f(j)\lambda_{jk}(\mathbf{n}) h + \mathrm{o}(h) \end{aligned} \tag{2.70}$$

thus, using (2.70), we conclude that $(\mathbf{X}^A, J^A)$ has $(AQ, \{A\Lambda_{\mathbf{n}}\}_{\mathbf{n}>\mathbf{0}})$-source.

(b). First note that since $B(t) \le x \Leftrightarrow A(x) \ge t$, it follows that $B(t)$ is a stopping time for J. Using the fact that $(\mathbf{X}, J)$ is a strong MAP and (A, J) is a time-homogeneous Markov subordinator, this implies that properties (i) and (ii) of MAPs for $(\mathbf{X}^B, J^B)$ follow directly from the correspondent properties for $(\mathbf{X}, J)$, and that $(\mathbf{X}^B, J^B)$ is a time-homogeneous process. Thus $(\mathbf{X}^B, J^B)$ is a time-homogeneous MAP with the same state space as $(\mathbf{X}, J)$. We denote by P^B the transition probability measure of $(\mathbf{X}^B, J^B)$. For $(j, k, \mathbf{n}) \in S_r(E)$ we have

$$P_{jk}^B(\mathbf{n}; h) = \int_0^{[f(j)]^{-1}h} e^{-\gamma_j u} \lambda_{jk}(\mathbf{n}) e^{-\gamma_k \frac{h - f(j)u}{f(k)}}\, du + \mathrm{o}(h)$$

$$= \quad [f(j)]^{-1}\lambda_{jk}(\mathbf{n})h + o(h). \tag{2.71}$$

Using (2.71), we conclude that $(\mathbf{X}^B, J^B)$ has $(A^{-1}Q, \{A^{-1}\Lambda_{\mathbf{n}}\}_{\mathbf{n}>\mathbf{0}})$-source. □

The functional A used in Theorem 9 is a continuous additive functional of the Markov component of (X, J) (see Definition 2). Functionals of the same type as A in (2.68) have have been considered in the applied literature, e.g. Neuts [26]; they may be used in particular to model the service in queueing systems with variable service rate. We note that the Markov additive property would have been preserved in Theorem 9 with $(\mathbf{X}, J)$ being a strong MAP and A being an additive functional of J, not necessarily of the form (2.68).

Corollary 5 *Suppose that $(\mathbf{X}, J_1)$ is an MAP of arrivals on $I\!N^r \times E_1$ with source $(Q^1, \{\Lambda^1_{\mathbf{n}}\}_{\mathbf{n}>\mathbf{0}})$, J_2 is a non-explosive Markov chain on E_2 with generator matrix Q^2, and $(\mathbf{X}, J_1)$ and J_2 are independent. We let $J = (J_1, J_2)$ and assume that for a nonnegative function f on E_2*

$$A(t) = \int_0^t f(J_2(s))\, ds.$$

(a). If we define $(\mathbf{X}^A(t), J^A(t)) = (\mathbf{X}(A(t)), J(A(t)))$, the process $(\mathbf{X}^A, J^A)$ is an MAP of arrivals on $I\!N^r \times E_1 \times E_2$ with $(AQ, \{A\Lambda_{\mathbf{n}}\}_{\mathbf{n}>\mathbf{0}})$-source, where

$$Q = Q^1 \oplus Q^2, \qquad A = \left(f(j_2)\, \delta_{(j_1,j_2)(k_1,k_2)}\right), \qquad \Lambda_{\mathbf{n}} = \Lambda^1_{\mathbf{n}} \oplus 0.$$

(b). If $\inf_{j_2 \in E_2} f(j_2) > 0$ and we let $(\mathbf{X}^B(t), J^B(t)) = (\mathbf{X}(B(t)), J(B(t)))$, with $B(t)$ as in (2.69), then the process $(\mathbf{X}^B, J^B)$ is an MAP of arrivals on $I\!N^r \times E_1 \times E_2$ with source

$$(A^{-1}Q, \{A^{-1}\Lambda_{\mathbf{n}}\}_{\mathbf{n}>\mathbf{0}}).$$

Proof: If we define $O(t) = \mathbf{0} \in \mathbb{N}^r$, it is easy to see that (O, J_2) is an MAP of arrivals on $\mathbb{N}^r \times E_2$ with $(Q^2, \{0\}_{\mathbf{n}>\mathbf{0}})$-source, and is independent of $(\mathbf{X}, J_1)$. Using Theorem 8 (b), it then follows that $(\mathbf{X}, J)$ is an MAP of arrivals on $\mathbb{N}^r \times E$ with source

$$(Q, \{\Lambda_{\mathbf{n}}\}_{\mathbf{n}>\mathbf{0}}) = (Q^1 \oplus Q^2, \{\Lambda^1_{\mathbf{n}} \oplus 0\}_{\mathbf{n}>\mathbf{0}}).$$

The statements now follow by using Theorem 9. □

Neuts [26] considered a random time transformation of particular MRPs associated with the simple BMAP, which is related with the transformations considered in Corollary 5, but Neuts approach does not work for multivariate arrival processes.

If in Corollary 5 $\mathbf{X}(= X)$ is a Poisson process with rate λ, $J_2 = J$ is a Markov chain on $\{1, 2, \ldots, m\}$ with generator matrix Q, and X and J are independent, then (X^A, J^A) is an MMPP with $(AQ, \lambda A)$-source and (X^B, J^B) is an MMPP with $(A^{-1}Q, \lambda^{-1}A)$-source. Thus the transformations described in Theorem 9 transform Poisson processes into MMPPs. It is also easy to see that MMPPs are also transformed into MMPPs, and the same is true for the transformations in Theorem 8 and Corollary 3; this shows that MMPPs have many closure properties.

2.6 Markov-Bernoulli recording of MAPs of arrivals

Secondary recording of an arrival process is a mechanism that from an *original* arrival process generates a *secondary* arrival process. A classical example of secondary recording is the Bernoulli thinning which records each arrival in the original process, independently of all others, with a given probability p or removes it with probability $1-p$. For MAPs of arrivals $(\mathbf{X}, J)$ we consider secondary recordings that leave J unaffected.

A simple example for which the probability of an arrival being recorded varies with time is the case where there is a recording station (or control process) which is *on* and *off* from time to time, so that arrivals in the original process are recorded in the secondary process during periods in which the station is operational (*on*). The control process may be internal or external to the original process and may be of variable complexity (e.g. rules for access of customer arrivals into a queueing network may be simple or complicated and may depend on the state of the network at arrival epochs or not).

Another example is the one in which the original process counts arrivals of batches of customers into a queueing system, while the secondary process counts the number of individual customers, which is more important than the original arrival process in case service is offered to customers one at a time (this is a way in which the compound Poisson process may be obtained from the simple Poisson process). Here the secondary process usually counts more arrivals than the original process, whereas in the previous examples the secondary process always records fewer arrivals than the original process (which corresponds to thinning).

Secondary recording is related with what is called *marking* of the original process (see e.g. Kingman [19]). Suppose that each arrival in the original process is given a *mark* from a space of marks M, independently of the marks given to other arrivals (e.g. for Bernoulli thinning with probability p each point is marked "recorded" with probability p or "removed" with probability $1-p$). If we consider the process that accounts only for arrivals which have marks on a subset C of M, it may be viewed as a secondary recording of the original process. A natural mark associated with cells moving in a network is the pair of origin and destination nodes of the cell. If the original process accounts for traffic generated in the network, then the secondary process may represent the traffic generated between a given pair of nodes, or between two sets of nodes. The marks may represent priorities. These are commonly used in modelling access regulators to communications network systems. A simple example of an access regulator is the one in which cells (arrivals) which are judged in violation of the "contracts" between the network and the user are marked and may be dropped from service under specific situations of congestion in the network; other cells are not marked and carried through the network (see e.g. Elwalid and Mitra [10]). The secondary processes of arrivals of marked and unmarked cells are of obvious interest.

Consider a switch which receives inputs from a number of sources and suppose that the original process accounts for inputs from those sources. If we are interested in studying one of the sources in particular we should observe a secondary process which accounts only for input from this source; this corresponds to a marginal of the original process. In case all sources generate the same type of input, what is relevant for the study of performance of the system is the total input arriving to the switch; this is a secondary process which corresponds to the sum of the coordinates of the original arrival process. This shows that some of the transformations of arrival processes considered in section 2.5 may also be viewed as special cases of secondary recording of an arrival process.

Suppose that $(\mathbf{X}, J)$ is an MAP of arrivals on $\mathbb{N}^r \times E$ with $(Q, \{\Lambda_\mathbf{n}\}_{\mathbf{n}>\mathbf{0}})$-source. We consider a recording mechanism that independently records with probability $r_{jk}(\mathbf{n}, \mathbf{m})$ an arrival in $\mathbf{X}$ of a batch of size $\mathbf{n}$ associated with a transition from j to k in J as an arrival of size $\mathbf{m}$ in the secondary process (with $\mathbf{m} \in \mathbb{N}^s$ for some $s \geq 1$). The operation is identified by the set R of recording probabilities

$$R = \{R_{(\mathbf{n},\mathbf{m})} = (r_{jk}(\mathbf{n}, \mathbf{m})) : \quad \mathbf{n} \in \mathbb{N}^r, \, \mathbf{m} \in \mathbb{N}^s\} \tag{2.72}$$

where $r_{jk}(\mathbf{n}, \cdot)$ is a probability function on $\mathbb{N}^s$, and $r_{jk}(\mathbf{0}, \mathbf{m}) = \delta_{\mathbf{m0}}$. We call this recording *Markov-Bernoulli recording with probabilities in* R and denote the resulting secondary process as $(\mathbf{X}^R, J)$. Thus $(\mathbf{X}^R, J)$ is a process on $\mathbb{N}^s \times E$ which is non-decreasing in $\mathbf{X}^R$, and which increases only when $\mathbf{X}$ increases, i.e.

$$\mathbf{X}^R(t) = \mathbf{X}^R(T_p^\circ) \qquad (T_p^\circ \leq t < T_{p+1}^\circ). \tag{2.73}$$

Moreover, for $\mathbf{n} > \mathbf{0}$

$$r_{jk}(\mathbf{n}, \mathbf{m}) = P\{\mathbf{X}^R(T_{p+1}^\circ) - \mathbf{X}^R(T_p^\circ) = \mathbf{m} \mid A_{jk}(\mathbf{n})\} \tag{2.74}$$

with $A_{jk}(\mathbf{n}) = \{\mathbf{X}(T_{p+1}^\circ) - \mathbf{X}(T_p^\circ) = \mathbf{n}, \, J(T_{p+1}^\circ-) = j, \, J(T_{p+1}^\circ) = k\}$.

Theorem 10 Markov-Bernoulli recording.

Suppose that $(\mathbf{X}, J)$ *is an MAP of arrivals on* $\mathbb{N}^r \times E$ *with* $(Q, \{\Lambda_n\}_{n>0})$*-source and* $R = \{R_{(\mathbf{n},\mathbf{m})}, \, \mathbf{n} \in \mathbb{N}^r, \, \mathbf{m} \in \mathbb{N}^s\}$ *is a set of recording probabilities. We have the following.*

(a). The process $(\mathbf{X}, \mathbf{X}^R, J)$ *is an MAP of arrivals on* $\mathbb{N}^{r+s} \times E$ *with source*

$$(Q, \{\Lambda_\mathbf{n} \bullet R_{(\mathbf{n},\mathbf{m})}\}_{(\mathbf{n},\mathbf{m})>\mathbf{0}}). \tag{2.75}$$

(*b*). *The process* $(\mathbf{X}^R, J)$ *is an MAP of arrivals on* $\mathbb{N}^s \times E$ *with source*

$$(Q, \{\Lambda^R_{\mathbf{m}}\}_{\mathbf{m}>\mathbf{0}}) = \left(Q, \left\{\sum_{\mathbf{n}>\mathbf{0}} \Lambda_n \bullet R_{(\mathbf{n},\mathbf{m})}\right\}_{\mathbf{m}>\mathbf{0}}\right). \tag{2.76}$$

Proof: (*a*). From (2.73) and (2.74), it is easy to see that $(\mathbf{X}, \mathbf{X}^R, J)$ is an MAP of arrivals on $\mathbb{N}^{r+s} \times E$. If $P^{\mathbf{X}}$ is the probability transition measure of $(\mathbf{X}, J)$, the transition probability measure P of $(\mathbf{X}, \mathbf{X}^R, J)$ is such that for $j, k \in E$ and $(\mathbf{n}, \mathbf{m}) > \mathbf{0}$

$$P_{jk}((\mathbf{n}, \mathbf{m}); h) = P^{\mathbf{X}}_{jk}(\mathbf{n}; h)\, r_{jk}(\mathbf{n}, \mathbf{m}) + o(h) \tag{2.77}$$

since the probability of two or more transitions in $(\mathbf{X}, J)$ in time h is of order $o(h)$. In any case, since $r_{jk}(\mathbf{0}, \mathbf{m}) = 0$ for $\mathbf{m} > \mathbf{0}$, (2.77) gives for $(\mathbf{n}, \mathbf{m}) > \mathbf{0}$

$$P_{jk}((\mathbf{n}, \mathbf{m}); h) = \lambda_{jk}(\mathbf{n})\, r_{jk}(\mathbf{n}, \mathbf{m}) h + o(h).$$

This implies that in fact $(\mathbf{X}, \mathbf{X}^R, J)$ has $(Q, \{\Lambda_n \bullet R_{(\mathbf{n},\mathbf{m})}\}_{(\mathbf{n},\mathbf{m})>\mathbf{0}})$-source.

(*b*). The statement follows from (*a*) using Corollary 3 (*a*). □

We give now two examples of Markov-Bernoulli recording. In Example 5 we view the overflow process from a state dependent $M/M/1/K$ system as a special case of secondary recording of the arrival process of customers to the system, and in Example 6 we give a more elaborated example of secondary recording.

Example 5 Overflow from a state dependent $M/M/1/K$ system

We consider a Markov-modulated $M/M/1/K$ system with batch arrivals with (independent) size distribution $\{p_n\}_{n>0}$. When there is an arrival of a batch with n customers and only $m < n$ positions are available only m customers from the batch enter the system. Assume that the service rate is μ_j and the arrival rate of batches is α_j, whenever the number of customers in the system is j. We let $J(t)$ be the number of customers in the system at time t and $X(t)$ be the number of customer arrivals in $(0,t]$.

The process (X, J) is an MAP of arrivals on $\mathbb{N} \times \{0, 1, \ldots, K\}$ with rates

$$\lambda_{jk}(n) = \begin{cases} \alpha_j\, p_n & k = \min(j+n, K) \\ \mu_j & k = j-1,\, n = 0 \\ 0 & \text{otherwise} \end{cases}.$$

Its source is $(Q, \{\Lambda_n\}_{n>0})$, where Q is obtained from the matrices $\{\Lambda_n\}_{n\geq 0}$. If we define $X^R(t)$ as the overflow from the system in $(0,t]$, then it is readily seen that (X^R, J) is a Markov-Bernoulli recording of (X, J) with recording probabilities

$$r_{jk}(n,m) = 1_{\{m=n-(k-j)\}} \qquad (n, m > 0).$$

Thus (X^R, J) is an MAP of arrivals on $\mathbb{N} \times \{0, 1, \ldots, K\}$ with $(Q, \{\Lambda^R_m\}_{m>0})$-source, where $\lambda^R_{jk}(m) = \alpha_j\, 1_{\{k=K\}}\, p_{m+(K-j)}$ for $m > 0$. □

Example 6 Suppose $\mathbf{X} = (X_1, \ldots, X_r)$, $\mathbf{Y} = (Y_1, \ldots, Y_s)$ and $(\mathbf{X}, \mathbf{Y}, J_2)$ is an MAP of arrivals on $\mathbb{N}^{r+s} \times E_2$ with $(Q, \{\Lambda_{(\mathbf{n},\mathbf{m})}\}_{(\mathbf{n},\mathbf{m})>\mathbf{0}})$-source. We are interested in keeping only the arrivals in $\mathbf{X}$ which are preceeded by an arrival in $\mathbf{Y}$ without a simultaneous arrival in $\mathbf{X}$. Denote by $\mathbf{X}^R$ be the arrival counting process which we obtain by this operation. For simplicity, we assume that E_2 is finite and $\Lambda_{(\mathbf{n},\mathbf{m})} = 0$ if $\mathbf{n} > \mathbf{0}$ and $\mathbf{m} > \mathbf{0}$.

We let $\{T^{\mathbf{X}}_p\}_{p\geq 0}$ and $\{T^{\mathbf{Y}}_p\}_{p\geq 0}$ be the sucessive arrival epochs in $\mathbf{X}$ and $\mathbf{Y}$, respectively, denote $J = (J_1, J_2)$ with

$$J_1(t) = \begin{cases} 1 & \max\{T^{\mathbf{Y}}_p : T^{\mathbf{Y}}_p \leq t\} > \max\{T^{\mathbf{X}}_p : T^{\mathbf{X}}_p \leq t\} \\ 0 & \text{otherwise} \end{cases}.$$

It can be checked in a routine fashion that $((\mathbf{X}, \mathbf{Y}), J)$ is an MAP of arrivals on $I\!N^{r+s} \times \{0,1\} \times E_2$ with $(Q^\star, \{\Lambda^\star_{(\mathbf{n},\mathbf{m})}\}_{(\mathbf{n},\mathbf{m})>\mathbf{0}})$-source (with the states of (J_1, J_2) ordered in lexicographic order), where $\Lambda^\star_{(\mathbf{n},\mathbf{m})} = 0$ if $\mathbf{n} > \mathbf{0}$ and $\mathbf{m} > \mathbf{0}$, and

$$Q^\star = \begin{bmatrix} Q - \sum_{\mathbf{m}>\mathbf{0}} \Lambda_{(\mathbf{0},\mathbf{m})} & \sum_{\mathbf{m}>\mathbf{0}} \Lambda_{(\mathbf{0},\mathbf{m})} \\ \sum_{\mathbf{n}>\mathbf{0}} \Lambda_{(\mathbf{n},\mathbf{0})} & Q - \sum_{\mathbf{n}>\mathbf{0}} \Lambda_{(\mathbf{n},\mathbf{0})} \end{bmatrix}, \qquad \Lambda^\star_{(\mathbf{n},\mathbf{m})} = \begin{bmatrix} \Lambda_{(\mathbf{n},\mathbf{0})} & \Lambda_{(\mathbf{0},\mathbf{m})} \\ \Lambda_{(\mathbf{n},\mathbf{0})} & \Lambda_{(\mathbf{0},\mathbf{m})} \end{bmatrix}$$

if either $\mathbf{n} = \mathbf{0}$ or $\mathbf{m} = \mathbf{0}$. It is easy to see that $(\mathbf{X}^R, J)$ is a Markov-Bernoulli recording of $((\mathbf{X}, \mathbf{Y}), J)$ with recording probabilities

$$r_{(p,j)(q,k)}((\mathbf{n}, \mathbf{m}), \mathbf{l}) = 1_{\{p=1,\, \mathbf{n}=\mathbf{l}\}}$$

for $(\mathbf{n}, \mathbf{m}) > \mathbf{0}$. Thus $(\mathbf{X}^R, J)$ is an MAP of arrivals on $I\!N^r \times \{0,1\} \times E_2$ with $(Q^\star, \{\Lambda^R_{\mathbf{n}}\}_{\mathbf{n}>\mathbf{0}})$-source, where

$$\Lambda^R_{\mathbf{n}} = \begin{bmatrix} 0 & 0 \\ \Lambda_{(\mathbf{n},\mathbf{0})} & 0 \end{bmatrix}.$$

A particular case of this example with $r = 1$ and Y being a Poisson process was considered briefly by Neuts [26]. □

We now consider the special case of *Markov-Bernoulli marking* for which the space of marks M is discrete. Specifically, each arrival of a batch of size $\mathbf{n}$ associated with a transition from state j to state k in J is given a mark $m \in M$ with *(marking)* probability $c_{jk}(\mathbf{n}, m)$, independently of the marks given to other arrivals.

If we assume w.l.o.g. that $M = \{0, 1, \ldots, K\} \subseteq I\!N$, Markov-Bernoulli marking becomes a special case of Markov-Bernoulli recording for which the *marked process*

$(\mathbf{X}, \mathbf{X}^R, J)$ is such that $(\mathbf{X}^R, J)$ is a Markov-Bernoulli recording of $(\mathbf{X}, J)$ with recording probabilities

$$R_{(\mathbf{n},m)} = (r_{jk}(\mathbf{n}, m)) = (1_{\{m \in M\}}\, c_{jk}(\mathbf{n}, m))$$

for $\mathbf{n} > \mathbf{0}$ and $m \in I\!N$.

Theorem 11 *With the conditions described we have the following.*

(a). **Markov-Bernoulli marking.**

The marked process $(\mathbf{X}, \mathbf{X}^R, J)$ *is an MAP of arrivals on* $I\!N^{r+1} \times E$ *with source*

$$(Q, \{\Lambda_{\mathbf{n}} \bullet R_{(\mathbf{n},m)}\}_{(\mathbf{n},m)>\mathbf{0}}). \tag{2.78}$$

(b). **Markov-Bernoulli colouring.**

Suppose that the set of marks is finite. i.e. $K < \infty$. *We identify mark* $m \in M$ *as colour* m, *and let* $\mathbf{X}^{(m)}$ *be the counting process of arrivals of batches coloured* m, *for* $0 \le m \le K$. *The process* $(\mathbf{X}^{(0)}, \ldots, \mathbf{X}^{(K)}, J)$ *is an MAP of arrivals on* $I\!N^{(K+1)r} \times E$ *with* $(Q, \{\Lambda^\star_{\mathbf{a}}\}_{\mathbf{a}>\mathbf{0}})$*-source. Here for* $\mathbf{a} = (\mathbf{a}^{(0)}, \ldots, \mathbf{a}^{(K)}) > \mathbf{0}$ *with* $\mathbf{a}^{(m)} \in I\!N^r$ $(0 \le m \le K)$

$$\Lambda^\star_{\mathbf{a}} = \left(\sum_{m=0}^{K} 1_{\{A_{\mathbf{a}} = \{\mathbf{a}^{(m)}\}\}}\, \Lambda_{\mathbf{a}^{(m)}} \bullet R_{(\mathbf{a}^{(m)},m)} \right) \tag{2.79}$$

where $A_{\mathbf{a}} = \{\mathbf{a}^{(l)} : 0 \le l \le K,\ \mathbf{a}^{(l)} > \mathbf{0}\}$.

Proof: (a). The statement is a consequence of Theorem 10 (a).

(b). The statement follows from (a), Theorem 10 (b), and the fact that the process $(\mathbf{X}^{(0)}, \ldots, \mathbf{X}^{(K)}, J)$ is a Markov-Bernoulli recording of $(\mathbf{X}, \mathbf{X}^R, J)$ with recording probability matrices

$$R^\star_{((\mathbf{n},m)(\mathbf{n}^{(0)},\ldots,\mathbf{n}^{(K)}))} = \left(1_{\{\mathbf{n}^{(m)} = \mathbf{n}, \mathbf{n}^{(l)} = \mathbf{0}\ (0 \le l \le K, l \ne m)\}}\right)$$

for $0 \leq m \leq K$ and $\mathbf{n}, \mathbf{n}^{(0)}, \ldots, \mathbf{n}^{(K)} \in \mathbb{N}^r$, with $(\mathbf{n}, m) > \mathbf{0}$. $\square$

Theorem 11 (*b*) is an extension of the Colouring Theorem for Poisson processes (see Kingman ([19], Section 5.1)). An important consequence of Theorem 11 is the following result.

Corollary 6 *Suppose* $(\mathbf{X}, J)$ *is an MAP of arrivals on* $\mathbb{N}^r \times E$ *with* $(Q, \{\Lambda_{\mathbf{n}}\}_{\mathbf{n}>\mathbf{0}})$*-source, and* $\{B_m\}_{0 \leq m \leq K}$ *is a finite partition of* $\mathbb{N}^r - \{\mathbf{0}\}$.

If for $0 \leq m \leq K$ *we let* $\mathbf{X}^{(m)}$ *be the counting process of arrivals with batch sizes in* B_m*, then* $(\mathbf{X}^{(0)}, \ldots, \mathbf{X}^{(K)}, J)$ *is an MAP of arrivals on* $\mathbb{N}^{(K+1)r} \times E$ *with* $(Q, \{\Lambda^\star_{\mathbf{a}}\}_{\mathbf{a}>\mathbf{0}})$*-source, where for* $\mathbf{a} = (\mathbf{a}^{(0)}, \ldots, \mathbf{a}^{(K)}) > \mathbf{0}$*, with* $\mathbf{a}^{(m)} \in \mathbb{N}^r$ $(0 \leq m \leq K)$

$$\Lambda^\star_{\mathbf{a}} = \begin{cases} \Lambda_{\mathbf{a}^{(m)}} & A_{\mathbf{a}} = \{\mathbf{a}^{(m)}\} \subseteq B_m \\ 0 & \textit{otherwise} \end{cases} \tag{2.80}$$

where $A_{\mathbf{a}} = \{\mathbf{a}^{(l)} : 0 \leq l \leq K,\ \mathbf{a}^{(l)} > \mathbf{0}\}$.

Proof: The statement follows from Theorem 11 (*b*) by considering a Markov-Bernoulli marking of $(\mathbf{X}, J)$ with marking probabilities $c_{jk}(\mathbf{n}, m) = 1_{\{\mathbf{n} \in B_m\}}$. $\square$

Note that the processes $(\mathbf{X}^{(p)}, J)$ and $(\mathbf{X}^{(q)}, J)$ with $p \neq q$ have no common arrival epochs a.s. When J has more than one state, $\mathbf{X}^{(p)}$ and $\mathbf{X}^{(q)}$ $(0 \leq p, q \leq K)$ are not independent, except perhaps in very special cases. In case J has only one state and $r = 1$, X is a compound Poisson process and Corollary 6 implies that $X^{(m)}$ $(0 \leq m \leq K)$ are independent compound Poisson processes.

When the primary and secondary processes are defined on the same state space and the recording probabilities are such that $R_{(\mathbf{n},\mathbf{m})} = 0$ for $\mathbf{m} > \mathbf{n}$ we say that the associated secondary recording is a *thinning*, and the secondary process is a *thinned* *process.* If $\mathbf{X}$ and $\mathbf{X}^R$ are the counting processes of arrivals in the original and the thinned process we may also consider the process $\mathbf{X}^L = \mathbf{X} - \mathbf{X}^R$ which counts lost arrivals. From Theorem 10 we know that the *loss* process $(\mathbf{X}^L, J)$ is also an MAP of arrivals, since we may interchange recorded arrivals with non-recorded arrivals. However for Markov-Bernoulli thinning we are able to give sharper results.

Theorem 12 Markov-Bernoulli thinning

Suppose that $(\mathbf{X}, J)$ *is an MAP of arrivals on* $\mathbb{N}^r \times E$ *with* $(Q, \{\Lambda_{\mathbf{n}}\}_{\mathbf{n}>0})$*-source and* $R = \{R_{(\mathbf{n},\mathbf{m})}\}$ *is a set of thinning probabilities. Then the process* $(\mathbf{X}^R, \mathbf{X}^L, J)$ *is an MAP of arrivals on* $\mathbb{N}^{2r} \times E$ *with source*

$$(Q, \{\Lambda_{\mathbf{n}+\mathbf{m}} \bullet R_{(\mathbf{n}+\mathbf{m},\mathbf{n})}\}_{(\mathbf{n},\mathbf{m})>\mathbf{0}}). \tag{2.81}$$

Proof: Using arguments similar to the ones in the proof of Theorem 10 (a), we may conclude that the process $(\mathbf{X}, \mathbf{X}^R, \mathbf{X}^L, J)$ is an MAP of arrivals on $\mathbb{N}^{3r} \times E$ with source

$$\left(Q, \left\{[\Lambda_{\mathbf{n}} \bullet R_{(\mathbf{n},\mathbf{m})}]\, 1_{\{\mathbf{p}=\mathbf{n}-\mathbf{m}\}}\right\}_{(\mathbf{n},\mathbf{m},\mathbf{p})>\mathbf{0}}\right). \tag{2.82}$$

Using Corollary 3 (a), the statement follows. □

Example 7 Suppose that X is a Poisson process with rate λ, J is a stable finite Markov chain with generator matrix Q, and X and J are independent. We view J as the state of a recording station, and assume that each arrival in J is recorded, independently of all other arrivals, with probability p_j or non-recorded with probability $1 - p_j$, whenever the station is in state j. We let $X^R(t)$ be the number of recorded arrivals in $(0, t]$.

Using Theorem 8 (*b*), Theorem 12, and Corollary 3 (*a*) we may conclude that (X^R, J) is an MMPP with $(Q, (\lambda\, p_j\, \delta_{jk}))$-source. In the special case where the probabilities p_j are either 1 or 0, the station is *on* and *off* from time to time, with the distributions of the on and off periods being Markov dependent. □

Chapter 3

Properties of the storage model

In this chapter we study properties of the storage model described in section 1.4. The input process (X, J) and the demand process (D, J) are MAPs given by

$$X(t) = X_0(t) + \int_0^t a\,(J(s))\,ds, \qquad D(t) = \int_0^t d(J(s))\,ds \tag{3.1}$$

with (X_0, J), the part of the input in $(0, t]$ associated with jumps, being a pure jump process, and a and d being nonnegative functions. We define d_- and d_+ as follows:

$$0 \leq d_- = \inf_{j \in E} d(j) \leq \sup_{j \in E} d(j) = d_+ \leq +\infty. \tag{3.2}$$

The storage equation of the model is

$$Z(t) = Z(0) + Y(t) + \int_0^t y(J(s))^- 1_{\{Z(s) \leq 0\}}\,ds \tag{3.3}$$

where Y is the net input, and $y = a - d$. The Markov component of the input and demand processes J is an MJP, which from (1.14) has infinitesimal generator

$$\mathcal{A}_J f(j) = \lambda_j \sum_{k \in E} [f(k) - f(j)]\, p_{jk}, \qquad j \in E \tag{3.4}$$

where

$$\Sigma = (\sigma_{jk}) = \Lambda\, P = (\lambda_j\, \delta_{jk})\, (p_{jk}) \tag{3.5}$$

with P being a stochastic matrix. We call Σ the *transition rate matrix*, the diagonal matrix Λ the *intensity matrix*, and P the *transition probability matrix* of J. In case J is ergodic, we denote by $\pi = \{\pi_j,\, j \in E\}$ its stationary distribution and let $\Pi = \mathbf{e}\pi$, so that

$$\Pi = \left(\lim_{t\to\infty} P[J(t) = k \mid J(0) = j]\right).$$

We let $\tilde{J} = \{J_n,\, n \geq 0\}$ be the associated discrete time Markov chain (DTMC) with (stationary) transition probability matrix P. We may then consider $J_n = J(T_{n+1}-)$, $(n \geq 0)$ with

$$T_n = \text{ time at which the n-th transition in } J \text{ occurs.}$$

We solve (3.3) in section 3.1. In section 3.2 we investigate the properties of the actual and unsatisfied demands, and in section 3.3 the inverse of the actual demand. If we denote by $D(t)$ the total demand in $(0, t]$ and by $B(x)$ its inverse, then the main result is that $\{(B, J \circ B)(x),\, x \geq 0\}$ is an MAP consisting of just a drift.

3.1 Solution of the storage integral equation

We assume that the input process (X, J) is defined on some probability space $(\Omega, \mathcal{F}, \mathcal{P})$.

Theorem 13 *With the conditions described, we have ($\mathcal{P}$-a.s.):*

(a). The integral equation (3.3) has the unique solution

$$Z(t) = Z(0) + Y(t) + I(t) \tag{3.6}$$

where

$$I(t) = [Z(0) + m(t)]^- = \left[Z(0) + \inf_{0 \le \tau \le t} Y(\tau)\right]^- . \tag{3.7}$$

(b). D and Y are $\mathcal{F}$-measurable processes, and Z and I are $(\mathcal{F} \bigvee \sigma(Z(0)))$-measurable.

(c). D and I have continuous sample paths, and X, Y and Z have càdlàg *sample paths with all jumps being positive.*

(d). If in addition we are given a filtration $\{\mathcal{F}_t, t \ge 0\}$ and (X, J) is adapted to the filtration $\{\mathcal{F}_t\}$ then, in addition to the results of $(a) - (c), J, X, D$ and Y are progressively measurable with respect to $\{\mathcal{F}_t\}$, and I and Z are progressively measurable with respect to $\{\mathcal{F}_t \bigvee \sigma(Z(0))\}$.

(e). If $\mathcal{F}_t^o = \mathcal{F}_t^{(X,J)}$ and $\mathcal{F}_t$ is the proper completion of $\mathcal{F}_t^o$, then D and Y are progressively measurable with respect to $\{\mathcal{F}_t\}$, and I and Z are progressively measurable with respect to the proper completion of $\{\mathcal{F}_t^o \bigvee \sigma(Z(0))\}$.

Proof: During the proof, it should be kept in mind that $\mathcal{P}$-a.s.: *(i)* (X, J) has *càdlàg* sample paths, *(ii)* in a finite time interval, the number of jumps of X is finite (since X is non-explosive) and *(iii)* all jumps of X are positive. For simplicity we omit $\mathcal{P}$-a.s. in the proof.

(a). This proof is based on Prabhu [30]. Let ω, a point in the sample space satisfying *(i)-(iii)*, and $t > 0$ be fixed. From the setting of the theorem and equation (3.3) we easily see that $Z(t) \geq 0, \forall t$ and for $\tau \in [0,t]$, with $Z(0-) = Z(0)$,

$$\begin{aligned} Z(t) &\geq Z(t) - Z(\tau-) \\ &= Y(t) - Y(\tau-) + \int_\tau^t d(J(s))\, 1_{\{Z(s)\leq 0\}}\, ds \geq Y(t) - Y(\tau-). \quad (3.8) \end{aligned}$$

If $Z(\tau)$ is never zero for $\tau \in (0,t)$ then from (3.3) and (3.8)

$$Z(t) = Z(0) + Y(t) \geq \sup_{0\leq\tau\leq t} [Y(t) - Y(\tau-)]; \quad (3.9)$$

otherwise let $t_0 = \max\{\tau : 0 < \tau \leq t, Z(\tau-) = 0\}$ and use (3.8) and (3.3) to get

$$Z(t) \geq \sup_{0\leq\tau\leq t} [Y(t) - Y(\tau-)] \geq [Y(t) - Y(t_0-)] = Z(t). \quad (3.10)$$

Since $Z(t) \geq Z(0) + Y(t), \forall t \geq 0$, using (3.9) and (3.10), it follows easily that

$$\begin{aligned} Z(t) &= \max\{Z(0) + Y(t), \sup_{0\leq\tau\leq t} [Y(t) - Y(\tau-)]\} \\ &= \max\{Z(0) + Y(t), Y(t) - m(t)\} \quad (3.11) \end{aligned}$$

which with (3.3) implies that

$$I(t) = \int_0^t d(J(s))\, 1_{\{Z(s)\leq 0\}}\, ds = Z(t) - Z(0) - Y(t) = \max\{0, -m(t) - Z(0)\}$$

as required.

(b). Since (X, J) is right-continuous it is also $\mathcal{F}$-measurable by Proposition 1.13 and Remark 1.14 (Chapter 1) in Karatzas and Shreve [17]. The $\mathcal{F}$-measurability of J implies the $\mathcal{F}$-mesurability of $d(J(t))$ which in turn implies the $\mathcal{F}$-measurability of D by Remark 4.6(i) (Chapter 1) in Karatzas and Shreve [17], and therefore $Y = X - D$ is also $\mathcal{F}$-measurable as does $m = \{m(t), t \geq 0\}$ given by (3.7). Since m is also $(\mathcal{F} \bigvee \sigma(Z(0)))$-measurable so are I, by (3.7), and Z by (3.6).

(*c*). Since J has a finite number of jumps in a finite interval, the continuity of the sample paths of D and I follows. Since X has *càdlàg* non-decreasing sample paths and D has continuous sample paths, $Y = X - D$ has *càdlàg* sample paths with all jumps being positive; this and the fact that I has continuous sample paths imply by (3.6) that Z has *càdlàg* sample paths with all jumps being positive.

(*d*). The proof follows along the same lines of (*b*); the fact that (X, J) is progressively measurable with respect to $\{\mathcal{F}_t\}$ follows from the fact that it is adapted to $\{\mathcal{F}_t\}$ and has *càdlàg* sample functions, by Proposition 1.13 (Chapter 1) in Karatzas and Shreve [17].

(*e*). The statement follows from (*d*). □

3.2 The actual and the unsatisfied demands

We recall the definition of an additive functional of a Markov process (see Blumenthal [3]).

Definition 2 *Let $\{R(t),\, t \geq 0\}$ be a Markov process, $\mathcal{F}_t^o = \mathcal{F}_t^R$ and $\mathcal{F}_t$ the proper completion of $\mathcal{F}_t^o$. Then $\{\mathcal{F}_t, t \geq 0\}$ is a right-continuous filtration for $\{R(t),\, t \geq 0\}$.*

A family $\{A(t),\, t \geq 0\}$ of functions with values in $[0, \infty]$ is called an additive functional of $\{R(t),\, t \geq 0\}$ provided:

(i) $t \rightarrow A(t), t \geq 0$ is non-decreasing, right-continuous and $A(0) = 0$, a.s. for all initial distributions.

(ii) For each $t, A(t) \in \mathcal{F}_t$.

(iii) *For each* $s, t \geq 0, A(t+s) = A(t) + A(s) \circ \phi_t$, *a.s. for any initial distribution, where* ϕ *is a shift operator,*

$$\phi_t : \Omega \to \Omega \text{ such that } (\phi_t(\omega))(r) = \omega(r+t),\ r \geq 0. \ \Box$$

Remark 2 With the setting of Definition 2, a functional A of $\{R(t),\ t \geq 0\}$ of the form

$$A(t) = \int_0^t f(R(s))\, ds, \tag{3.12}$$

where f is a bounded nonnegative measurable function, is called a *classical* additive functional by Blumenthal [3]. As remarked in Blumenthal and Gettor [4] (p.151), A is still an additive functional of R even when f is not bounded provided $A(t)$ is finite for all t a.s. $\Box$

Lemma 3 *The demand process* $\{D(t),\ t \geq 0\}$ *has the following properties:*

(*a*). *D is a continuous additive functional of the MJP J.*

(*b*). *D has strictly increasing sample paths, for all initial distributions.*

(*c*). *If $d_- > 0$, then $D(\infty) = \lim_{t\to\infty} D(t) = \infty$, for all initial distributions.*

(*d*). $d_- \leq \frac{D(t)}{t} \leq d_+$, *for all* $t > 0$.

(*e*). *If J is ergodic, then* $\lim_{t\to\infty} \frac{D(t)}{t} = \sum_{j\in E} d(j)\,\pi_j$, *a.s.*

(*f*). *If J has a persistent state, then* $\lim_{t\to\infty} D(t) = \infty$, *a.s.*

Proof: (*a*). Since J is non-explosive, J has a finite number of transitions in finite time a.s. and this implies that $D(t) < \infty$ a.s., $\forall t \geq 0$; the statement now follows from Remark 2.

(*b*). Since $d(j) > 0, \forall j \in E$, then all sample paths of $\{D(t),\ t \geq 0\}$ are strictly increasing regardless of the initial distribution and the result follows.

(c). The statement follows since, if $d_- > 0$, for any $t \geq 0$ and any initial distribution,

$$D(t) = \int_0^t d(J(s))\, ds \geq \int_0^t d_-\, ds = d_-\, t.$$

(d). Since for all $t > 0$, $d_- t \leq D(t) \leq d_+ t$, the statement follows.

(e). d is a positive measurable function on E, so that, if J is ergodic,

$$\begin{aligned} \frac{D(t)}{t} &= \frac{1}{t} \int_0^t d(J(s))\, ds = \frac{1}{t} \int_0^t \sum_{j \in E} d(j)\, 1_{\{J(s)=j\}}\, ds \\ &= \sum_{j \in E} \frac{d(j)}{t} \int_0^t 1_{\{J(s)=j\}}\, ds \to \sum_{j \in E} d(j)\, \pi_j \text{ , as } t \to \infty, \text{ a.s.} \end{aligned}$$

(f). Given a persistent state $j \in E$ (and since $d(j) > 0$),

$$D(t) = \int_0^t d(J(s))\, ds \geq \int_0^t d(j)\, 1_{\{J(s)=j\}}\, ds \to \infty \text{ , as } t \to \infty, \text{ a.s.}$$

and the statement follows. □

The demand process can be identified as a random clock, its speed at t being $d(J(t))$, and with transformed time scale $[0, D(\infty))$. Lemma 3 (c) and (f) gives sufficient conditions for the transformed time scale to be infinite a.s.; we note the condition $d(j) > 0, \forall j \in E$ is not a sufficient one (in general) as the following example shows.

Example 8 Suppose $E = \{1, 2, 3, \ldots\}$, $d(j) = j^{-2}$ and Σ, the intensity matrix of J, is such that for $j = 1, 2, \ldots$, $\sigma_{j,j+1} = 1$; $\sigma_{jk} = 0$ otherwise. Then, for any initial distribution,

$$E(D(\infty)) \leq \sum_{j=1}^{\infty} j^{-2} < \infty,$$

thus $D(\infty) < \infty$ a.s. □

We give in Theorem 14 some properties of the unsatisfied demand and of the infimum of the net input, which is naturally related to the unsatisfied demand by Theorem 13.

Theorem 14 *The unsatisfied demand $I = \{I(t), t \geq 0\}$ is a continuous additive functional of (Z, J), moreover, if $Z(0) = 0$ a.s., then I is a continuous additive functional of (X, J).*

The negative of the infimum of the net input process $-m$ is a continuous additive functional of (X, J).

Proof: Since (Z, J) is a Markov process, the fact that the unsatisfied demand I is a continuous additive functional of (Z, J) follows from the definition of the unsatisfied demand, and Remark 2. Suppose now $Z(0) = 0$ a.s., then (since $X(0) = 0$ a.s.) $\sigma\{(X, J)(s), s \leq t\}$ is equal to $\sigma\{(Z, J)(s), s \leq t\}$ and I becomes a continuous additive functional of (X, J). The rest of the theorem follows trivially since, by Theorem 13, $I(t) = -m(t)$ a.s. in case $Z(0) = 0$ a.s. □

3.3 The inverse of the demand

It is of interest to study the right-continuous inverse of the demand process $D(t)$, i.e. for $x \geq 0$ we want to know the amount of time needed for the demand to exceed x viewed as a stochastic process. Denoting this process as B we thus have

$$B(x) = \inf\{t : D(t) > x\}. \tag{3.13}$$

We study below some of properties of the right-continuous inverse of an additive functional of a Markov process. The following result for inverses of additive functionals is partially mentioned by Blumenthal [3] but seems to call for a detailed proof.

Lemma 4 *If $\{A(t),\, t \geq 0\}$ is an additive functional of the Markov process R and $\tilde{A}(x)$ is its right-continuous inverse, $\tilde{A}(x) = \inf\{t : A(t) > x\}$, then:*

(a). $\tilde{A}$ has non-decreasing and a.s. right-continuous sample paths.

(b). If A has continuous and strictly increasing sample paths a.s. then (a.s.):

(i) $\tilde{A}(x) = \infty$ for $x \in [A(\infty), \infty)$.

(ii) $\tilde{A}(0) = 0$ and $\tilde{A}$ has continuous and strictly increasing sample paths on $[0, A(\infty))$.

(c). For each $x, \tilde{A}(x)$ is a stopping time.

(d). For any $t \in [0, \infty)$, a.s., $A(t) = \inf\{x : \tilde{A}(x) > t\}$.

Proof: (a). If $x_1 \leq x_2$, then

$$\tilde{A}^*(x_2) = \{t : A(t) > x_2\} \subseteq \{t : A(t) > x_1\} = \tilde{A}^*(x_1)$$

so the indexed sets $\tilde{A}^*(x)$ are non-increasing on x, and

$$\tilde{A}(x_1) = \inf \tilde{A}^*(x_1) \leq \inf \tilde{A}^*(x_2) = \tilde{A}(x_2).$$

This shows that $\tilde{A}(x)$ is non-decreasing. In the following we refer only to non-decreasing sample paths of A. Let $\{x_n\}$ be a decreasing sequence with limit $x \geq 0$; since $\{\tilde{A}(x_n)\}$ is a non-increasing sequence bounded below by $\tilde{A}(x)$ the sequence converges in $[0, +\infty]$ and

$$\lim \tilde{A}(x_n) \geq \tilde{A}(x). \tag{3.14}$$

Now note that for $y \geq 0$, $\tilde{A}^*(y)$ is either $(\tilde{A}(y), +\infty)$ or $[\tilde{A}(y), +\infty)$, so

$$(\tilde{A}(y), +\infty) \subseteq \tilde{A}^*(y) \subseteq [\tilde{A}(y), +\infty), \forall y \geq 0.$$

Consequently

$$(\tilde{A}(x), +\infty) \subseteq \tilde{A}^*(x) = \bigcup_n \tilde{A}^*(x_n) \subseteq \bigcup_n [\tilde{A}(x_n), +\infty) \subseteq [\lim \tilde{A}(x_n), +\infty)$$

which implies that $\lim \tilde{A}(x_n) \leq \tilde{A}(x)$. In view of (3.14) it follows that

$$\lim \tilde{A}(x_n) = \tilde{A}(x).$$

Therefore $\tilde{A}$ has right-continuous sample paths almost surely.

(b). We consider only continuous and strictly increasing sample paths of A such that $A(0) = 0$ (for which the associated sample paths of $\tilde{A}$ are non-decreasing and right-continuous) and keep the notation used in (i). If $A(\infty) < x < \infty$ then clearly $\tilde{A}(x) = \infty$ and by right-continuity also $\tilde{A}(A(\infty)) = \infty$; this proves (a). Trivially $\tilde{A}(0) = 0$ since $A(t) > 0, \forall t > 0$, thus to prove (b) we need to show that $\tilde{A}$ is left-continuous and strictly increasing on $\mathcal{S} = [0, A(\infty))$. We have, $\tilde{A}^*(y) = (\tilde{A}(y), +\infty), \forall y \in \mathcal{S}$. Let $\{x_n\}$ be a non-negative increasing sequence with limit $x \in \mathcal{S}$, then $\{\tilde{A}(x_n)\}$ is a non-decreasing sequence bounded above by $\tilde{A}(x) < \infty$, so the sequence converges on $[0, +\infty)$ and

$$\lim \tilde{A}(x_n) \leq \tilde{A}(x). \tag{3.15}$$

Consequently

$$\begin{aligned} [\tilde{A}(x), +\infty) &= \{t : A(t) \geq x\} = \bigcap_n \{t : A(t) > x_n\} \\ &= \bigcap_n (\tilde{A}(x_n), +\infty) \supseteq (\lim \tilde{A}(x_n), +\infty) \end{aligned}$$

which implies that $\lim \tilde{A}(x_n) \geq \tilde{A}(x)$. Again, in view of (3.15) it follows that

$$\lim \tilde{A}(x_n) = \tilde{A}(x).$$

This proves that $\tilde{A}(x)$ is left-continous and thus continuous on $\mathcal{S}$. Suppose, with the same conditions, that $\tilde{A}(x)$ is not strictly increasing on $\mathcal{S}$, that is

$$\exists y < x < A(\infty) \text{ such that } \tilde{A}(y) = \tilde{A}(x).$$

Then $(\tilde{A}(y), +\infty) = (\tilde{A}(x), +\infty)$ or equivalently

$$\{t : A(t) > y\} = \{t : A(t) > x\}.$$

Since $A(0) = 0$, this contradicts the fact that $A(t)$ is continuous and strictly increasing. Thus $\tilde{A}(x)$ is strictly increasing on $\mathcal{S}$. We have thus proved that $\tilde{A}$ has continuous and strictly increasing sample paths on $[0, A(\infty))$ a.s..

(c). Since for any $t, x \geq 0$

$$\{\tilde{A}(x) \leq t\} = \bigcap_n \{ \sup_{0 \leq \tau \leq t + \frac{1}{n}} A(\tau) > x\} \in \mathcal{F}_{t_+} = \mathcal{F}_t$$

$\tilde{A}(x)$ is a stopping time.

(d). We prove the result for non-decreasing and right-continuous sample paths of A. For $t \geq 0$, let

$$\overline{A}(t) = \inf\{x : \tilde{A}(x) > t\}$$

then, using the fact that A and $\tilde{A}$ are non-decreasing and right-continuous and the definitions of $\tilde{A}$ and $\overline{A}$,

$$x < (>)\, \overline{A}(t) \Rightarrow \tilde{A}(x) \leq (>)\, t \Rightarrow x \leq (\geq)\, A(t).$$

This shows that $\overline{A}(t) = A(t)$, $\forall t \geq 0$, which proves the statement. □

Theorem 15 *For all initial distributions, the inverse of the demand has the following properties:*

(a). $B(0) = 0$ and B has continuous and strictly increasing sample paths on $[0, D(\infty))$ and $B(x) = \infty$ for $x \in [D(\infty), \infty)$.

(b). For each x, $B(x)$ is a stopping time.

(c). For any $t \in [0, \infty)$,

$$D(t) = \inf\{x : B(x) > t\}. \tag{3.16}$$

(d). $\frac{1}{d_+} \leq \frac{B(x)}{x} \leq \frac{1}{d_-}$, for all $x > 0$.

(e). If J is ergodic and $d_+ < \infty$, then $\lim_{x\to\infty} \frac{B(x)}{x} = \left[\sum_{j\in E} d(j)\,\pi_j\right]^{-1}$, a.s.

(f). $\{(B, J\circ B)(x),\, 0 \leq x < D(\infty)\}$ is an MAP with infinitesimal generator:

$$\mathcal{A}_{(B,J\circ B)}\, f(x,j) = \frac{1}{d(j)}\frac{\partial}{\partial x} f(x,j) + \sum_{k\neq j} \frac{\sigma_{jk}}{d(j)}\left[f(x,k) - f(x,j)\right]. \tag{3.17}$$

Proof: The statements $(a) - (c)$ follow in the same way as Lemma 4 (b)-(d) by Lemma 3 (a) and (b).

(d). It suffices to prove the result for $0 < d_- \leq d_+ < +\infty$; if so, given $x > 0$ and arbitrary t_1, t_2 such that $t_2 > x/d_-$ and $t_1 < x/d_+$,

$$D(t_2) \geq t_2\, d_- > x \text{ and } D(t_1) \leq t_1\, d_+ < x.$$

Now, by the definition of $B(x)$, we have $B(x) \leq t_2$ and $B(x) \geq t_1$. This gives the desired result since t_1, t_2 are arbitrary.

(e). From Lemma 3(e),

$$\lim_{t\to\infty} \frac{D(t)}{t} = \sum_{j\in E} d(j)\,\pi_j, \text{ a.s.}$$

and since, by Lemma 3(a) and (b), D has continuous and strictly increasing sample paths, then for all $x > 0$,

$$B(x) = t \Leftrightarrow D(t) = x.$$

Since by (d) $\lim_{x\to\infty} B(x) = \infty$, this gives

$$\lim_{x\to\infty} \frac{B(x)}{x} = \lim_{t\to\infty} \frac{t}{D(t)} = \frac{1}{\sum_{j\in E} d(j)\,\pi_j}.$$

(f). The statement follows as a consequence of the results obtained in Example 11 below. □

Remark 3 If we let J^B be the MJP with intensity matrix $\Sigma^B = D^{-1}\Lambda P$, where $D = (\delta_{jk} d(j))$, then (up to the explosition time $D(\infty)$ and by (3.17)) $(B, J \circ B)$ is equivalent to the MCPP (B, J^B) with drift $[d(j)]^{-1}$, $j \in E$, and no jumps. Note that if J is ergodic and $d_+ < \infty$ then J^B is also ergodic and has stationary distribution π^B given by:

$$\pi_j^B = \frac{\pi_j\, d(j)}{\sum_{k\in E} \pi_k\, d(k)};$$

this would lead to Theorem 15 (e) using a procedure similar to the one used to obtain Lemma 3 (e). Since the infinitesimal generator of a process characterizes the process, using Theorem 1.17, we obtain

$$B(x) = \int_0^x \frac{1}{d(J^B(y))}\, dy\ , \text{ for } x < D(\infty).$$

As was done for the demand, we can identify the inverse of the demand as a random clock with $[d(J^B(y))]^{-1}$ being its time speed at y. □

Chapter 4

A fluid storage model

In this chapter we study the fluid storage model presented in subsection 1.4.1, which occurs in communications systems. The storage level $Z(t)$ satisfies the following equation

$$Z(t) = Z(0)+Y(t)+I(t) = Z(0)+\int_0^t y(J(s))\,ds+\int_0^t y(J(s))^- 1_{\{Z(s)\leq 0\}}\,ds \quad (4.1)$$

where Y and I are the net input and unsatisfied demand processes, respectively, and J is an MJP. The sample functions of the net input are continuous a.s., and differentiable everywhere except at the transition epochs T_n $(n \geq 0)$ of the Markov chain J. As argued in section 1.4, in order to study the process (Z, I, J) it may be of interest to first study the properties of the embedded process $(Z(T_n), I(T_n), J(T_n))$.

From Theorem 13, it follows that equation (4.1) has an unique solution given by

$$I(t) = [Z(0) + m(t)]^- = \left[Z(0) + \inf_{0\leq\tau\leq t} Y(\tau)\right]^- \quad (4.2)$$

which implies in particular that

$$Z(t) = \max\{Z(0) + Y(t), Y(t) - m(t)\} \geq [Z(0) + Y(t)]^{+} \geq 0 \tag{4.3}$$

with $y^{+} = \max\{y, 0\}$. We note that from (4.2) it follows that the storage level and the unsatisfied demand have the properties that Harrison [16] (Section 2.2) postulates for the same processes. However our model description states more effectively the mechanics of the system since it states the storage policy in a more precise way.

A particular consequence of solution (4.2) is that $Z(T_n)$ and $I(T_n)$ may be identified as functionals on the process $(T_n, Y(T_n), J(T_n))$, which is an MRW. So the properties of this MRW are investigated in section 4.2, the key result being a Wiener-Hopf factorization due to Presman [35]; see Prabhu and Tang [32] and Prabhu, Tang and Zhu [33]. These properties are then used in section 4.3 to study the properties of the storage level and the unsatisfied demand. This constitutes a novel use of MRWs in the study of models envolving MAPs.

We shall denote by $Q = (q_{jk})$ the generator matrix of J and assume that J has a stationary distribution $(\pi_j, j \in E)$. For analytical convenience we assume that $y(j) \neq 0$ for $j \in E$. Finally, we let

$$E_{\mathbf{1}} = \{j \in E : y(j) \leq 0\}, \qquad E_{\mathbf{0}} = \{j \in E : y(j) > 0\} \tag{4.4}$$

and

$$Z_n = Z(T_n), \qquad I_n = I(T_n), \qquad S_n = Y(T_n), \qquad Y_{n+1} = S_{n+1} - S_n \tag{4.5}$$

so that $Y_{n+1} = y(J_n)(T_{n+1} - T_n)$.

4.1 Preliminary results

We prove some preliminary results for the embedded process $(Z(T_n), I(T_n), J(T_n))$.

Theorem 16 *For $T_n \le t \le T_{n+1}$ we have*

$$Z(t) = [Z_n + y(J_n)(t - T_n)]^+, \qquad I(t) = I_n + [Z_n + y(J_n)(t - T_n)]^- \quad (4.6)$$

where

$$Z_n = Z_0 + S_n + I_n, \qquad I_n = [Z_0 + m_n]^- = \left[Z_0 + \min_{0 \le r \le n} S_r\right]^- . \quad (4.7)$$

Proof: If $J_n \in E_0$, then $y(J_n)^- = 0$, and (1.23) and the nonnegativity of Z give

$$Z(t) = Z_n + y(J_n)(t - T_n) \ge 0.$$

This expression for $Z(t)$ holds if $J_n \in E_1$ and $Z(t) > 0$ for $T_n \le t \le T_{n+1}$. Suppose now that $J_n \in E_1$ and $Z(t) = 0$ for some $t \in [T_n, T_{n+1}]$, and let t_0 be the smallest of such t. Then from (1.23), since $y(J_n)^- = -y(J_n)$, we obtain

$$Z(t) = Z_n + y(J_n) \int_{T_n}^{t} 1_{\{Z(s)>0\}}\, ds \qquad (T_n \le t \le T_{n+1})$$

which shows that $Z(t)$ is nonincreasing in $[T_n, T_{n+1}]$. This implies that $Z(t) = 0$ for $t \in [t_0, T_{n+1}]$ by the nonnegativity of Z. Moreover from (1.23)

$$Z(t) = Z_n + y(J_n)(t - T_n) > 0 \qquad (T_n \le t < t_0)$$

and for $t_0 \le t \le T_{n+1}$,

$$Z(t) = 0 = Z(t_0) = Z_n + y(J_n)(t_0 - T_n) \ge Z_n + y(J_n)(t - T_n).$$

All cases considered above are summarized in the result

$$Z(t) = \max\{0, Z_n + y(J_n)(t - T_n)\} \qquad (n \geq 0,\, t \in [T_n, T_{n+1}])$$

which proves (4.6). This implies in particular that

$$Z_{n+1} = \max\{0, Z_n + Y_{n+1}\} \qquad (n \geq 0).$$

The desired result follows from the following lemma familiar in queueing systems, since

$$S_n = Y_1 + Y_2 + \ldots + Y_n. \;\Box$$

Lemma 5 *Let* $\{x_n,\, n \geq 1\}$ *be a sequence of real numbers and suppose that a sequence of real numbers* $\{z_n,\, n \geq 0\}$*, with* $z_0 \geq 0$*, is defined by the recurrence relations*

$$z_{n+1} = \max\{0, z_n + x_{n+1}\} \qquad (n \geq 0).$$

Then

$$z_n = \max\{z_0 + s_n, s_n - \min_{0 \leq r \leq n} s_r\}$$

where

$$s_0 = 0, \qquad s_n = x_1 + x_2 + \ldots + x_n \quad (n \geq 1). \qquad \Box$$

Theorem 16 shows that in order to study the process (Z_n, I_n, J_n) it suffices to investigate the MRW (T_n, S_n, J_n), which we do in the next section. We note that if $T_n \leq t \leq T_{n+1}$, then $Y(t) = S_n + y(J_n)\,(t - T_n)$, thus $\min(S_n, S_{n+1}) \leq Y(t) \leq \max(S_n, S_{n+1})$ which in turns implies that a.s.

$$\liminf Y(t) = \liminf S_n, \qquad \limsup Y(t) = \limsup S_n. \tag{4.8}$$

If we denote for $t \geq 0$ and $n = 0, 1, \ldots,$

$$M(t) = \sup_{0 \leq \tau \leq t} Y(\tau), \qquad M_n = \max_{0 \leq r \leq n} S_r \tag{4.9}$$

we may conclude that as in (4.8),

$$\lim_{t \to \infty} M(t) = \lim_{n \to \infty} M_n = M \leq +\infty \tag{4.10}$$

$$\lim_{t \to \infty} m(t) = \lim_{n \to \infty} m_n = m \geq -\infty. \tag{4.11}$$

These statements show that some conclusions about the fluctuation behaviour of the net input process may be drawn from the associated MRW (S_n, J_n). This in turn has implications for the storage level and unsatisfied demand since these processes depend on the net input. We denote by $(\pi_j^\star, j \in E)$ the stationary distribution of (J_n), so that

$$\pi_j^\star = \frac{(-q_{jj})\,\pi_j}{\sum_{k \in E}(-q_{kk})\,\pi_k}. \tag{4.12}$$

Also, let $\overline{y}$ be the *net input rate*

$$\overline{y} = \sum_{j \in E} \pi_j\, y(j) \tag{4.13}$$

where we assume the sum exists, but may be infinite. We have then the following.

Theorem 17 *(Fluctuation behaviour of $Y(t)$) We have a.s.*

(a). $\frac{Y(t)}{t} \to \overline{y}$.

(b). If $\overline{y} > 0$, then $\lim Y(t) = +\infty$, $m > -\infty$, and $M = +\infty$.

(c). If $\overline{y} = 0$, then $\liminf Y(t) = -\infty$, $\limsup Y(t) = +\infty$, $m = -\infty$, and $M = +\infty$.

(d). If $\overline{y} < 0$, then $\lim Y(t) = -\infty$, $m = -\infty$, and $M < +\infty$.

Proof: (a). The proof of the statement is standard, but is given here for completeness. We have

$$\frac{Y(t)}{t}=\frac{1}{t}\int_0^t y(J(s))\,ds=\frac{1}{t}\int_0^t\{[y(J(s))]^+-[y(J(s))]^-\}\,ds. \tag{4.14}$$

Now since J is ergodic

$$\begin{aligned}\frac{1}{t}\int_0^t [y(J(s))]^+\,ds &= \frac{1}{t}\int_0^t \sum_{j\in E}[y(j)]^+ 1_{\{J(s)=j\}}\,ds\\ &=\sum_{j\in E}\frac{[y(j)]^+}{t}\int_0^t 1_{\{J(s)=j\}}\,ds\to\sum_{j\in E}[y(j)]^+\,\pi_j\ ,\ \text{as } t\to\infty,\ \text{a.s.}\end{aligned}$$

and similarly

$$\frac{1}{t}\int_0^t [y(J(s))]^-\,ds\to\sum_{j\in E}[y(j)]^-\,\pi_j\ ,\ \text{as } t\to\infty,\ \text{a.s.}$$

These and (4.14) imply that a.s.

$$\lim_{t\to\infty}\frac{Y(t)}{t}=\sum_{j\in E}\{[y(j)]^+-[y(j)]^-\}\,\pi_j=\sum_{j\in E}y(j)\,\pi_j=\overline{y}.$$

Now let $\mu^\star$ be the mean increment in the MRW (S_n,J_n)

$$\mu^\star=\sum_{j\in E}\pi_j^\star\,E\,[Y_1\mid J_0=j]\,. \tag{4.15}$$

Since $E\,[Y_1\mid J_0=j]=y(j)/(-q_{jj})$, $\overline{y}=\mu^\star\sum_{k\in E}(-q_{kk})\,\pi_k$. The statements (b) – (d) follow from this, (4.8), and (4.10)-(4.11), by using Proposition 2 of Prabhu and Tang [32] and Theorem 8 of Prabhu, Tang and Zhu [33]. These last two results describe the fluctuation behaviour of the MRW (S_n,J_n). □

4.2 The MRW (T_n,S_n,J_n)

In this section we investigate the properties of the MRW (T_n,S_n,J_n). We note that the conditional distribution of the increments $(T_n-T_{n-1},S_n-S_{n-1})$, given

J_{n-1} is singular, since $Y_n = S_n - S_{n-1} = y(J_n)(T_n - T_{n-1})$ a.s. The distribution of (T_1, Y_1, J_1) is best described by the transform matrix

$$\Phi(\theta,\omega) = \left(\phi_{jk}(\theta,\omega)\right) = \left(E\left[e^{-\theta T_1 + i\omega Y_1}; J_1 = k \mid J_0 = j\right]\right) \tag{4.16}$$

for $\theta > 0$, ω real and $i = \sqrt{-1}$. We find that

$$\Phi(\theta,\omega) = \left(\phi_{jk}(\theta,\omega)\right) = \left(\alpha_j(\theta,\omega)\, p_{jk}\right) = \left(\alpha_j(\theta,\omega)\, \delta_{jk}\right)\left(p_{jk}\right) = \alpha(\theta,\omega)P \tag{4.17}$$

where

$$\alpha_j(\theta,\omega) = \frac{-q_{jj}}{-q_{jj} + \theta - i\omega y(j)}, \qquad p_{jk} = \frac{q_{jk}}{(-q_{jj})} \;\; (k \neq j), \qquad p_{jj} = 0. \tag{4.18}$$

For the time-reversed MRW $(\hat{T}_n, \hat{S}_n, \hat{J}_n)$ corresponding to the given MRW we have

$$\begin{aligned} \hat{\Phi}(\theta,\omega) &= \left(\hat{\phi}_{jk}(\theta,\omega)\right) = \left(E\left[e^{-\theta \hat{T}_1 + i\omega \hat{Y}_1}; \hat{J}_1 = k \mid \hat{J}_0 = j\right]\right) \\ &= \left(\frac{\pi_k^\star}{\pi_j^\star} E\left[e^{-\theta T_1 + i\omega Y_1}; J_1 = j \mid J_0 = k\right]\right) = \hat{P}\,\alpha(\theta,\omega) \end{aligned} \tag{4.19}$$

where $\hat{P}$ is the transition probability matrix of the time-reversed chain $\hat{J}$, namely

$$\hat{P} = \left(\hat{p}_{jk}\right) = \left(\frac{\pi_k^\star}{\pi_j^\star}\, p_{kj}\right). \tag{4.20}$$

Since the T_n are non-decreasing a.s., the flutuacting theory of (T_n, S_n, J_n) is adequately described by (S_n). We now define the descending ladder epoch $\overline{N}$ of this MRW (T_n, S_n, J_n), and the ascending ladder epoch N of the time-reversed MRW $(\hat{T}_n, \hat{S}_n, \hat{J}_n)$

$$\overline{N} = \min\{n : S_n < 0\}, \qquad N = \min\{n : \hat{S}_n > 0\}. \tag{4.21}$$

(Here we adopt the convention that the minimum of an empty set is $+\infty$.) It should be noted that both N and $\overline{N}$ are strong ladder epochs, which is reasonable since the increments of S_n and $\hat{S}_n$ in each case have an absolutely continuous distribution. The random variables $S_{\overline{N}}$ and $\hat{S}_N$ are the ladder heights corresponding to $\overline{N}$ and N. We also denote the transforms (in the matrix form)

$$\overline{\chi} = (\overline{\chi}_{jk}(z,\theta,\omega)) = \left(E\left[z^{\overline{N}}\, e^{-\theta T_{\overline{N}}+i\omega S_{\overline{N}}}; J_{\overline{N}} = k \mid J_0 = j\right]\right) \tag{4.22}$$

$$\chi = (\chi_{jk}(z,\theta,\omega)) = \left(\frac{\pi_k^\star}{\pi_j^\star}\, E\left[z^{N}\, e^{-\theta \hat{T}_{N}+i\omega \hat{S}_{N}}; \hat{J}_{N} = j \mid \hat{J}_0 = k\right]\right) \tag{4.23}$$

where $0 < z < 1$, $\theta > 0$, $i = \sqrt{-1}$ and ω is real. Connecting these two transforms is the Wiener-Hopf factorization, first established by Presman [35] analytically, and interpreted in terms of the ladder variables defined above by Prabhu, Tang and Zhu [33]. The result is the following

Lemma 6 *(Wiener-Hopf factorization) For the MRW (T_n, S_n, J_n) with $0 < z < 1$, $\theta > 0$ and ω real*

$$I - z\Phi(\theta,\omega) = [I - \chi(z,\theta,\omega)]\,[I - \overline{\chi}(z,\theta,\omega)]\,. \quad \square \tag{4.24}$$

We shall use this factorization and the special structure of our MRW to indicate how the transforms χ and $\overline{\chi}$ can be computed in the general case. It turns out that our results contain information concerning the descending ladder epoch $\overline{T}$ of the net input process (Y, J) and the ascending ladder epoch T of the time-reversed process $(\hat{Y}, \hat{J})$, which is defined as follows:

$$\hat{J}(t) = \hat{J}_n \quad (\hat{T}_{n-1} < t \leq \hat{T}_n), \qquad\qquad \hat{Y}(t) = \int_0^t y(\hat{J}(s))\, ds. \tag{4.25}$$

Thus

$$\overline{T} = \inf\{t > 0 \,:\, Y(t) \leq 0\}, \qquad\qquad T = \inf\{t > 0 \,:\, \hat{Y}(t) \geq 0\}. \tag{4.26}$$

We note that $Y(\overline{T}) = 0$ and $\hat{Y}(T) = 0$ a.s. For $0 < z < 1$, $\theta > 0$ we define the transforms

$$\zeta = (\zeta_{jk}(z,\theta)) = \left(E\left[z^{\overline{N}} e^{-\theta \overline{T}}; J(\overline{T}) = k \mid J(0) = j\right]\right) \tag{4.27}$$

$$\eta = (\eta_{jk}(z,\theta)) = \left(\frac{\pi_k^\star}{\pi_j^\star} E\left[z^N e^{-\theta T}; \hat{J}(T) = j \mid \hat{J}(0) = k\right]\right). \tag{4.28}$$

Theorem 18 *For $0 < z < 1$, $\theta > 0$ and ω real we have*

$$\overline{\chi}(z,\theta,\omega) = \zeta(z,\theta)\,\Phi(\theta,\omega), \qquad \chi(z,\omega) = \alpha(\theta,\omega)\,\eta(z,\theta). \tag{4.29}$$

Proof: An inspection of the sample paths of (Y,J) will show that $J(\overline{T}) = J_{\overline{N}-1}$ and

$$T_{\overline{N}} - \overline{T} = \frac{S_{\overline{N}}}{y(J(\overline{T}))} \quad \text{a.s.}$$

Since $\overline{T}$ is a stopping time for (Y,J), we see that given $J(\overline{T}) = l$, $S_{\overline{N}}/y(J(\overline{T}))$ is independent of $\overline{T}$ and has the same distribution as T_1, given $J_0 = l$. Therefore $\overline{\chi}_{jk}(z,\theta,\omega)$ is equal to

$$\sum_{l\in E} E\left[z^{\overline{N}} \exp\left(-\theta\left[\overline{T} + \frac{S_{\overline{N}}}{x\left(J_{\overline{N}-1}\right)}\right] + i\omega S_{\overline{N}}\right); J_{\overline{N}-1} = l, J_{\overline{N}} = k \mid J_0 = j\right]$$

$$\begin{aligned}
&= \sum_{l\in E} E\left[z^{\overline{N}} e^{-\theta\overline{T}}; J(\overline{T}) = l \mid J_0 = j\right] \\
&\qquad .E\left[\exp\left(-\theta\frac{S_{\overline{N}}}{x\left(J_{\overline{N}-1}\right)} + i\omega S_{\overline{N}}\right); J_{\overline{N}} = k \mid J_{\overline{N}-1} = l\right] \\
&= \sum_{l\in E} \zeta_{jl}(z,\theta) E\left[e^{-\theta T_1 + i\omega Y_1}; J_1 = k \mid J_0 = l\right] = \sum_{l\in E} \zeta_{jl}(z,\theta)\,\phi_{lk}(\theta,\omega).
\end{aligned}$$

Thus $\overline{\chi}(z,\theta,\omega) = \zeta(z,\theta)\,\Phi(\theta,\omega)$. The proof of the result $\chi(z,\theta,\omega) = \alpha(\theta,\omega)\,\eta(z,\theta)$ is similar. □

In general, for an $(| E | \times | E |)$-matrix A we block-partition A in the form

$$A = \begin{pmatrix} A_{\mathbf{00}} & A_{\mathbf{01}} \\ A_{\mathbf{10}} & A_{\mathbf{11}} \end{pmatrix}$$

with the rows and columns of $A_{\mathbf{00}}$ corresponding to the states in $E_{\mathbf{0}}$. We have now

$$\zeta = \begin{pmatrix} \mathbf{0} & \zeta_{\mathbf{01}} \\ \mathbf{0} & \zeta_{\mathbf{11}} \end{pmatrix}, \quad \eta = \begin{pmatrix} \eta_{\mathbf{00}} & \eta_{\mathbf{01}} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}, \quad I = \begin{pmatrix} I_{\mathbf{00}} & \mathbf{0} \\ \mathbf{0} & I_{\mathbf{11}} \end{pmatrix} \tag{4.30}$$

where I is the identity matrix. From (4.24) and Theorem 18 we have the following.

Theorem 19 *We have for $0 < z < 1$, $\theta > 0$ and ω real*

$$\chi = \begin{pmatrix} \alpha_{\mathbf{00}}\eta_{\mathbf{00}} & \alpha_{\mathbf{00}}\eta_{\mathbf{01}} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}, \qquad \overline{\chi} = \begin{pmatrix} \zeta_{\mathbf{01}}\Phi_{\mathbf{10}} & \zeta_{\mathbf{01}}\Phi_{\mathbf{11}} \\ z\Phi_{\mathbf{10}} & z\Phi_{\mathbf{11}} \end{pmatrix}, \tag{4.31}$$

and

$$I_{\mathbf{00}} - \overline{\chi}_{\mathbf{00}} = (I_{\mathbf{00}} - \chi_{\mathbf{00}})^{-1}\left[(I_{\mathbf{00}} - z\Phi_{\mathbf{00}}) - z\chi_{\mathbf{01}}\Phi_{\mathbf{10}})\right] \tag{4.32}$$

$$\overline{\chi}_{\mathbf{01}} = (I_{\mathbf{00}} - \chi_{\mathbf{00}})^{-1}\left[z\Phi_{\mathbf{01}} - \chi_{\mathbf{01}}(I_{\mathbf{11}} - z\Phi_{\mathbf{11}})\right] \tag{4.33}$$

where the inverse exists in the specified domain. □

We note that if we let

$$\overline{\gamma} = (\overline{\gamma}_{jk}(z,\theta)) = \left(E\left[z^{\overline{N}} e^{-\theta T_{\overline{N}}}; J_{\overline{N}-1} = k \mid J_0 = j\right]\right) \tag{4.34}$$

$$\gamma = (\gamma_{jk}(z,\theta)) = \left(\frac{\pi_k^\star}{\pi_j^\star} E\left[z^N e^{-\theta \hat{T}_N}; \hat{J}_N = j \mid \hat{J}_0 = k\right]\right) \tag{4.35}$$

then

$$\overline{\gamma} = \zeta\, I^\theta, \qquad \gamma = I^\theta\, \eta \tag{4.36}$$

with

$$I^{\theta} = \left(\frac{-q_{jj}}{-q_{jj} + \theta}\, \delta_{jk} \right), \qquad R = \left(\frac{\mid y(j) \mid}{-q_{jj} + \theta}\, \delta_{jk} \right). \tag{4.37}$$

Due to (4.36) the results for (ζ, η) are equivalent to those for $(\overline{\gamma}, \gamma)$. As a matter of convenience we express some of the remaining results of this section in terms of $(\overline{\gamma}, \gamma)$.

Corollary 7 *For* $0 < z < 1$ *and* $\theta > 0$ *we have*

$$\gamma_{00} + (I_{00} - \gamma_{00})\, \overline{\gamma}_{01}\, P_{10} = z \left[I^{\theta}_{00}\, P_{00} + \gamma_{01}\, I^{\theta}_{11}\, P_{10} \right] \tag{4.38}$$

$$\gamma_{01} + (I_{00} - \gamma_{00})\, \overline{\gamma}_{01}\, P_{11} = z \left[I^{\theta}_{00}\, P_{01} + \gamma_{01}\, I^{\theta}_{11}\, P_{11} \right] \tag{4.39}$$

$$R_{00}\, \overline{\gamma}_{01}\, P_{10} + (I_{00} - \gamma_{00})\, \overline{\gamma}_{01}\, R_{11}\, P_{10} = z \gamma_{01}\, R_{11}\, I^{\theta}_{11}\, P_{10} \tag{4.40}$$

$$R_{00}\, \overline{\gamma}_{01}\, P_{11} + (I_{00} - \gamma_{00})\, \overline{\gamma}_{01}\, R_{11}\, P_{11} = z \gamma_{01}\, R_{11}\, I^{\theta}_{11}\, P_{11}. \tag{4.41}$$

Proof: We equate the real parts of the identity (4.32) and put $\omega = 0$. This yields (4.38). We also equate the imaginary parts of (4.32), divide by ω and let $\omega \to 0$. This yields (4.40), in view of (4.38).

The proof of (4.39) and (4.41) is similar, starting with the identity (4.33). □

Theorem 19 shows that the submatrices $\overline{\chi}_{00}$ and $\overline{\chi}_{01}$ are determined by χ_{00} and χ_{01}. Corollary 7 can be used in some important cases to reduce the computation to a single (matrix) equation for $\overline{\gamma}_{01}$, as we will show in the following. Case (a) arises in models with $\mid E_1 \mid > 1$, while case (b) covers the situation with $\mid E_1 \mid = 1$. Details of the computations are omitted.

Case (i). If the submatrix P_{11} has an inverse, then

$$\gamma_{00} = z I^{\theta}_{00}\, P_{00} + R_{00}\, \overline{\gamma}_{01}\, R^{-1}_{11}\, P_{10} \tag{4.42}$$

$$\gamma_{01} = zI^{\theta}_{00}\, P_{01} + R_{00}\,\overline{\gamma}_{01}\, R^{-1}_{11}\, P_{11} \tag{4.43}$$

where $\overline{\gamma}_{01}$ satisfies the equation

$$\begin{aligned}\overline{\gamma}_{01}\left[R^{-1}_{11} P_{10}\right]\overline{\gamma}_{01} \quad &- \quad \left[R^{-1}_{00}\left(I_{00} - zI^{\theta}_{00} P_{00}\right)\overline{\gamma}_{01} + \overline{\gamma}_{01} R^{-1}_{11}\left(I_{11} - zP_{11} I^{\theta}_{11}\right)\right] \\ &+ \quad z^2 R^{-1}_{00} I^{\theta}_{00} P_{01} I^{\theta}_{11} = 0.\end{aligned} \tag{4.44}$$

Case (ii). If $P_{11} = \mathbf{0}$ and $r_{jj}(\theta) = r^{\theta}_1 \ (j \in E_1)$, then

$$\gamma_{00} = zI^{\theta}_{00}\, P_{00} + \frac{1}{r^{\theta}_1} R_{00}\,\overline{\gamma}_{01}\, P_{10}, \qquad \gamma_{01} = zI^{\theta}_{00}\, P_{01} \tag{4.45}$$

$$\begin{aligned}\frac{1}{r^{\theta}_1}\left(\overline{\gamma}_{01}\, P_{10}\right)^2 \quad &- \quad \left[R^{-1}_{00}\left(I_{00} - zI^{\theta}_{00}\, P_{00}\right) + \frac{1}{r^{\theta}_1}\right]\left(\overline{\gamma}_{01}\, P_{10}\right) \\ &+ \quad z^2 R^{-1}_{00}\, I^{\theta}_{00}\, P_{01}\, I^{\theta}_{11}\, P_{10} = 0.\end{aligned} \tag{4.46}$$

Example 9 Consider the Gaver-Lehoczky [13] model with a single output channel, in which the channel is shared by data and voice calls (with calls having preemptive priority over data). Here $J(t) = 0$ if a call is in progress at time t (i.e. the channel is not available for data transmission), and $J(t) = 1$ otherwise. Thus J has a two state space $\{0, 1\}$ and

$$a(0) = a(1) = c_0, \qquad d(0) = 0,\ d(1) = c_2 \quad (c_0 < c_2),$$

so that $E_0 = \{0\}$, $E_1 = \{1\}$, $y(0) = c_0$ and $y(1) = c_0 - c_2 = -c_1$. Let the arrival and service rate of calls be denoted by λ and μ respectively, then

$$P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad Q = (q_{jk}) = \begin{pmatrix} -\mu & \mu \\ \lambda & -\lambda \end{pmatrix}.$$

We may now use (4.36) and (4.45)-(4.46) to conclude that

$$\eta_{01} = z, \qquad \eta_{00} = \frac{\sigma_1}{\sigma_0}\zeta_{01}, \qquad \sigma_1\,\zeta^2_{01} - \left[(\sigma_0 + \sigma_1) + \left(\frac{1}{c_0} + \frac{1}{c_1}\right)\theta\right]\zeta_{01} + z^2\,\sigma_0 = 0$$

where $\sigma_0 = \frac{\mu}{c_0}$ and $\sigma_1 = \frac{\lambda}{c_1}$. This implies that

$$\zeta_{01}(z,\theta) = \frac{(\sigma_0+\sigma_1) + \left(\frac{1}{c_0}+\frac{1}{c_1}\right)\theta - \sqrt{\left[(\sigma_0+\sigma_1)+\left(\frac{1}{c_0}+\frac{1}{c_1}\right)\theta\right]^2 - 4\,z^2\,\sigma_0\,\sigma_1}}{2\,\sigma_1}. \tag{4.47}$$

For an $M/M/1$ queue with arrival and service rates σ_1 and σ_0 respectively, we denote the busy period by $\overline{T}^\star$ and the number of customers served during the busy period by $\overline{N}^\star$. From (4.47) we have the following (see Section 2.8 of Prabhu [29])

$$\zeta_{01}(z,\theta) = E\left[z^{\overline{N}} e^{-\theta\overline{T}}; J(\overline{T}) = 1 \mid J(0) = 0\right] = E\left[z^{2\,\overline{N}^\star} e^{-\theta\left(\frac{1}{c_0}+\frac{1}{c_1}\right)\overline{T}^\star}\right].$$

We note that this case has been considered also by Chen and Yao [9], Gaver and Miller [14], and Kella and Whitt [18] in the context of storage models for which the net input is alternating nonincreasing and non-decreasing. □

4.3 The main results

With the properties for the MRW (T_n, S_n, J_n) established in section 4.2, we are now in a position to derive the main results of the paper. We first state the following results for the embedded process (Z_n, I_n, J_n), which follow easily from Theorem 3 and 4 of Prabhu and Tang [32].

Lemma 7 *If $Z_0 = 0$ a.s. then for $\theta > 0$ and ω_1, ω_2 real we have*

$$\left(\sum_{n=0}^{\infty} E\left[e^{-\theta T_n + i\omega_1 Z_n + i\omega_2 I_n}; J_n = k \mid J_0 = j\right]\right)^{-1} = \left[I - \chi(z,\theta,\omega_1)\right]\left[I - \overline{\chi}(z,\theta,-\omega_2)\right]. \;\square \tag{4.48}$$

Lemma 8 *Suppose that $\mu^\star < 0$. Then for all initial distributions*

$$(Z_n, J_n) \xrightarrow{\mathcal{D}} (Z^\star_\infty, J^\star_\infty) \text{ as } n \to \infty,$$

where $J^\star_\infty$ *is the stationary version of* (J_n) *and* $\left(E\left[e^{i\omega Z^\star_\infty}; J^\star_\infty = k\right]\right)$ *is the* k*th element of the row vector*

$$\pi^\star \left[I - \chi(1,0,0)\right] \left[I - \chi(1,0,\omega)\right]$$

with $\pi^\star = (\pi^\star_j,\ j \in E)$. □

For finite t, the distribution of $(Z(t), I(t), J(t))$ can be found from equation (4.6) using Lemma 7. We note that since $\overline{T} = \inf\{t > 0 \,:\, Y(t) \leq 0\}$ we have

$$\overline{T} = \inf\{t > 0 \,:\, Z(t) = 0 \mid Z(0) = 0\}. \tag{4.49}$$

Thus $\overline{T}$ is the busy period of the storage. The transform of $(\overline{T}, J(\overline{T}))$ is $\zeta(1,\theta)$ given by

$$\zeta(1,\theta) = (\zeta_{jk}(1,\theta)) = \left(E\left[e^{-\theta\overline{T}}; J(\overline{T}) = k \mid J(0) = j\right]\right) \tag{4.50}$$

In section 4.2 it was shown how this transform can be computed. We recall that $\overline{y}$ is the *net input rate* and is given by (4.13).

Theorem 20 *The busy period* $\overline{T}$ *defined by (4.49) has the following properties:*

(*a*). *Given* $J(0) \in E_1$, $\overline{T} = 0$ *a.s.*

(*b*). *If* $\overline{y} < 0$ *then* $\overline{T} < \infty$ *a.s.*

Proof: (*a*). The statement follows immediately from the fact that $y(j) < 0$ for $j \in E_1$.

(*b*). Since $\overline{y} < 0$ we have $\mu^\star < 0$ (see the proof of Theorem 17 (*b*)). This implies that the descending ladder epoch of the MRW (T_n, S_n, J_n) is finite ($\overline{N} < \infty$ a.s.) in virtue of Proposition 2 of Prabhu and Tang [32]. This in turn implies that $T_{\overline{N}} < \infty$ a.s. The statement now follows since $\overline{T} \leq T_{\overline{N}}$. □

The limit behaviour of the process $(Z(t), I(t), J(t))$ as $t \to \infty$ can also be obtained from that of the embedded process (Z_n, I_n, J_n) as $n \to \infty$, by using Lemma 8. The following theorems characterize this limit behaviour.

Theorem 21 *The process $(Z(t), I(t), J(t))$ has the following properties:*

(a). If $\overline{y} > 0$, then $I(t) \to (Z(0)+m)^- < +\infty$ and $\frac{Z(t)}{t} \to \overline{y}$ a.s.;

in particular $Z(t) \to +\infty$ a.s.

(b). If $\overline{y} = 0$, then $I(t) \to +\infty$ and $\limsup Z(t) = +\infty$ a.s.

(c). If $\overline{y} < 0$, then $\frac{I(t)}{t} \to -\overline{y}$ and $\frac{Z(t)}{t} \to 0$ a.s.; in particular $I(t) \to +\infty$ a.s.

Proof: *(a)*. We first note that $I(t)$ converges as indicated by Theorem 17 *(b)* and (4.2). The rest of the statement follows directly from Theorem 17 *(a)* and (4.1).

(b). From Theorem 17 *(iii)* and (4.2) $I(t) \to (Z(0)+m)^- = +\infty$. Also, from (4.1), since $I(t)$ is nonnegative, $Z(t) \geq Z(0) + Y(t)$. Using Theorem 17 *(iii)* we obtain

$$\limsup Z(t) \geq \limsup [Z(0) + Y(t)] = +\infty.$$

(c). Since $Y(t)$ has continuous sample functions and $\frac{Y(t)}{t} \to \overline{y} < 0$, by Theorem 17 *(a)*, standard analytical arguments show that

$$\lim \frac{m(t)}{t} = \lim \frac{Y(t)}{t} = \overline{y} < 0.$$

The desired results now follow from (4.2)-(4.3). □

Theorem 22 *If $\overline{y} < 0$ then for $z_0, z \geq 0$ and $j, k \in E$, and with $(Z_\infty^\star, J_\infty^\star)$ being the limit distribution of (Z_n, J_n) as given in Lemma 8,*

$$\lim_{t\to\infty} P\{Z(t) \leq z; J(t) = k \mid Z(0) = z_0, J(0) = j\}$$
$$= \pi_k \int_0^\infty P\{Z_\infty^\star \in dv \mid J_\infty^\star = k\}\; P\{Y_1 \leq z - v \mid J(0) = k\}. \quad (4.51)$$

Proof: Let $N(t) = \sup\{n : T_n \leq t\}$. We have

$$P\{Z(t) \leq z; J(t) = k \mid Z(0) = z_0, J(0) = j\}$$
$$= \int_0^\infty P\left\{Z_{N(t)} \in dv; J_{N(t)} = k \mid Z_0 = z_0, J_0 = j\right\}$$
$$.P\left\{Z(t) \leq z \mid Z(T_{N(t)}) = v, J(T_{N(t)}) = k, Z(0) = z_0, J(0) = j\right\}$$
$$= P\left\{J_{N(t)} = k \mid J_0 = j\right\} \int_0^\infty P\left\{Z_{N(t)} \in dv \mid Z_0 = z_0, J_{N(t)} = k, J_0 = j\right\}$$
$$.P\left\{Z(t) \leq z \mid Z(T_{N(t)}) = v, J(T_{N(t)}) = k\right\}$$
$$= P\{J(t) = k \mid J(0) = j\} \int_0^\infty P\left\{Z_{N(t)} \in dv \mid Z_0 = z_0, J_{N(t)} = k, J_0 = j\right\}$$
$$.P\left\{[v + y(k)(t - T_{N(t)})]^+ \leq z \mid J(T_{N(t)}) = k\right\}$$

The statement follows from the fact that as $t \to \infty$ the following results hold.

$$P\{J(t) = k \mid J(0) = j\} \to \pi_k \quad \text{a.s.}$$

Given $J(T_{N(t)}) = k$, $(t - T_{N(t)})$ has by limit distribution T_1, given $J(0) = k$, so that

$$P\left\{[v + y(k)(t - T_{N(t)})]^+ \leq z \mid J(T_{N(t)}) = k\right\} \;\to\; P\left\{[v + Y_1]^+ \leq z \mid J(0) = k\right\}$$
$$= \; P\{Y_1 \leq z - v \mid J(0) = k\}.$$

Since $\overline{y} < 0$ we have $\mu^\star < 0$, and $N(t) \to \infty$. Thus using Lemma 8 we conclude that

$$P\left\{Z_{N(t)} \in dv \mid Z_0 = z_0, J_{N(t)} = k, J_0 = j\right\} \to P\{Z_\infty^\star \in dv \mid J_\infty^\star = k\}. \quad \square$$

In case $\overline{y} < 0$, we denote by (Z_∞, J_∞) the limit random variable of $(Z,J)(t)$, which in view of Theorem 22 is independent of the initial distribution.

Corollary 8 *If $\overline{y} < 0$ we have the following:*

(a). For $z \geq 0$ we have

$$\begin{aligned} P\{Z_\infty \leq z \mid J_\infty = k\} &= P\{Z_\infty^\star \leq z \mid J_\infty^\star = k\} \\ &\cdot \left(1 - E\left[e^{-\frac{y(k)}{q_{kk}}(Z_\infty^\star - z)} \mid Z_\infty^\star \leq z,\, J_\infty^\star = k\right]\right) \quad (k \in E_0) \end{aligned} \tag{4.52}$$

$$\begin{aligned} P\{Z_\infty > z \mid J_\infty = k\} &= P\{Z_\infty^\star > z \mid J_\infty^\star = k\} \\ &\cdot \left(1 - E\left[e^{-\frac{y(k)}{q_{kk}}(Z_\infty^\star - z)} \mid Z_\infty^\star > z,\, J_\infty^\star = k\right]\right) \quad (k \in E_1) \end{aligned} \tag{4.53}$$

(b). We have

$$P\{Z_\infty = 0 \mid J_\infty = k\} = \begin{cases} 0 & k \in E_0 \\ E\left[e^{-\frac{y(k)}{q_{kk}} Z_\infty^\star} \mid J_\infty^\star = k\right] & k \in E_1 \end{cases} \tag{4.54}$$

Proof: (a). Let $z \geq 0$ and $k \in E_0$. From (4.51) we have

$$\begin{aligned} P\{Z_\infty \leq z \mid J_\infty = k\} &= \int_0^z P\left\{Z_\infty^\star \in dv \mid J_\infty^\star = k\right\} P\left\{Y_1 \leq z - v \mid J(0) = k\right\} \\ &= \int_0^z P\left\{Z_\infty^\star \in dv \mid J_\infty^\star = k\right\} \left(1 - e^{\frac{y(k)}{q_{kk}}(z-v)}\right) \\ &= P\{Z_\infty^\star \leq z \mid J_\infty^\star = k\} \\ &\quad \cdot \left(1 - \int_0^z P\left\{Z_\infty^\star \in dv \mid Z_\infty^\star \leq z,\, J_\infty^\star = k\right\} e^{-\frac{y(k)}{q_{kk}}(v-z)}\right) \\ &= P\{Z_\infty^\star \leq z \mid J_\infty^\star = k\} \left(1 - E\left[e^{-\frac{y(k)}{q_{kk}}(Z_\infty^\star - z)} \mid Z_\infty^\star \leq z,\, J_\infty^\star = k\right]\right). \end{aligned}$$

This gives (4.52). The proof of (4.53) is similar.

(b). From (4.51) we have

$$P\{Z_\infty = 0 \mid J_\infty = k\} = \int_0^\infty P\left\{Z_\infty^\star \in dv \mid J_\infty^\star = k\right\} P\left\{Y_1 \leq -v \mid J(0) = k\right\}$$

$$= \begin{cases} 0 & k \in E_0 \\ \int_0^\infty P\left\{Z_\infty^\star \in dv \mid J_\infty^\star = k\right\} e^{-\frac{y(k)}{q_{kk}}v} & k \in E_1 \end{cases}$$

which proves the statement. □

We denote by $I_k(t)$ the unsatisfied demand in state k during $(0, t]$, so that

$$I_k(t) = \int_0^t y(J(s))^- 1_{\{Z(s)=0,\, J(s)=k\}} ds = y(k)^- \int_0^t 1_{\{Z(s)=0,\, J(s)=k\}} ds. \qquad (4.55)$$

If $k \in E_0$, then $y(k)^- = 0$ and $I_k(t) = 0$. If $k \in E_1$, we have the following important result for the performance analysis of the system.

Corollary 9 *If $\overline{y} < 0$ and $k \in E_1$, then*

$$\lim_{t\to\infty} \frac{I_k(t)}{t} = -y(k)\, \pi_k\, E\left[e^{-\frac{y(k)}{q_{kk}} Z_\infty^\star} \mid J_\infty^\star = k\right] \qquad (4.56)$$

and

$$\lim_{t\to\infty} \frac{I_k(t)}{I(t)} = \frac{y(k)\, \pi_k\, E\left[e^{-\frac{y(k)}{q_{kk}} Z_\infty^\star} \mid J_\infty^\star = k\right]}{\sum_{j\in E_1} y(j)\, \pi_j\, E\left[e^{-\frac{y(j)}{q_{jj}} Z_\infty^\star} \mid J_\infty^\star = j\right]}. \qquad (4.57)$$

Proof: Using Theorem 22 we conclude that

$$\frac{1}{t}\int_0^t 1_{\{Z(s)=0, J(s)=k\}} ds \longrightarrow P\{Z_\infty = 0,\, J_\infty = k\}.$$

This implies (4.56) in view of (4.54)-(4.55). Also, since $I(t) = \sum_{j\in E_1} I_j(t)$, (4.57) follows by using (4.56). □

Chapter 5

A storage model with jump input process

In this chapter we consider the storage model introduced in subsection 1.4.2. The input process (X, J) is a Markov-compound Poisson process with no drift; its infinitesimal generator is given by (1.14), with $a = 0$, whereas the input process (D, J) is given by

$$D(t) = \int_0^t d(J(s))\, ds \tag{5.1}$$

with d being a nonnegative function. Here the Markov component J is a pure jump process, with infinitesimal generator given by (3.4).

The storage equation of the model is

$$Z(t) = Z(0) + X(t) - D(t) + \int_0^t d(J(s))\, 1_{\{Z(s) \leq 0\}}\, ds \tag{5.2}$$

and, from Theorem 13, it follows that the unsatisfied demand is given by

$$I(t) = \int_0^t d(J(s))\, 1_{\{Z(s) \leq 0\}}\, ds = \left[Z(0) + \inf_{0 \leq \tau \leq t} Y(\tau)\right]^- \tag{5.3}$$

with Y being the net input.

We note that from Theorem 5.3 it follows that

$$I(t) = \int_0^t d(J(s))\, 1_{\{Z(s)=0\}}\, ds \tag{5.4}$$

since $Z(t) \geq 0$, $\forall t \geq 0$. If we replace equation (5.2) by

$$Z(t) = Z(0) + X(t) - D(t) + \int_0^t d(J(s))\, 1_{\{Z(s)=0\}}\, ds \tag{5.5}$$

(5.3) is also a solution of (5.5). However, (5.5) may not have an unique solution as we show in the following.

Example 10 Consider $E = \{1\}$, $\Sigma = (0)$, $M_{11}\{0\} = 1$, $d(1) = 1$ and $Z(0) = 0$ a.s.. Note that $X(t) = 0$ and $D(t) = t$, a.s.. Trivially, $Z(t) = -t$ and $I(t) = 0$ is a solution of (5.5) (in addition to the solution $Z(t) = 0$ and $I(t) = t$, correspondent to (5.3)). Note that the uniqueness of the solution (5.3) of (5.5) has been systematically referred to in the literature for models that include this example.

□

The busy period problem is investigated in section 5.1. If T denotes the busy period, it follows that $(T, J \circ T)$ is an MAP with both drift and jumps. Section (5.2) contains some additional notations and results of a technical nature. Transforms of the various processes of interest are obtained in sections 5.3-5.5. The steady state behaviour of the model is investigated in section 5.6.

5.1 The busy period process

The busy period is defined as the time the storage system would become empty if the initial level was $x(\geq 0)$. Denoting this time as $T(x, k)$, when the initial state of J is k, we have

$$T(x,k) \ = \inf \{t \ : Z(t) \ = 0\} \quad \text{on} \quad \{Z(0) = x, J(0) = k\} \tag{5.6}$$

for $x \in I\!R_+$, $k \in E$, with the convention that the infimum of an empty set is $+\infty$. We are also interested in the state of J at the time of emptiness, namely $J \circ T(x,k)$. If $T(x,k) = \infty$ we define $J \circ T = \Delta$. Of special interest is the situation when the initial level x has the distribution M_{jk} with $j \in E$. This busy period may represent (for example) the busy period initiated by the input arriving with a transition from state j to state k in J. In this case, we shall call this the busy period initiated by the transition j to k and denote it as T_{jk}. Let

$$G_{jk}(A,l) = P\{(T_{jk}, J \circ T_{jk}) \in A \times \{l\}\} \tag{5.7}$$

be the distribution of $(T_{jk}, J \circ T_{jk})$. We have

$$G_{jk}(A,l) = \int_0^{\infty+} P\{(T, J \circ T)(x,k) \in A \times \{l\}\} \, M_{jk}\{dx\} \tag{5.8}$$

for any Borel subset A of $[0,\infty]$, $j,k \in E$ and $l \in E \bigcup \Delta$. Note that T_{jk} has possibly an atom at (∞, Δ). We let $G_{jkl}(A) = G_{jk}(A,l)$ and for simplicity we attach the atom at (∞, Δ) (with value $G_{jk\Delta}\{\infty\}$) to the measure G_{jkj}, so that for $j,k,l \in E$ with j,k fixed,

$$\sum_{l \neq j} G_{jkl}([0,\infty)) + G_{jkj}([0,+\infty]) = 1. \tag{5.9}$$

We next consider the process $(T, J \circ T) = \{(T, J \circ T)(x), \, x \geq 0\}$ where $T(x)$ is now defined as $T(x, J(0))$ with $J(0)$ having an arbitrary distribution. Thus we are studying the effect of the initial storage level, regardless of the initial state of J. We have the following.

Lemma 9 *$(T, J \circ T)$ is a MAP with infinitesimal generator given by*

$$\mathcal{A} = \frac{1}{d(j)} \left[\frac{\partial}{\partial x} f(x,j) + \sum_{l \in E} \int_0^{\infty +} [f(x+y,l) - f(x,j)] \sum_{k \in E} \sigma_{jk} G_{jkl}\{dy\} \right] \quad (5.10)$$

where $G_{jkl}\{\infty\} = 0$ for $l \neq j$.

Proof: On the set $\{J(0) = j\}$, $j \in E$, consider for $A \in \mathcal{B}([0, +\infty))$ the events

$$C(j,h;A,l) = \{T(h) \in A, J \circ T(h) = l\}$$

$$H(j,h;A,l) = \left\{ \begin{array}{l} T(h) \in A, J \circ T(h) = l \text{ and on } [0, B(h)] \\ \text{there are no transitions in } J \end{array} \right\}$$

and for $k \in E$,

$$E_k(j,h;A,l) = \left\{ \begin{array}{l} T(h) \in A, J \circ T(h) = l \text{ and the first transition in } J \\ \text{is to state } k \text{ and occurs before time } B(h) \end{array} \right\}.$$

It is easily seen that with $\lambda_j = \sum_{k \in E} \sigma_{jk}$ for $j \in E$ we have

$$P\left[C(j,h;A,l)\right] = \sum_{k \in E} P\left[E_k(j,h;A,l)\right] + P\left[H(j,h;A,l)\right]$$

$$= \sum_{k \in E} \int_0^{\frac{h}{d(j)}} \sigma_{jk} e^{-\lambda_j \tau} \int_0^{\infty +} \sum_{m \in E} G_{jkm}\left\{A - \tau - ds\right\} P[C(m, h - \tau d(j); ds, l)] d\tau$$
$$+ \delta_{jl} \left[1 - \lambda_j \frac{h}{d(j)}\right] 1_{\left\{\frac{h}{d(j)} \in A\right\}} + o(h).$$

We are now able to obtain the infinitesimal generator of the process $(T, J \circ T)$,

$$\mathcal{A}_{(T, J \circ T)} f(x,j) = \lim_{h \to 0+} \left[\frac{1}{h} \sum_{l \in E} \int_0^{\infty +} [f(x+y,l) - f(x,j)] P\left[C(j,h;dy,l)\right] \right]$$

$$\frac{1}{d(j)} \lim_{h \to 0+} \left[\int_0^{\infty +} \frac{f(x+y,j) - f(x,j)}{h/d(j)} \left(1 - \lambda_j \frac{h}{d(j)}\right) 1_{\left\{\frac{h}{d(j)} \in dy\right\}} \right]$$

$$+\sum_{l\in E}\int_0^{\infty+}[f(x+y,l)-f(x,j)]\sum_{k\in E}\frac{\sigma_{jk}}{d(j)}$$

$$\lim_{h\to 0+}\left[\frac{d(j)}{h}\int_0^{\frac{h}{d(j)}}e^{-\lambda_j\tau}\int_0^{y-\tau}\sum_{m\in E}G_{jkm}\{dy-\tau-ds\}\ P[C(m,h-\tau d(j);ds,l)]d\tau\right]$$

$$\frac{1}{d(j)}\frac{\partial}{\partial x}f(x,j)+\sum_{l\in E}\int_0^{\infty+}[f(x+y,l)-f(x,j)]\sum_{k\in E}\frac{\sigma_{jk}}{d(j)}G_{jkl}\{dy\}$$

which is the desired result. □

Example 11 As an application of Lemma 9 we derive the infinitesimal generator of the inverse of the demand process $\{(B,J\circ B)(x),\ x\geq 0\}$ which coincides with the storage busy period process if no input arrives to the system, i.e.

$$M_{jk}=\varepsilon_0,\quad \forall j,k\in E$$

where ε_0 is a distribution concentrated at the origin. For $A\in\mathcal{B}([0,+\infty])$, this implies that

$$G_{jkl}(A)=\begin{cases}0 & \text{if } k\neq l\\ \varepsilon_0(A) & \text{if } k=l\end{cases}\tag{5.11}$$

so that the generator of $\{(B,J\circ B)(x),\ x\geq 0\}$ is given by

$$\mathcal{A}_{(B,J\circ B)}f(x,j)=\frac{1}{d(j)}\left[\frac{\partial f}{\partial x}(x,j)+\sum_{l\neq j}\sigma_{jl}\left[f(x,l)-f(x,j)\right]\right]$$

thus proving Theorem 15 (f). □

We consider for $j\in E$ the random variable $(\tilde{T},J\circ\tilde{T})_j$, where $\tilde{T}_j$ represents the duration of the storage busy period initiated by a new input if at present J is in state j. Denote by $\tilde{G}_j$ its distribution and by $p^{\star}_{jl}$ the probability that this busy period ends in state l. Our setting implies that for $A\in\mathcal{B}([0,\infty]),l\in E$

$$\tilde{G}_{jl}(A)\equiv\tilde{G}_j(A,\{l\})=\sum_{k\in E}p_{jk}G_{jkl}(A)\tag{5.12}$$

and

$$p^\star_{jl} = P\{J \circ \tilde{T}_j = l\} = \tilde{G}_{jl}([0,\infty]). \tag{5.13}$$

In terms of the definitions above, the inifinitesimal generator of $(T, J \circ T)$ (given by (5.10)) can be expressed, with $\tilde{G}_{jl}\{\infty\} = 0$ for $l \neq j$, as follows:

$$\mathcal{A}_{(T,J\circ T)}\, f(x,j) = \frac{1}{d(j)}\frac{\partial f}{\partial x}(x,j) + \frac{\lambda_j}{d(j)} \sum_{l\in E} \int_0^{\infty+} \left[f(x+y,l) - f(x,j)\right] \tilde{G}_{jl}\{dy\}. \tag{5.14}$$

We elaborate a bit more to increase the interpretability of (5.14). Suppose the present state of the MJP J is j. We are now able to consider the (conditional) distribution of the length of the busy period initiated by a new input given that at the end of the same busy period the Markov chain J is in a given state l. We denote such a variable by $T^\star_{jl}$ and its distribution by $M^\star_{jl}$ and for completeness we set $T^\star_{jl} = 0$ if $p^\star_{jl} = 0$. Thus, for $A \in \mathcal{B}([0,\infty])$, $j, l \in E$,

$$M^\star_{jl}(A) = \begin{cases} \frac{\tilde{G}_{jl}(A)}{\tilde{G}_{jl}([0,\infty])} & p^\star_{jl} \neq 0 \\ 1_{\{0\in A\}} & p^\star_{jl} = 0 \end{cases} \tag{5.15}$$

where we note that

$$p^\star_{jl} M^\star_{jl}(A) = \tilde{G}_{jl}(A). \tag{5.16}$$

We define for $j, l \in E$,

$$\lambda^\star_j = \frac{\lambda_j}{d(j)}, \qquad \sigma^\star_{jl} = \lambda^\star_j\, p^\star_{jl} \tag{5.17}$$

and let $J^\star$ be a MJP on the state space E with intensity matrix $\Sigma^\star = D^{-1}\Lambda P^\star$, where $P^\star = (p^\star_{jk})$, $D = (\delta_{jk}\, d(j))$ and $\Lambda = (\lambda_j\, \delta_{jk})$. As a direct consequence of Lemma 9, or even more imediately from (5.14) we have:

Theorem 23 *$(T, J \circ T)$ is an MAP whose generator is given by*

$$\mathcal{A}_{(T,J\circ T)}\, f(x,j) = \frac{1}{d(j)}\frac{\partial f}{\partial x}(x,j) + \sum_{l\in E}\int_0^\infty \left[f(x+y,l) - f(x,j)\right]\sigma^\star_{jl} M^\star_{jl}\{dy\}. \tag{5.18}$$

Thus $(T, J^\star)$ is an MCPP with drift function $[d(j)]^{-1}, j \in E$, and jumps with distribution $M^\star_{jl}$ occuring at rate $\sigma^\star_{jl}$. □

Theorem 23 provides insights that lead to a more meaningful probabilistic interpretation of the results obtained in Lemma 9. Among other facts, it is easy to see that the motion of the process $(T, J^\star)$ is related with the motion of the process $(T, J \circ T)$ (as it should) but the perspective from which we see the busy period evolving is quite different in the two processes. Some comments are in order.

Remark 4 The transitions between states of E occuring during a busy period initiated by a new input in $(T, J \circ T)$ are eliminated in the process $(T, J^\star)$. In addition, the jumps in the additive component of the process $(T, J^\star)$ (the busy period) are lenghts of busy periods initiated by new arrivals. □

Remark 5 Time is rescaled; when $J^\star$ is in state j, time is measured in units $[d(j)]^{-1}$ if refered to the Markov chain J. Thus the busy period process increases locally at rate equal to the reciprocal of the present demand rate $[d(j)]^{-1}$. In particular, the busy period local timing is only a function of the present state of the MJP $J^\star$. □

Remark 6 With respect to the length of busy periods initiated by new arrivals, the actual values of the jump rates $\{\lambda_j,\ j \in E\}$ and of the demand rates $\{d(j),\ j \in E\}$ are immaterial; these quantities impact the length of busy periods initiated by new arrivals only through the ratios $\{\lambda_j^\star = \frac{\lambda_j}{d(j)}, j \in E\}$. This expresses a relativity property of the jump rates and demand rates. Neverthless the demand rates impact the length of busy periods (locally) as a function of their absolute values in terms of the drift of the busy period process. □

As shown in Theorem 23, the storage busy period process $(T, J^\star)$ is an MCPP, thus a relatively simple process. Neverthless, at this point, we cannot make direct use of this property since we do not know the distributions $M_{jl}^\star$ of the jumps, nor the values of the probabilities $p_{jl}^\star$. In the following our aim will be to get a better description of the objects we have just mentioned.

5.2 Some more notations and results

In this section we mainly introduce some more notations and results that we use in the following sections. The reader may prefer to proceed directly to the next section now and return whenever the need arises.

Definition 3 If for $j, l \in E$, $A_{jl}\{dx\}$ is a finite measure on $[0, \infty]$ and $\theta > 0$, then we denote by $A(\theta) = (A_{jl}(\theta))$ the Laplace transform matrix of the matrix measure $A = (A_{jl}\{dx\})$,

$$A(\theta) = \left(\int_0^\infty e^{-\theta x} A_{jl}\{dx\} \right). \tag{5.19}$$

For completeness, if $A = (A_{jl})$ is a matrix of nonnegative numbers we consider that A_{jl} is an atomic measure concentrated solely at the origin to which it assigns measure A_{jl}, so that $A(\theta) \equiv A$. □

Definition 4 We let

$$S = \{f = \left(f_{jk}\right), j, k \in E \text{ such that } f_{jk} : \mathbb{R}_+ \to \mathbb{R} \text{ is bounded and continuous}\}$$

and we define a norm in S,

$$\| f \| = \sup_{k \in E} \sum_{j \in E} \| f_{jk} \|,$$

where $\| f_{jk} \| = \sup_{x \geq 0} | f_{jk}(x) |$. Then $(S, \| \cdot \|)$ is a Banach space. We define, for $f \in S$ and $s > 0$,

$$\mathcal{L}(f, s) = \left(\int_0^\infty e^{-st} f_{jk}(t) dt\right). \quad \square$$

Definition 5 Suppose J is a pure jump process on a countable state space E, that V and W are nonnegative process such that (V, J) and (W, J) are adapted to a given filtration $\mathcal{F} = \{\mathcal{F}_t, t \geq 0\}$, and that for all $x \geq 0$, $T(x)$ is an $\mathcal{F}$-stopping time. We define (with the operator E denoting the usual expectation and A being $\mathcal{F}_0$-measurable) the following types of expectation matrices:

$$\mathbb{E}_J[W(t)] = (E[W(t); J(t) = k | J(0) = j])$$

$$\mathbb{E}_J[W(t)|A] = (E[W(t); J(t) = k | A, J(0) = j])$$

$$\mathbb{E}_J[T(x)] = (E[T(x); J \circ T(x) = k | J(0) = j]).$$

In a similar way we define for $\theta, \theta_1, \theta_2 > 0$ the following types of Laplace transform matrices:

$$\psi_W(s; \theta) = \mathcal{L}\left(\mathbb{E}_J\left[e^{-\theta W(.)}\right], s\right) = \left(\int_0^\infty e^{-st} E\left[e^{-\theta W(t)}; J(t) = k | J(0) = j\right] dt\right),$$

$$\psi_{(W,S)}(s; \theta_1, \theta_2) = \mathcal{L}\left(\mathbb{E}_J\left[e^{-\theta_1 W(.) - \theta_2 S(.)}\right], s\right);$$

obvious definitions are considered for the conditional case (e.g. $\psi_{W|A}(s; \theta)$). $\square$

We define for $\theta > 0$

$$\Phi(\theta) = \Lambda\left[I - \Gamma(\theta)\right], \qquad \Phi^{\star}(\theta) = D^{-1}\Lambda\left[I - \Gamma^{\star}(\theta)\right] \tag{5.20}$$

where,

$$\Gamma(\theta) = \left(p_{jk}M_{jk}(\theta)\right), \qquad \Gamma^{\star}(\theta) = \left(p^{\star}_{jk}M^{\star}_{jk}(\theta)\right) \tag{5.21}$$

and in addition, we let

$$\Psi(\theta) = \Phi(\theta) + D\theta, \qquad \Psi^{\star}(\theta) = \Phi^{\star}(\theta) + D^{-1}\theta. \tag{5.22}$$

Suppose J is an MJP on a countable state space E with intensity matrix Σ, and (W, J) is an MCPP with nonnegative jumps whose infinitesimal generator is given by

$$\mathcal{A}_{(W,J)}\, f(x,j) = d(j)\frac{\partial f}{\partial x}(x,j) + \sum_{k\in E}\sigma_{jk}\int_0^{\infty}[f(x+y,k) - f(x,j)]M_{jk}\{dy\}. \tag{5.23}$$

The distribution measures M_{jk} are thus concentrated on $[0,\infty)$. We assume that $\{d(j)\}$ and $\{\lambda_j\}$ are uniformly bounded.

Lemma 10 *The process W with generator given by equation (5.23) has the following Laplace transform matrix*

$$I\!\!E_J\left[e^{-\theta W(t)}\right] = e^{-t\Psi_W(\theta)} = e^{-t[\Phi_W(\theta)+D_W\theta]} \tag{5.24}$$

where $\Psi_W = \Psi$, $\Phi_W = \Phi$, $D_W = D$, so that

$$\Psi_W(\theta) = \Lambda\left[I - \Gamma(\theta)\right] + D\theta. \tag{5.25}$$

Proof: The result follows similarly to Theorem 19 in Prabhu [31]. □

Remark 7 The Laplace transform matrix of the additive component of an MCPP with nonnegative jumps is a direct extension (to a matrix-like form) of the Laplace transform of a compound-Poisson process, as it is easily seen from equations (5.24) and (5.25).

The presence of the matrices Λ and D in equations (5.24) and (5.25) expresses the fact that the (compound) arrival rate and the drift of the additive component of an MCPP depend on the state of the underlying Markov chain, as opposed to the compound Poisson case where they are fixed constants. □

Remark 8 Assume $d_+ < \infty$, $s, \theta > 0$, and let $Q = \Sigma - \Lambda$ and 0 denote the null process $0(t) \equiv 0$ then, using Lemma 10, we have:

$$\pi(t) = (P\left[J(t) = k | J(0) = j\right]) = \mathbb{E}_J\left[e^{-\theta 0(t)}\right] = e^{tQ} \tag{5.26}$$

$$\hat{\pi}(s) = \int_0^\infty e^{-st}\pi(t)\,dt = \psi_0(s;.) = [sI - Q]^{-1} \tag{5.27}$$

$$\mathbb{E}_J\left[e^{-\theta X(t)}\right] = e^{-t\Phi(\theta)}, \qquad \psi_X(s;\theta) = [sI + \Phi(\theta)]^{-1} \tag{5.28}$$

$$\mathbb{E}_J\left[e^{-\theta D(t)}\right] = e^{-t[D\theta - Q]}, \qquad \psi_D(s;\theta) = [sI + D\theta - Q]^{-1} \tag{5.29}$$

and,

$$\mathbb{E}_J\left[e^{-\theta Y(t)}\right] = e^{-t[\Phi(\theta) - D\theta]}, \qquad \psi_Y(s;\theta) = [sI - D\theta + \Phi(\theta)]^{-1}. \;\square \tag{5.30}$$

Next we give a definition that is very useful for the remainder of the paper.

Definition 6 If $-\eta$ is the generator of a semigroup as described in Lemma 2 and $A(\theta)$ is a Laplace transform matrix as described in Definition 3, we define

$$A \circ \eta = \left(\int_0^\infty \sum_{k \in E} A_{jk}\{dx\}\left(e^{-x\eta}\right)_{kl}\right). \quad \square$$

Remark 9 Note that if A is a matrix of nonnegative numbers, then $A \circ \eta = A$. Very important is the fact that, if in Definition 6 A is such that $(A_{jl}[0, \infty])$ is a substochastic matrix and $-\eta$ is the generator of a contraction semigroup, then $A \circ \eta$ is a contraction map from S to S. □

5.3 Laplace transform of the busy period

The process $(T, J^\star)$ is an MCPP with nonnegative jumps and positive drift. We now use the definitions and results of the previous section to derive the Laplace transform matrix of T, but in order to use Lemma 10 we need to add some additional conditions to the storage process, which we call the *standard conditions*,

$$0 < d_- \leq d_+ < \infty \qquad \text{and} \qquad \lambda_+ = \sup_{j \in E} \lambda_j < \infty. \tag{5.31}$$

In the rest of the paper we assume implicitly that the *standard conditions* hold.

Theorem 24 *The Laplace transform matrix of the storage busy period is given by*

$$\mathbb{E}_J \left[e^{-\theta T(x)} \right] = e^{-x \Psi^\star(\theta)} \tag{5.32}$$

where as previously defined,

$$\Psi^\star(\theta) = \Phi^\star(\theta) + D^{-1}\theta = D^{-1}\Lambda \left[I - \Gamma^\star(\theta) \right] + D^{-1}\theta. \tag{5.33}$$

Moreover, and more importantly, $\Gamma^\star$ *satisfies the functional equation*

$$\Gamma^\star(\theta) = \Gamma \circ \Psi^\star(\theta) \tag{5.34}$$

so that

$$\Gamma^\star(\theta) = \Gamma \circ \left(D^{-1}\theta + D^{-1}\Lambda \left[I - \Gamma^\star(\theta) \right] \right) \tag{5.35}$$

and

$$D\Psi^\star(\theta) = \theta I + \Phi \circ \Psi^\star(\theta). \tag{5.36}$$

Proof: From the conditions of the theorem, Theorem 23 and since (as a consequence of the standard conditions) the rates $\{\lambda_j^\star = \frac{\lambda_j}{d(j)}\}$ and $\{[d(j)]^{-1}\}$ are uniformly bounded, we conclude by Lemma 10 that $I\!E_J\left[e^{-\theta T(x)}\right] = e^{-x\Psi_T(\theta)}$, where

$$\Psi_T(\theta) = \Phi_T(\theta) + D_T\theta = D^{-1}\Lambda\left[I - \Gamma^\star(\theta)\right] + D^{-1}\theta = \Psi^\star(\theta)$$

and

$$\begin{aligned}\Gamma^\star(\theta) &= \left(\int_0^\infty \sum_{k\in E} p_{jk}M_{jk}\{dx\}E\left[e^{-\theta T(x)}; J\circ T(x) = l|J(0)=k\right]\right)\\ &= \left(\int_0^\infty \sum_{k\in E} p_{jk}M_{jk}\{dx\}\left(e^{-x\Psi^\star(\theta)}\right)_{kl}\right) = \Gamma\circ\Psi^\star(\theta).\end{aligned}$$

Therefore, equations (5.32), (5.33), and (5.34) follow, and (5.35) is an imediate consequence of equations (5.33) and (5.34). From (5.33), (5.34) and Remark 9, we have

$$D\Psi^\star(\theta) = \theta I + \Lambda\left[I - \Gamma\right]\circ\Psi^\star(\theta) = \theta I + \Phi\circ\Psi^\star(\theta). \qquad \square$$

As (5.33) shows, to study the distribution of the busy period initiated by a given initial storage level all we need to study is the distribution of the busy period initiated by new arrivals (which may or may not be associated with a transition on J). Such information is given in our study by $\Gamma^\star(\theta)$. We show in Theorem 25 how $\Gamma^\star(\theta)$ may be computed.

Theorem 25 *There exists a unique bounded continuos solution $\Gamma^\star(\theta)$ of the functional equation (5.35). Moreover, if for $A \in S$ we define*

$$\Psi^\star(A(\theta)) = D^{-1}\theta + D^{-1}\Lambda\left[I - A(\theta)\right] \tag{5.37}$$

and let $\Gamma_0^\star(\theta) = 0$ *say, and*

$$\Gamma_{n+1}^\star(\theta) = \Gamma \circ \Psi^\star(\Gamma_n^\star(\theta)), \quad n = 0, 1, 2, ...$$

then

$$\Gamma^\star(\theta) = \lim_{n\to\infty} \Gamma_n^\star(\theta).$$

Proof: In view of (5.37), (5.35) can be written as

$$\Gamma^\star(\theta) = \Gamma \circ \Psi^\star(\Gamma^\star(\theta))$$

which is a fixed point equation. Moreover, by Remark 9, $\Gamma \circ \Psi^\star$ is a contraction map from S to S. The rest of the proof now follows from a Banach fixed point theorem such as Theorem 3.8.2 in Friedman [12]. □

Remark 10 Consider an $M/G/1$ system with arrival rate λ and let $G^\star$ and $B^\star$ be the Laplace-Stieltjes transforms of the service time and busy period initiated by a new arrival, respectively. Let $\phi = \lambda(1 - G^\star)$ and $E[e^{-sT(x)}] = e^{-x\eta(s)}$, where $T(x)$ is the duration of a busy period initiated by a given amount of work x. It should be noted that equation (5.35) is a generalization of Takács functional equation for the $M/G/1$ busy period initiated by a new input,

$$B^\star(s) = G^\star(s + \lambda[1 - B^\star(s)])$$

whereas equation (5.36) is a generalization of the functional equation associated with the busy period initiated by a given amount of work,

$$\eta(s) = s + \phi\,(\eta(s))$$

as given in Prabhu [30]. □

Remark 11 From Theorem 24 and since $(B, J \circ B)$ is equal to $(T, J \circ T)$ if there is no input (i.e. the distribution measures M_{jk} are probability measures concentrated at the origin. In this case $\Phi(\theta) \equiv \Lambda - \Sigma = -Q$ and, by (5.33), $\Psi^\star(\theta) = D^{-1}(\theta I - Q)$), it follows that

$$I\!E_J\left[e^{-\theta B(x)}\right] = e^{-xD^{-1}[\theta I - Q]}, \qquad \psi_B(s;\theta) = \left[sI + D^{-1}\left(\theta I - Q\right)\right]^{-1}. \quad \square$$

5.4 The unsatisfied demand and demand rejection rate

We note that $I(t)$ is the local time at zero of the storage level, where the clock has variable speed; the clock speed being $d(j)$ when J is in state j. We let $\zeta(t)$ be the demand rejection rate at time t, so that

$$\zeta(t) = d(J(t))\, 1_{\{Z(t)=0\}} \tag{5.38}$$

and

$$I(t) = \int_0^t \zeta(s)\, ds. \tag{5.39}$$

Theorem 26 *For $\theta, s > 0$, we have:*

$$\mathcal{L}\left(I\!E_J\left[e^{-\theta I(.)}\zeta(.)|Z(0) = x\right], s\right) = e^{-x\Psi^\star(s)}\left[\theta I + \Psi^\star(s)\right]^{-1}. \tag{5.40}$$

Proof: We have, using the fact that $\zeta(t) = 0$ for $t < T(x)$,

$$\mathcal{L}\left(I\!E_J\left[e^{-\theta I(.)}\zeta(.)|Z(0) = x\right], s\right)$$
$$= \left(E\int_{T(x)}^{\infty}\left[e^{-st-\theta I(t)}\zeta(t); J(t) = k|Z(0) = x, J(0) = j\right]\, dt\right)$$

Now we carry out the transformation $I(t) = \tau$. We have $t = T(x+\tau) = T(x) + T(\tau) \circ \phi_{T(x)}$ (where ϕ is a shift operator as in Definition 2) and $d\tau = dI(t) = \zeta(t)dt$. Therefore the right hand side of the last expression becomes

$$\begin{aligned}
&\left(E \int_0^\infty \left[e^{-sT(x)-s[T(\tau)\circ\phi_{T(x)}]-\theta\tau}; J(t) = k | Z(0) = x, J(0) = j\right] d\tau\right) \\
&= \left(\sum_{l \in E} \left(\mathbb{E}_J \left[e^{-sT(x)}\right]\right)_{jl} \left(\int_0^\infty e^{-\theta\tau} \mathbb{E}_J \left[e^{-sT(\tau)}\right] d\tau\right)_{lk}\right) \\
&= \left(\sum_{l \in E} \left(e^{-x\Psi^\star(s)}\right)_{jl} \left(\int_0^\infty e^{-\tau[\theta I + \Psi^\star(s)]} d\tau\right)_{lk}\right) \\
&= \left(\sum_{l \in E} \left(e^{-x\Psi^\star(s)}\right)_{jl} \left([\theta I + \Psi^\star(s)]^{-1}\right)_{lk}\right) \\
&= e^{-x\Psi^\star(s)} [\theta I + \Psi^\star(s)]^{-1}. \qquad \square
\end{aligned}$$

Corollary 10 *For $\theta, s > 0$, we have:*

$$\mathcal{L}\left(\mathbb{E}_J [\zeta(.) | Z(0) = x], s\right) = e^{-x\Psi^\star(s)} [\Psi^\star(s)]^{-1} \tag{5.41}$$

and

$$\mathcal{L}\left(\mathbb{E}_J \left[1_{\{Z(.)=0\}} | Z(0) = x\right], s\right) = e^{-x\Psi^\star(s)} [\Psi^\star(s)]^{-1} D^{-1}. \tag{5.42}$$

Proof: By letting $\theta \to 0$ in Theorem 26, (5.41) follows. If $J(t) = k$, $\zeta(t) = d(k) 1_{\{Z(t)=0\}}$, thus

$$\begin{aligned}
\mathcal{L}\left(\mathbb{E}_J [\zeta(.) | Z(0) = x], s\right) &= \mathcal{L}\left(\mathbb{E}_J \left[1_{\{Z(.)=0\}} | Z(0) = x\right] D, s\right) \\
&= \mathcal{L}\left(\mathbb{E}_J \left[1_{\{Z(.)=0\}} | Z(0) = x\right], s\right) D
\end{aligned}$$

(5.42) follows from this and (5.41). $\square$

5.5 The storage level and unsatisfied demand

In this section we study the process $((Z, I), J)$; we are thus interested in the storage level and unsatisfied demand processes along with their interaction. We start this study with a lemma.

Lemma 11 *We have for* $\theta_1, \theta_2 > 0$

$$e^{-\theta_1 Z(t) - \theta_2 I(t)} = e^{-\theta_1 [Z(0)+Y(t)]} - (\theta_1 + \theta_2) \int_0^t e^{-\theta_1 [Y(t)-Y(\tau)] - \theta_2 I(\tau)} \zeta(\tau)\, d\tau. \quad (5.43)$$

Proof: For $\theta > 0$, we have,

$$\theta \int_0^t e^{-\theta I(\tau)} \zeta(\tau)\, d\tau = \left[-e^{-\theta I(\tau)}\right]_0^t = 1 - e^{-\theta I(t)}. \quad (5.44)$$

Since $Z(t) = Z(0) + Y(t) + I(t)$ and using (5.44), we have for $\theta_1, \theta_2 > 0$,

$$\begin{aligned} e^{-\theta_1 Z(t) - \theta_2 I(t)} &= e^{-\theta_1 [Z(0)+Y(t)]} e^{-(\theta_1+\theta_2) I(t)} \\ &= e^{-\theta_1 [Z(0)+Y(t)]} \left[1 - (\theta_1 + \theta_2) \int_0^t e^{-(\theta_1+\theta_2) I(\tau)} \zeta(\tau)\, d\tau\right] \\ &= e^{-\theta_1 [Z(0)+Y(t)]} - (\theta_1 + \theta_2) \int_0^t e^{-\theta_1 [Y(t)-Y(\tau)] - \theta_2 I(\tau)} \zeta(\tau)\, d\tau. \quad \square \end{aligned}$$

Theorem 27 *We have for* $s, \theta_1, \theta_2 > 0$

$$\psi_{(Z,I)|Z(0)=x}(s; \theta_1, \theta_2) = \{e^{-\theta_1 x} I - (\theta_1 + \theta_2)\, e^{-x \Psi^\star(s)} \left[\theta_2 I + \Psi^\star(s)\right]^{-1}\}\, \psi_Y(s; \theta_1) \quad (5.45)$$

where, as given by (5.30), $\psi_Y(s; \theta_1) = \left[sI - D\theta_1 + \Phi(\theta_1)\right]^{-1}$.

Proof: Using Lemma 11 and since (Y, J) is independent of $Z(0)$,

$$\begin{aligned} \psi_{(Z,I)|Z(0)=x}(s; \theta_1, \theta_2) &= \psi_{x+Y}(s; \theta_1) \\ -(\theta_1 + \theta_2)\, \mathcal{L}\left(\mathbb{E}_J \left[\int_0^{(.)} e^{-\theta_1 [Y(.)-Y(\tau)] - \theta_2 I(\tau)} \zeta(\tau)\, d\tau | Z(0) = x\right], s\right). \quad (5.46) \end{aligned}$$

Now, note that

$$\psi_{x+Y}(s;\theta_1) = e^{-\theta_1 x}\psi_Y(s;\theta_1) \tag{5.47}$$

and since $Y(t) - Y(\tau)$ and $(I,\zeta)(\tau)$ are independent as are (Y,J) and $Z(0)$, by interchanging the order of integration in the expression defining the second term in (5.46), we get

$$\left[\mathcal{L}\left(I\!E_J\left[\int_0^{(.)} e^{-\theta_1[Y(.)-Y(\tau)]-\theta_2 I(\tau)}\zeta(\tau)\,d\tau | Z(0)=x\right],s\right)\right]_{jk}$$

$$\begin{aligned}
&= \sum_{l\in E}\left[\mathcal{L}\left(I\!E_J\left[e^{-\theta_2 I(.)}\zeta(.)|Z(0)=x\right],s\right)\right]_{jl} \\
&\qquad \int_\tau^\infty e^{-s(t-\tau)}\,E\left[e^{-\theta_1[Y(t)-Y(\tau)]};J(t)=k|J(\tau)=l\right]\,dt \\
&= \sum_{l\in E}\left(e^{-x\Psi^\star(s)}\left[\theta_2 I+\Psi^\star(s)\right]^{-1}\right)_{jl}\left[\psi_Y(s;\theta_1)\right]_{lk} \\
&= \left[e^{-x\Psi^\star(s)}\left[\theta_2 I+\Psi^\star(s)\right]^{-1}\,\psi_Y(s;\theta_1)\right]_{jk}
\end{aligned} \tag{5.48}$$

by Theorem 26. The statement follows from (5.46), (5.47) and (5.48). □

Corollary 11 *For $s,\theta > 0$, we have:*

$$\psi_{Z|Z(0)=x}(s;\theta) = \{e^{-\theta x}I - \theta\, e^{-x\Psi^\star(s)}\left[\Psi^\star(s)\right]^{-1}\}\,\psi_Y(s;\theta) \tag{5.49}$$

and

$$\psi_{I|Z(0)=x}(s;\theta) = \left\{I - e^{-x\Psi^\star(s)}\left[I+\theta^{-1}\,\Psi^\star(s)\right]^{-1}\right\}\hat{\pi}(s). \tag{5.50}$$

Proof: Equations (5.49) and (5.50) follow by letting $\theta_2 \to 0$ and $\theta_1 \to 0$, respectively, in equation (5.45). □

If at time 0 the storage unit is empty, Theorem 27 and Corollary 11 characterize the process $((Y-m,-m),J)$, where m is the infimum of the net input. Corollary 12 gives this characterization.

Corollary 12 *For $s, \theta_1, \theta_2 > 0$, we have,*

$$\psi_{(Y-m,-m)}(s;\theta_1,\theta_2) = \left\{I - (\theta_1+\theta_2)\left[\theta_2 I + \Psi^\star(s)\right]^{-1}\right\}\psi_Y(s;\theta_1) \tag{5.51}$$

$$\psi_{Y-m}(s;\theta) = \left\{I - \theta\left[\Psi^\star(s)\right]^{-1}\right\}\psi_Y(s;\theta) \tag{5.52}$$

and

$$\psi_{-m}(s;\theta) = \left\{I - \theta\left[\theta I + \Psi^\star(s)\right]^{-1}\right\}\hat{\pi}(s). \ \square \tag{5.53}$$

The unsatisfied demand up to time t, $I(t)$, is zero if and only if the storage facility is nonempty on $[0,t]$. This fact is used to deduce Theorem 28 which characterizes the distribution of $I(t)$ as t goes to infinity.

Theorem 28 *If J is ergodic and $x, y \geq 0$, then*

$$\left(\lim_{t\to\infty} P[I(t)\leq y, J(t)=k \mid Z(0)=x, J(0)=j]\right) = \left[I - e^{-(x+y)\Psi^\star(0)}\right]\Pi. \tag{5.54}$$

Proof: If $j,k \in E$ and $t,x,y\geq 0$, then $P(I(t)\leq y, J(t)=k \mid Z(0)=x, J(0)=j)$ is equal to

$$P(I(t)=0, J(t)=k \mid Z(0)=x+y, J(0)=j).$$

From this, Corollary 11, and a Tauberian theorem, we obtain

$$\left(\lim_{t\to\infty} P[I(t)\leq y, J(t)=k \mid Z(0)=x, J(0)=j]\right) = \lim_{s\to 0+}\left\{s\,\psi_{1_{\{I(.)\leq y\}}|Z(0)=x}(s;\theta)\right\}$$

$$\begin{aligned}
&= \lim_{s\to 0+}\left\{s\psi_{1_{\{I(.)=0\}}|Z(0)=x+y}(s;\theta)\right\} = \lim_{s\to 0+}\left\{s\lim_{\theta\to\infty}\psi_{I|Z(0)=x+y}(s;\theta)\right\}\\
&= \lim_{s\to 0+}\left[s\lim_{\theta\to\infty}\left\{I - e^{-(x+y)\Psi^\star(s)}\left[I+\theta^{-1}\,\Psi^\star(s)\right]^{-1}\right\}\hat{\pi}(s)\right]\\
&= \left[I - e^{-(x+y)\Psi^\star(0)}\right]\lim_{s\to 0+}\{s\,\hat{\pi}(s)\} = \left[I - e^{-(x+y)\Psi^\star(0)}\right]\Pi. \ \square
\end{aligned}$$

5.6 The steady state

In this section we study the storage process in steady state when a *congestion measure* ρ is less than one. We assume the state space E of the MJP J is finite but some of our results may hold for countable E. Note that, since E is finite, the *standard conditions* hold. In addition, if $\tilde{J}$ is irreducible, we may assume without loss of generality that $\tilde{J}$ is aperiodic (by introducing self-transitions in J that bring no input to the storage system).

Let $\bar{d}$ be the *demand rate* and $\bar{\alpha}$ be the *offered load*,

$$\bar{d} = \sum_{k \in E} \pi_k d(k), \qquad \bar{\alpha} = \sum_{k \in E} \pi_k \alpha_k, \tag{5.55}$$

where α_k is the offered load when J is in state k,

$$\alpha_k = \lambda_k \sum_{l \in E} p_{kl}\, \beta_{kl}, \qquad \beta_{kl} = \int_0^\infty x\, M_{kl}\{dx\} \tag{5.56}$$

with β_{kl} being the average value of a jump associated with a transition from state k to state l in J, which we assume finite. We now define the *congestion measure* ρ as

$$\rho = \frac{\bar{\alpha}}{\bar{d}}\,. \tag{5.57}$$

Lemma 12 *If $\rho < 1$, E is finite and $\tilde{J}$ is irreducible (and aperiodic w.l.o.g.), then*

(*a*). $\frac{Y(t)}{t} \to \bar{\alpha} - \bar{d}$ (< 0) *a.s., for all initial distributions.*

(*b*). $\inf\{t > 0 : Z(t) = 0\} < \infty$ *a.s., for all initial distributions.*

(*c*). *For all $j \in E$, $\tilde{T}_j < \infty$ a.s.*

(*d*). $\int_0^t 1_{\{Z(s)=0\}}\, ds \to \infty$ *a.s. as $t \to \infty$, for all initial distributions.*

(*e*). *For all $j, k \in E$, $p_{jk} > 0 \Rightarrow p^\star_{jk} > 0$.*

(f). The DTMC with transition probability matrix $P^\star = \left(p^\star_{jk}\right)$ *is persistent nonnull and we denote its limit distribution by* $\pi^\star$.

(g). The MJP J° *with intensity matrix* $\Lambda P^\star$ *is persistent nonnull and has limit distribution* π°, *with*

$$\pi^\circ_k = \frac{\frac{\pi^\star_k}{\lambda_k}}{\sum_{l\in E}\frac{\pi^\star_l}{\lambda_l}}. \tag{5.58}$$

(h). $(Z,J)(t) \xrightarrow{\mathcal{D}} (Z_\infty, J_\infty)$ *as* $t\to\infty$, *for all initial distributions, where* J_∞ *is the stationary version of* J.

(i). $\frac{I(t)}{t} \to \bar{d} - \bar{\alpha}$ *a.s., for all initial distributions.*

Proof: (a). Since Y can be seen as an accumulated reward process of the ergodic MJP J the result follows imediately from a limit theorem for this type of processes.

(b). Suppose $Z(0) = x \geq 0$, and let $T_x = \inf\{t \geq 0 : Z(t) = 0\}$. We have,

$$Y(s) \leq -x \Rightarrow T_x \leq s. \tag{5.59}$$

The statement follows from (a) and (5.59).

(c). The statement is an imediate consequence of (b).

(d). From (b) and (c), the storage level process returns an infinite number of times to zero a.s.; moreover, since each visit to the zero level has a duration which is stochastically greater or equal to an exponential random variable with mean $[\sup_{j\in E}\lambda_j]^{-1}$, the statement follows.

(e). If $p_{jk} > 0$, we have

$$p^\star_{jk} \geq p_{jk}\int_0^\infty e^{-\lambda_k\frac{x}{d(k)}}\, M_{jk}\{dx\} > 0.$$

(f). From (e) and the fact that $\tilde{J}$ is irreducible and aperiodic, it follows that the DTMC with transition probability matrix $P^{\star}$ is irreducible and aperiodic, and therefore persistent nonnull, since E is finite.

(g). The statement follows directly from (f) and the relation between the stationary distribution of an ergodic MJP and its associated DTMC.

(h). The statement follows from the relation between limit properties of our storage process and the embedded Markov-Lindley process $Z_n = Z(T_{n+1}-)$, as given by Theorem 1 of Van Doorn and Regterschot [41] and the limit properties of this Markov-Lindley process as given in Theorem 4 of Prabhu and Tang [32].

(i). From (h), $\frac{Z(t)}{t} \to 0$, a.s.. From this, (a), and the fact that $Z(t) = Z(0) + Y(t) + I(t)$ the statement follows. □

Remark 12 Since our aim is to study the steady state behaviour of the storage process, and using Lemma 12 (b) and (h), we may without loss of generality assume $Z(0) = 0$ a.s.. From Lemma 12 (d), J° can be identified as J on the time set $\mathcal{T}^{\circ}$,

$$\mathcal{T}^{\circ} = \{t \geq 0 : Z(t) = 0\}, \tag{5.60}$$

and we will assume J° has time set $\mathcal{T}^{\circ}$. J° gives thus the evolution of J when all busy periods are removed from consideration. □

Theorem 29 *If $\rho < 1$, E is finite and $\tilde{J}$ is irreducible (and aperiodic without loss of generality), then*

(a).

$$P(Z_{\infty} = 0) = (1-\rho)\frac{\sum_{k \in E} \pi_k \, d(k)}{\sum_{k \in E} \pi_k^{\circ} \, d(k)}. \tag{5.61}$$

(b). For all initial distributions,

$$\frac{I(t)}{t} \longrightarrow (1-\rho)\sum_{k\in E}\pi_k\, d(k) = (1-\rho)\,\bar{d}, \text{ a.s..} \tag{5.62}$$

(c). With ζ_∞ being the limit distribution of $\zeta(t)$ and $k \in E$, we have

$$P\left(Z_\infty = 0, J_\infty = k\right) = P\left(Z_\infty = 0\right)\pi_k^\circ, \tag{5.63}$$

$$E\left[\zeta_\infty\, 1_{\{J_\infty = k\}}\right] = P\left(Z_\infty = 0\right)\pi_k^\circ\, d(k). \tag{5.64}$$

(d). With (from (c)) $E_J\left[\zeta_\infty\right] = P\left(Z_\infty = 0\right)\Pi^\circ D$, where $\Pi^\circ = \mathbf{1}\pi^\circ$, we have

$$E_J\left[e^{-\theta Z_\infty}\right] = E_J\left[\zeta_\infty\right]\theta\left[D\theta - \Phi(\theta)\right]^{-1}. \tag{5.65}$$

Proof: (a). With $\mathcal{T}^\circ$ defined by (5.60) and $\mathcal{T}_t^\circ = [0,t]\bigcap\mathcal{T}^\circ$, we have, by Remark 12 and Lemma 12 (g),

$$\lim_{t\to\infty}\frac{I(t)}{\int_0^t 1_{\{Z(u)=0\}}\,du} = \lim_{t\to\infty}\frac{\int_{\mathcal{T}_t^\circ} d(J(u))\,du}{\int_{\mathcal{T}_t^\circ} 1\,du}$$

$$= \lim_{s\to\infty}\frac{\int_0^s d(J^\circ(u))\,du}{\int_0^s 1\,du} = \sum_{k\in E}\pi_k^\circ\, d(k), \text{ a.s.} \tag{5.66}$$

From Lemma 12 (h), and for all initial distributions,

$$\frac{1}{t}\int_0^t 1_{\{Z(u)=0\}}\,du \to P\left(Z_\infty = 0\right), \text{ a.s.} \tag{5.67}$$

From Lemma 12 (i), (5.66), (5.67) and the fact that

$$\frac{I(t)}{t} = \frac{I(t)}{\int_0^t 1_{\{Z(u)=0\}}\,du}\times\frac{\int_0^t 1_{\{Z(u)=0\}}\,du}{t} \tag{5.68}$$

we get

$$\bar{d} - \bar{\alpha} = P\left(Z_\infty = 0\right)\sum_{k\in E}\pi_k^\circ\, d(k)$$

from which the result follows.

(*b*). The statement is an imediate consequence of Lemma 12 (*i*).

(*c*). We have,

$$P(Z_\infty = 0, J_\infty = k) = P(Z_\infty = 0)\, P(J_\infty = k \mid Z_\infty = 0) = P(Z_\infty = 0)\, \pi_k^\circ$$

and (5.63) holds; (5.64) follows directly from (5.63).

(*d*). From (5.49) and by a Tauberian theorem, we have

$$\begin{aligned} E_J\left[e^{-\theta Z_\infty}\right] &= \lim_{s \to 0+} \psi_{Z|Z(0)=0}(s;\theta) = \left\{\lim_{s \to 0+} s\,[\Psi^\star(s)]^{-1}\right\}\, \theta\,[D\theta - \Phi(\theta)]^{-1} \\ &= E_J[\zeta_\infty]\, \theta\,[D\theta - \Phi(\theta)]^{-1}, \end{aligned}$$

where the last inequality follows from (5.41) and a Tauberian theorem. □

Remark 13 Note that (5.65) is a generalization of the Pollaczek-Khinchin transform formula for the steady-state waiting time in queue in the $M/G/1$ system that with the notation of Remark 10, $\rho = \lambda/\mu$ ($1/\mu$ being the expected value of the service time distribution), and $W^\star$ being the Laplace transform of the waiting time in queue, is given by:

$$W^\star(\theta) = (1-\rho)\, \frac{\theta}{\theta - \lambda\,[1 - G^\star(\theta)]}\,. \tag{5.69}$$

In particular, (5.65) and (5.69) show that in the Pollaczek-Khinchin formula $(1-\rho)$ stands for the limit demand rejection rate. Moreover, from (5.64), (5.61) and (5.62), the average demand rejection rate is $(1-\rho)\,\bar{d}$ which is (obviously) the limit rate of increase of $I(t)$. As (5.61) shows the limit probability of the storage system being empty may differ susbstantially from $(1-\rho)$. □

Bibliography

[1] D. Anick, D. Mitra, and M. M. Sondhi. Stochastic theory of a data-handling system with multiple sources. *Bell Syst. Tech. J.*, 61:1871–1894, 1982.

[2] S. Asmussen and G. Koole. Marked point processes as limits of markovian arrival streams. *J. Appl. Probab.*, 30:365–372, 1993.

[3] R. M. Blumenthal. *Excursions of Markov Processes.* Birkäuser, 1992.

[4] R. M. Blumenthal and R. Getoor. *Markov Processes and Potential Theory.* Academic Press, 1968.

[5] S. Browne and P. Zipkin. Inventory models with continuous, stochastic demands. *Annals Appl. Probab.*, 1:419–435, 1991.

[6] E. Çinlar. Markov renewal theory. *Adv. Appl. Probab.*, 1:123–187, 1969.

[7] E. Çinlar. Markov additive processes. I. *Z. Wahrscheinlichkeitstheorie verw. Geb.*, 24:85–93, 1972.

[8] E. Çinlar. Markov additive processes. II. *Z. Wahrscheinlichkeitstheorie verw. Geb.*, 24:95–121, 1972.

[9] H. Chen and D. D. Yao. A fluid model for systems with random disruptions. *Opns. Res.*, 40:S239–S247, 1992.

[10] A. I. Elwalid and D. Mitra. Analysis and design of rate-based congestion control of high speed networks, I: Stochastic fluid models, acess regulation. *Queueing Systems*, 9:29–64, 1991.

[11] W. Fischer and K. Meier-Hellstern. The Markov-modulated Poisson process (MMPP) cookbook. *Perform. Evaluation*, 18:149–171, 1993.

[12] A. Friedman. *Foundations of Modern Analysis.* Dover, 1982.

[13] D. P. Gaver and J. P. Lehoczky. Channels that cooperatively service a data stream and voice messages. *IEEE Trans. Commun.*, 30:1153–1162, 1982.

[14] D. P. Gaver and R. Miller, Jr. Limiting distributions for some storage problems. In J. Arrow, S. Karlin, and H. Scarf, editors, *Studies in Applied Probability and Management Science*, pages 110–126. Stanford University Press, 1962.

[15] J. A. Goldstein. *Semigroups of Linear Operators and Applications.* Oxford University Press, 1985.

[16] J. M. Harrison. *Brownian Motion and Stochastic Flow Systems.* Wiley, 1985.

[17] I. Karatzas and S. E. Shreve. *Brownian Motion and Stochastic Calculus.* Springer-Verlag, 1988.

[18] O. Kella and W. Whitt. A stochastic storage system in a random environment. *Opns. Res.*, 40:S257–S262, 1992.

[19] J. F. C. Kingman. *Poisson Processes.* Oxford University Press Inc., 1993.

[20] D. M. Lucantoni. The BMAP/G/1 queue: A tutorial. In *Models and Techniques for Performance Evaluation of Computer and Communication Systems*, pages 330–358. Springer Verlag, 1993.

[21] D. M. Lucantoni, K. S. Meier-Hellstern, and M. F. Neuts. A single-server queue with server vacations and a class of non-renewal arrival processes. *Adv. Appl. Probab.*, 22:676–705, 1990.

[22] F. Machihara. Completion time of service unit interrupted by PH-Markov renewal customers and its applications. In *Telletraffic Science for New Cost-Effective Systems, Networks and Services* (Proceedings of the 12th International Teletraffic Congress (ITC-12), Torino, Italy (1988)), pages 1508–1514. North-Holland, 1989.

[23] D. Mitra. Stochastic theory of a fluid model of producers and consumers coupled by a buffer. *Adv. Appl. Probab.*, 20:646–676, 1988.

[24] S. Narayana and M. F. Neuts. The first two moment matrices of the counts for the markovian arrival process. *Stochastic Models*, 8:459–477, 1992.

[25] M. F. Neuts. *Structered Stochastic Matrices of $M/G/1$ Type and their Applications.* Marcel Dekker, 1989.

[26] M. F. Neuts. Models based on the markovian arrival process. *IEICE Trans. Commun.*, E75-B:1255–1265, 1992.

[27] J. Neveu. Une généralisation des processus à accroissements positifs indépendants. *Abh. math. Sem. Univ. Hamburg*, 25:36–61, 1961.

[28] C. Olivier and J. Walrand. On the existence of finite dimensional filters for Markov modulated traffic. *J. Appl. Probab.*, 1994.

[29] N. U. Prabhu. *Queues and Inventories.* Wiley, 1965.

[30] N. U. Prabhu. *Stochastic Storage Processes.* Springer-Verlag, 1980.

[31] N. U. Prabhu. Markov renewal and markov-additive processes – a review and some new results. In B. D. Choi and J. Yim, editors, *Proceedings of KAIST Mathematics Workshop,* Volume 6, pages 57–94. Korea Advanced Institute of Science and Technology, 1991.

[32] N. U. Prabhu and L. C. Tang. Markov-modulated single server queueing systems. To appear in *J. Appl. Probab.*, 1994.

[33] N. U. Prabhu, L. C. Tang, and Y. Zhu. Some new results for the Markov random walk. *J. Math. Phys. Sci.*, 25:635–663, 1991.

[34] N. U. Prabhu and Y. Zhu. Markov-modulated queueing systems. *Queueing Systems*, 5:215–245, 1989.

[35] E. L. Presman. Factorization methods and boundary problems for sums of random variables given on Markov chains. *Matn. USSR Izvestija.*, 3:815–852, 1969.

[36] M. Rudemo. Point processes generated by transitions of Markov chains. *Adv. Appl. Probab.*, 5:262–286, 1973.

[37] R. F. Serfozo. Queueing networks with dependent nodes and concurrent movements. *Queueing Systems*, 13:143–182, 1993.

[38] B. A. Sevast'yanov. Influence of a storage bin capacity on the average standstill time of a production line. *Theory Prob. Appl.*, 7:429–438, 1962.

[39] T. Takine, B. Sengupta, and T. Hasegawa. A conformance measure for traffic shapping in high-speed networks with an application to the leaky bucket. To appear.

[40] E. A. van Doorn, A. A. Jagers, and J. S. J. de Wit. A fluid reservoir regulated by a birth-death process. *Stochastic Models*, 4:457–472, 1988.

[41] E. A. van Doorn and G. J. K. Regterschot. Conditional pasta. *Opns. Res. Lett.*, 7:229–232, 1988.

[42] J. Virtamo and I. Norros. Fluid queue driven by an $M/M/1$ queue. To appear, 1993.

[43] H. Yamada and F. Machihara. Performance analysis of a statistical multiplexer with control on input and/or service processes. *Perform. Evaluation*, 14:21–41, 1992.