

Operator–Geometric Solutions for the M/G/k Queue and its Variants

Lothar Breuer*
University of Trier

8th May 2002

One of the classical, but yet unsolved queueing systems is the M/G/k queue with Poisson input, general service time distribution and k serving facilities. In the present paper, this queue is analyzed as a piecewise–deterministic Markov process. First, an iteration formula for the transient distributions is derived. This formula is a new result for piecewise–deterministic Markov processes, too, since it yields their transition probability kernel. In the main part, recent results on stability of piecewise–deterministic Markov processes are used in order to show that the computation of the stationary distribution for the M/G/k queue can be reduced to the determination of the stationary distribution for a birth and death process with kernel entries. This results in operator–geometric solutions for the M/G/k queue. For the M/G/k retrial queue, a similar analysis is presented. Here, closed formulae for the transient distribution are derived. These expressions hold for the class of piecewise–deterministic processes without jumps from a border set, too. Additionally, it is sketched how these results can be generalized for BMAP/G/k queues.

1 Introduction

A very old method of reducing non–Markovian queueing systems to Markov processes is the method of supplementary variables introduced by Cox [8]. At the time of its introduction, this method lacked mathematical support, since it transformed the queueing process to a Markov process with continuous state space, and such processes have not been analyzed exhaustively until the 1970s. Now, after good results have been obtained for general Markov processes, the method of supplementary variables can be taken up again. This has been done first by Kovalenko in 1964 under the concept of piecewise–linear Markov processes (see Gnedenko, Kovalenko [12]). In 1984, Davis [9] introduced the more general class of piecewise–deterministic

*FB IV - Informatik, Universität Trier, 54286 Trier, Germany. email: breuer@info04.uni-trier.de

Markov processes, initiating subsequent research as listed in the bibliography of Davis [10]. Costa and Dufour [7, 11] contributed important results concerning stationary distributions of piecewise–deterministic Markov processes, which will be used in this paper extensively.

In the next section, basic notations are introduced and the M/G/k queue is described as a piecewise–deterministic Markov process. Then a derivation of the transient distributions for the M/G/k queue is given. The iteration formula is a new result valid for all piecewise–deterministic Markov processes. In subsection 2.2, the computation of stationary distribution for the M/G/k queue is reduced to the determination of the stationary distribution for a birth and death process with kernel entries. This yields operator–geometric solutions for the M/G/k queue for which existing results on operator–analytic methods can be used (e.g. Tweedie [16] or Nielsen and Ramaswami [15]). In section 3, a similar analysis is presented for the M/G/k retrial queue. For the transient distribution, a closed formula will be derived, which also holds for a large class of piecewise–deterministic Markov processes. Then it is shown how to determine the stationary distribution via a generalization of the methods in Hofmann [13]. For both kinds of queues, the models can be enhanced by inclusion of BMAP arrivals, for which the same method of analysis applies.

2 The M/G/k Queue

Consider an M/G/k queue with the following characteristics. The Poisson input shall have rate λ . The service time distribution shall be denoted by B , being equal for each of the k servers. It will be apparent that the convention of equal service time distributions is not necessary for the method of analysis. However, it simplifies notations and thus shall be adopted.

This queue can be described as a piecewise–deterministic Markov process in the following way. Let X be a piecewise–deterministic Markov process with state space $E := \mathbb{N}_0 \times (\mathbb{R}_0^+)^k$. A flow Φ on E shall be defined by

$$\Phi_t(n, x) := (n, (x_1 - t)^+, \dots, (x_k - t)^+) \quad (1)$$

for all $(n, x) = (n, x_1, \dots, x_k) \in E$ and $t \in \mathbb{R}_0^+$, with $(s - t)^+ := \max(0, s - t)$ for all $s, t \in \mathbb{R}$. Obviously, for the above $(n, x) = (n, x_1, \dots, x_k)$ the first component n represent the number of users waiting in the queue and the components x_i represent the remaining service time at the i th server. If the i th server is idle, then $x_i = 0$. According to Costa, Dufour [11], we define

$$t_*(x) := \begin{cases} \min\{x_i : 1 \leq i \leq k, x_i > 0\} & \text{for } x \neq 0 \\ \infty & \text{for } x = 0 \end{cases} \quad (2)$$

for all $x = (x_1, \dots, x_k) \in (\mathbb{R}_0^+)^k$. This denotes the time until the first server will become idle.

Differing from Davis [9] and Costa, Dufour [11], we will introduce two transition measures Q_1 and Q_2 for the jumps that can occur. This reflects the queueing process more transparently. Q_1 is the transition measure for arrivals, and thus we define for all $(n, x) =$

$(n, x_1, \dots, x_k) \in E$ and $A = A_1 \times \dots \times A_k$

$$Q_1((n, x), \{m\} \times A) := \begin{cases} \delta_{m, n+1} \cdot 1_A(x) & \text{for } \prod_{i=1}^k x_i > 0 \\ \delta_{m, n} \cdot \prod_{j=1, j \neq i}^k 1_{A_j}(x_j) \cdot B(A_i) & \text{for } i = \min\{l : x_l = 0\} \end{cases}$$

Note that the latter case in the definition of Q_1 is possible only for $n = m = 0$. The second transition measure Q_2 refers to the case of a server becoming idle. If there are any waiting users in the queue, it immediately will commence to serve a new user. Thus we have

$$Q_2((n, x), \{m\} \times A) := \begin{cases} \delta_{m, n-1} \cdot \prod_{j=1, j \neq i}^k 1_{A_j}(x_j) \cdot B(A_i) & \text{for } n \geq 1, x_i = 0 \\ \delta_{m, n} \cdot 1_A(x) & \text{for } n = 0 \end{cases}$$

Note that for the case $n \geq 1$, only one server can be idle at a time. Since the queue has Poisson single arrival input, the probability that two servers finish their work at the same time instant is zero. Furthermore, if one server had been idle before the other server and there had been any waiting users in the queue, it would have commenced serving one of them.

2.1 Transient Distributions

Denote $1_k \in \mathbb{R}^k$ as the vector with all entries being one and write $(x - y)^+ := ((x_1 - y_1)^+, \dots, (x_k - y_k)^+)$ for all $x, y \in \mathbb{R}^k$. Let $P(t; (n, x), \{m\} \times A)$ denote the probability that at time $t \in \mathbb{R}_0^+$, the queue process X is in the state set $\{m\} \times A$ under the condition that it was in state (n, x) at time 0. Further, let $P^{(l)}(t; (n, x), \{m\} \times A)$ denote the same probability, but restricted to the set of paths with $l \in \mathbb{N}_0$ jumps until time t . Then the transition probability kernel and hence the transient distribution of the queue process X is given iteratively by

$$P(t; (n, x), \{m\} \times A) = \sum_{l=0}^{\infty} P^{(l)}(t; (n, x), \{m\} \times A)$$

with

$$P^{(0)}(t; (n, x), \{m\} \times A) = \begin{cases} \delta_{m, n} \cdot e^{-\lambda t} \cdot 1_A(x - t \cdot 1_k) & \text{for } t < t_*(x), n \geq 1 \\ \delta_{m, 0} \cdot e^{-\lambda t} \cdot 1_A((x - t \cdot 1_k)^+) & \text{for } n = 0 \\ 0 & \text{else} \end{cases}$$

as start values and iterating by

$$P^{(l+1)}(t; (n, x), \{m\} \times A) = \int_0^t e^{-\lambda u} \lambda \int_E P^{(l)}(t - u; y, \{m\} \times A) dQ_1((n, x), y) du$$

if $t < t_*(x)$ or

$$P^{(l+1)}(t; (n, x), \{m\} \times A) = \int_0^{t_*(x)} e^{-\lambda u} \lambda \int_E P^{(l)}(t - u; y, \{m\} \times A) dQ_1((n, x), y) du \\ + e^{-\lambda t_*(x)} \int_E P^{(l)}(t - t_*(x); y, \{m\} \times A) dQ_2((n, x), y)$$

if $t > t_*(x)$.

A more efficient way to compute the transient distribution is the following: Define the operator f by $f(P^{(l)}) := P^{(l+1)}$. Then the transition probability kernel P can be computed as the limit

$$P = \lim_{n \rightarrow \infty} P_n$$

with

$$P_0 = P^{(0)} \quad \text{and} \quad P_{n+1} = f(P_n) + P_0 \quad (3)$$

Remark 1 The above formulae for the transient distribution can be stated for general piecewise-deterministic Markov processes as well. Using the notations as in Costa, Dufour [11], which are simpler than the ones in Davis [9], we have

$$P^{(0)}(t; x, A) = e^{-\Lambda(x,t)} 1_A(\Phi_t(x))$$

as start values and the iteration

$$P^{(l+1)}(t; x, A) = \int_0^t e^{-\Lambda(x,u)} \lambda(\Phi_u(x)) \int_E P^{(l)}(t-u; y, A) Q(\Phi_u(x), dy) du$$

if $t < t_*(x)$ or

$$\begin{aligned} P^{(l+1)}(t; x, A) &= \int_0^{t_*(x)} e^{-\Lambda(x,u)} \lambda(\Phi_u(x)) \int_E P^{(l)}(t-u; y, A) Q(\Phi_u(x), dy) du \\ &\quad + e^{-\Lambda(x,t_*(x))} \int_E P^{(l)}(t-t_*(x); y, A) Q(\Phi_{t_*(x)}(x), dy) \end{aligned}$$

if $t > t_*(x)$. Here, $\lambda(x)$ denotes the jump rate at $x \in E$, Q is the jump transition measure and $\Lambda(x, u) = \int_0^u \lambda(\Phi_v(x)) dv$.

Furthermore, the iteration as given in (3) is valid as well in the general setting of piecewise-deterministic Markov processes.

2.2 Stability and Stationary Distribution

Results concerning stability and the stationary distribution can be obtained by applying the method of Costa, Dufour [11] to the present process. In this paper, a crucial result by Azema et al. [1] is used, which states that the stationary distribution of a Markov process equals the invariant measure of its resolvent. In terms of the notation in Costa, Dufour [11], we have $\Lambda(x, t) = \lambda \cdot t$ and define the kernels

$$L((n, x), \{m\} \times A) = \delta_{m,n} \cdot \int_0^{t_*(x)} e^{-(s+\lambda s)} 1_A(x - s \cdot 1_k) ds$$

as well as

$$\begin{aligned} K((n, x), \{m\} \times A) &= \int_0^{t_*(x)} e^{-(s+\lambda s)} \lambda Q_1((n, x - s \cdot 1_k), \{m\} \times A) ds \\ &\quad + e^{-(t_*(x)+\lambda t_*(x))} Q_2((n, x - t_*(x) \cdot 1_k), \{m\} \times A) \end{aligned}$$

This means that for $A = A_1 \times \dots \times A_k$,

$$\begin{aligned} K((0, x), \{0\} \times A) &= e^{-(t_*(x) + \lambda t_*(x))} 1_A(x - t_*(x) \cdot 1_k) \\ K((0, x), \{1\} \times A) &= \lambda \int_0^{t_*(x)} e^{-(s + \lambda s)} 1_A(x - s \cdot 1_k) ds \end{aligned}$$

for $\prod_{i=1}^k x_i > 0$,

$$\begin{aligned} K((0, x), \{0\} \times A) &= \lambda \cdot B(A_i) \cdot \int_0^{t_*(x)} e^{-(s + \lambda s)} \cdot \prod_{j=1, j \neq i}^k 1_{A_j}(x_j - s) ds \\ &\quad + e^{-(t_*(x) + \lambda t_*(x))} 1_A(x - t_*(x) \cdot 1_k) \end{aligned}$$

for $i = \min\{l : x_l = 0\} > 1$ and

$$K((0, 0), \{0\} \times A) = \lambda \cdot B(A_i) \cdot \int_0^\infty e^{-(s + \lambda s)} \cdot \prod_{j=1, j \neq i}^k 1_{A_j}(x_j - s) ds$$

while for $n \in \mathbb{N}$ and $i = \arg \min\{x_1, \dots, x_k\}$,

$$\begin{aligned} K((n, x), \{n + 1\} \times A) &= \lambda \int_0^{t_*(x)} e^{-(s + \lambda s)} 1_A(x - s \cdot 1_k) ds \\ K((n, x), \{n - 1\} \times A) &= e^{-(t_*(x) + \lambda t_*(x))} \prod_{j=1, j \neq i}^k 1_{A_j}(x_j - t_*(x)) \cdot B(A_i) \end{aligned}$$

The kernels L and K can be regarded as the resolvents before the first jump of X and immediately after the first jump of X , respectively.

The main result in Costa, Dufour [11] is that the existence of a stationary distribution for the queue process X is equivalent to the existence of an invariant positive σ -finite measure for the kernel $G := K + L$ (see theorem 3.5 and corollary 3.6 in [11]). This kernel G can be arranged by its first component in an $\mathbb{N}_0 \times \mathbb{N}_0$ -matrix with kernel entries as

$$G(x, A) = \begin{pmatrix} G_{00}(x, A) & G_{01}(x, A) & 0 & 0 & 0 & \dots \\ G_{10}(x, A) & G_{11}(x, A) & G_{01}(x, A) & 0 & 0 & \dots \\ 0 & G_{10}(x, A) & G_{11}(x, A) & G_{01}(x, A) & 0 & \dots \\ 0 & 0 & G_{10}(x, A) & G_{11}(x, A) & G_{01}(x, A) & \ddots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots \end{pmatrix}$$

and the kernels

$$G_{00}(x, A) = \int_0^{t_*(x)} e^{-(s + \lambda s)} 1_A(x - s \cdot 1_k) ds + e^{-(t_*(x) + \lambda t_*(x))} 1_A(x - t_*(x) \cdot 1_k)$$

for $\prod_{i=1}^k x_i > 0$,

$$\begin{aligned} G_{00}(x, A) &= \int_0^{t_*(x)} e^{-(s+\lambda s)} 1_A((x - s \cdot 1_k)^+) ds \\ &\quad + \lambda \cdot B(A_i) \cdot \int_0^{t_*(x)} e^{-(s+\lambda s)} \cdot \prod_{j=1, j \neq i}^k 1_{A_j}((x_j - s)^+) ds \\ &\quad + e^{-(t_*(x)+\lambda t_*(x))} 1_A((x - t_*(x) \cdot 1_k)^+) \end{aligned}$$

for $i = \min\{l : x_l = 0\} > 1$ and

$$G_{00}(0, A) = \int_0^\infty e^{-(s+\lambda s)} 1_A((x - s \cdot 1_k)^+) ds + \lambda \cdot B(A_i) \cdot \int_0^\infty e^{-(s+\lambda s)} \cdot \prod_{j=1, j \neq i}^k 1_{A_j}(0) ds$$

as well as

$$\begin{aligned} G_{10}(x, A) &= e^{-(t_*(x)+\lambda t_*(x))} \prod_{j=1, j \neq i}^k 1_{A_j}(x_j - t_*(x)) \cdot B(A_i) \\ G_{11}(x, A) &= \int_0^{t_*(x)} e^{-(s+\lambda s)} 1_A(x - s \cdot 1_k) ds \\ G_{01}(x, A) &= \lambda \cdot G_{11}(x, A) \end{aligned}$$

for $\prod_{i=1}^k x_i > 0$, $i = \arg \min\{x_1, \dots, x_k\}$ and $A = A_1 \times \dots \times A_k$. Thus G has the form of a birth and death transition matrix with kernel entries.

Since the system $\{[0, y_1[\times \dots \times [0, y_k[: y_1, \dots, y_k \in \mathbb{R}_0^+\}$ generates $\sigma((\mathbb{R}_0^+)^k)$, the above kernels are determined by the values for $A = [0, y_1[\times \dots \times [0, y_k[$ for $y_1, \dots, y_k > 0$. First we have for $x = 0$

$$G_{00}(0, A) = \frac{1 + \lambda B([0, y_i])}{1 + \lambda}$$

Define $M(x - y)^+ := \max\{0, (x_1 - y_1), \dots, (x_k - y_k)\}$ for all $x, y \in (\mathbb{R}_0^+)^k$. Further note that the condition $t_*(x) > M(x - y)^+$ is equivalent to $y_n > x_n - t_*(x)$ for all $n \in \{1, \dots, k\}$. Hence for $t_*(x) \leq M(x - y)^+$ we get $G_{00}(x, A) = G_{10}(x, A) = G_{11}(x, A) = G_{01}(x, A)$ for the rest of the cases. Thus we assume $t_*(x) > M(x - y)^+$ and obtain

$$G_{00}(x, A) = \frac{1}{1 + \lambda} \left(e^{-M(x-y)^+(1+\lambda)} + \lambda e^{-t_*(x)(1+\lambda)} \right)$$

for $\prod_{i=1}^k x_i > 0$, and

$$G_{00}(x, A) = \frac{1}{1 + \lambda} \left((1 + \lambda B([0, y_i]) e^{-M(x-y)^+(1+\lambda)} + \lambda(1 - B([0, y_i]) e^{-t_*(x)(1+\lambda)}) \right)$$

for $i = \min\{l : x_l = 0\} > 1$, as well as

$$\begin{aligned} G_{10}(x, A) &= e^{-t_*(x)(1+\lambda)} \cdot B([0, y_i]) \\ G_{11}(x, A) &= \frac{1}{1+\lambda} \left(e^{-M(x-y)^+(1+\lambda)} - e^{-t_*(x)(1+\lambda)} \right) \\ G_{01}(x, A) &= \frac{\lambda}{1+\lambda} \left(e^{-M(x-y)^+(1+\lambda)} - e^{-t_*(x)(1+\lambda)} \right) \end{aligned}$$

for $\prod_{i=1}^k x_i > 0$, denoting $i = \arg \min\{x_1, \dots, x_k\}$. For $\prod_{i=1}^k x_i = 0$, we know that $G_{01}(x, A) = 0$ and that (n, x) is an inaccessible state for $n \in \mathbb{N}$.

Now in order to determine the stationary distribution of X , it suffices to obtain the stationary distribution for a quasi birth and death (QBD) process with a continuous phase variable. Denote this Markov process with transition kernel G by $Y = (Y_n : n \in \mathbb{N})$. This kind of process has been examined by Tweedie [16]. In this terminology, we have $B_0 = G_{00}$, $A_0 = G_{01}$, $B_1 = A_2 = G_{10}$, $A_1 = G_{11}$ and $A_i = B_j = 0$ for all other indices $i, j \in \mathbb{N}_0$. Further, define $A := A_0 + A_1 + A_2$ and

$$\beta(w) := A_1(w, (\mathbb{R}_0^+)^k) + 2 \cdot A_2(w, (\mathbb{R}_0^+)^k)$$

In Tweedie [16] it is shown that, under some technical assumptions, a sufficient condition for the QBD process to have a stationary distribution is that A has an invariant probability measure ν and that the drift condition

$$\int_{(\mathbb{R}_0^+)^k} \beta(w) \nu(dw) > 1 \quad (4)$$

is satisfied. Then the stationary distribution Π of the QBD is given by

$$\Pi(k, A) = c \int_{(\mathbb{R}_0^+)^k} {}_0\Pi(dy) S^k(y, A)$$

for all $k \in \mathbb{N}_0$ and $A \in \sigma((\mathbb{R}_0^+)^k)$, with $c = \Pi(0, (\mathbb{R}_0^+)^k) > 0$ and S^k denoting the k th iteration of the kernel S . Here, ${}_0\Pi$ denotes the stationary distribution of the imbedded Markov chain ${}_0Y$ at level zero with transition kernel

$$B(S)(x, A) = B_0(x, A) + \int_{(\mathbb{R}_0^+)^k} S(x, dy) B_1(y, A)$$

for all $x \in (\mathbb{R}_0^+)^k$ and $A \in \sigma((\mathbb{R}_0^+)^k)$. The kernel S on $(\mathbb{R}_0^+)^k$ is the minimal solution to the equation

$$S(x, A) = A_0(x, A) + \int_{(\mathbb{R}_0^+)^k} S(x, dy) A_1(y, A) + \int_{(\mathbb{R}_0^+)^k} S^2(x, dy) A_2(y, A)$$

for all $x \in (\mathbb{R}_0^+)^k$ and $A \in \sigma((\mathbb{R}_0^+)^k)$. Furthermore, c can be determined as the normalizing constant, i.e.

$$c = \left(\sum_{k=0}^{\infty} \int_{(\mathbb{R}_0^+)^k} {}_0\Pi(dy) S^k(y, (\mathbb{R}_0^+)^k) \right)^{-1}$$

Hence under condition (4), the QBD process Y has an operator–geometric solution Π . Questions of numerical computation of such solutions have been addressed by Nielsen and Ramaswami [15]. The probably most feasible approach to developing a numerically tractable calculus of such kernels would be a wavelet representation. Also applicable to the present case are the algorithms presented by Baum [3, 2] since they use only algebraic and probabilistic arguments.

Assume now that condition (4) is satisfied and denote the stationary distribution of the QBD process Y by Π . Then remarks 3.8 and 3.1 along with theorem 3.5 in Costa, Dufour [11] imply that $\mu = \Pi L$ is a stationary distribution for X . With the above notations, the structure of μ is

$$\mu(k, A) = c \int_{(\mathbb{R}_0^+)^k} \int_{(\mathbb{R}_0^+)^k} {}_0\Pi(dy) S^k(y, dz) L(z, A) \quad (5)$$

for all $k \in \mathbb{N}_0$ and $A \in \sigma((\mathbb{R}_0^+)^k)$. Thus the queue process X has an operator–geometric solution.

Remark 2 Analogously to the method of Tweedie [16], one can generalize known results for $M/G/1$ –type matrices towards such matrices with kernel entries. This yields the possibility of including *BMAP* arrivals (with m phases) into the model. With such arrivals, the above kernel G would assume $M/G/1$ –type form (of dimension $(\mathbb{N}_0 \times \{1, \dots, m\})^2$) with kernel entries if ordered according to the number of users in the system and the phase of the arrival process.

3 The M/G/k Retrial Queue

The same approach as above applies to the $M/G/k$ retrial queue. In terms of a representation as a piecewise–deterministic Markov process, this queue is simpler as it does not involve any jumps from a set of border states. In terms of the resulting matrices to be solved, it is more complicated since these matrices now are level–dependent.

In this section, the respective formulae for the transient and stationary distributions shall be given for the following system: Consider the $M/G/k$ retrial queue with Poissonian input of rate λ and service time distribution B , equal for each of the k servers. Again, the assumption of equal service time distributions is not necessary for the present method of analysis, but it simplifies notations. A user who finds the system busy shall persistently (i.e. until successfully entering the system) retry at time instants given by a Poisson process with rate μ . All users in the orbit, i.e. all users which have found the system busy, are assumed to behave independently.

This queue can be modelled as a piecewise–deterministic Markov process with state space $E := \mathbb{N}_0 \times (\mathbb{R}_0^+)^k$. Define the same flow Φ on E as given by (1). A state $(n, x) = (n, x_1, \dots, x_k)$ will be given the same interpretation as for the ordinary $M/G/k$ queue without retrials, the only difference being that n shall denote the number of users in the orbit. As in definition (2), denote by $t_*(x)$ the time until the first server becomes empty.

Unlike for the ordinary $M/G/k$ queue, we have only jumps induced by external Poisson events. This is due to the fact that a server becoming idle does not lead immediately to the

admission of a waiting user into the system, but such an admission can occur only due to the Poissonian arrival or retrial streams. In terms of the conventions in Costa, Dufour [11], this means that at the boundary

$$\partial E^0 = \{(n, x_1, \dots, x_k) \mid \exists 1 \leq i \leq k : x_i = 0\}$$

no jumps occur. Thus the M/G/k retrial queue can be modeled as a convenient special case of piecewise-deterministic Markov processes.

3.1 Transient Distribution

In this section, we want to derive an expression of the transient distribution in a form that is similar to the expression for the transition kernel of a Markov jump process as given in Breuer [5], p.70, or more explicitly in Breuer [4], p.92f.

Define the kernel

$$Q((n, x), \{m\} \times A) = \lambda(Q_1((n, x), \{m\} \times A) - \delta_{m,n}1_A(x)) \\ + n \cdot \mu(Q_2((n, x), \{m\} \times A) - \delta_{m,n}1_A(x))$$

along with the transition measures

$$Q_1((n, x), \{m\} \times A) = \begin{cases} \delta_{m,n} \prod_{j=1, j \neq i}^k 1_{A_j}(x_j) B(A_i) & \text{for } i = \min\{1 \leq j \leq k : x_j = 0\} \\ \delta_{m,n+1} 1_A(x) & \text{if } x_j > 0 \text{ for all } j \end{cases}$$

for arrival events, and

$$Q_2((n, x), \{m\} \times A) = \begin{cases} \delta_{m,n-1} \prod_{j=1, j \neq i}^k 1_{A_j}(x_j) B(A_i) & \text{for } i = \min\{1 \leq j \leq k : x_j = 0\} \\ \delta_{m,n} 1_A(x) & \text{if } x_j > 0 \text{ for all } j \end{cases}$$

for retrial events.

Similar to the kernel G in the previous section, we can write the kernel Q as a block matrix

$$Q(x, A) = \begin{pmatrix} Q_{00}(x, A) & Q_{01}(x, A) & 0 & 0 & 0 & \dots \\ Q_{10}(x, A) & Q_{11}(x, A) & Q_{12}(x, A) & 0 & 0 & \dots \\ 0 & Q_{21}(x, A) & Q_{22}(x, A) & Q_{23}(x, A) & 0 & \dots \\ 0 & 0 & Q_{32}(x, A) & Q_{33}(x, A) & Q_{34}(x, A) & \ddots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots \end{pmatrix}$$

with kernel entries

$$Q_{n,n}(x, A) = \lambda Q_1((n, x), \{n\} \times A) + n \cdot \mu Q_2((n, x), \{n\} \times A) - (\lambda + n \cdot \mu) 1_A(x) \\ = \lambda \prod_{j=1, j \neq i}^k 1_{A_j}(x_j) (B(A_i) - 1_{A_i}(x_i)) - n \cdot \mu 1_A(x)$$

for $i = \min\{1 \leq j \leq k : x_j = 0\}$,

$$Q_{n,n}(x, A) = -\lambda 1_A(x)$$

if $x_j > 0$ for all j , as well as

$$Q_{n,n+1}(x, A) = \lambda Q_1((n, x), \{n+1\} \times A) = \lambda 1_A(x)$$

if $x_j > 0$ for all j , and

$$Q_{n,n-1}(x, A) = n \cdot \mu Q_2((n, x), \{n-1\} \times A) = n \cdot \mu \prod_{j=1, j \neq i}^k 1_{A_j}(x_j) B(A_i)$$

for $i = \min\{1 \leq j \leq k : x_j = 0\}$. Note that this kernel Q is level-dependent, unlike the kernel G in the previous section.

Remark 3 The retrial queue is perhaps the most classical example for level-dependent queueing systems. Thus, in this section it is analyzed in order to exemplify the method of analysis for level-dependent systems. The same method can be applied to other types of level-dependent systems, e.g. finite capacity queues (see Hofmann [13], p.84ff).

Furthermore, this method of analysis also holds for BMAP arrivals (with m phases). Then the kernel Q can be written as a $(\mathbb{N}_0 \times \{1, \dots, m\})^2$ -block matrix with kernel entries if it is ordered according to the number of users in the system and the current phase of the arrival stream. The following arguments hold for this kind of kernel, too.

The kernel Q is the infinitesimal generator which characterizes the jump part of the queueing process. Q can be decomposed into a rate part R and a conditional jump part J as follows: Define R to be the negative diagonal of Q and $J := R^{-1}Q + I$. These definitions are completely analogous to the construction of a discrete-time Markov chain from a continuous-time Markov chain (with discrete state spaces), as described e.g. in Breuer, Dudin, Klimenok [6]. Then we have the relation

$$Q = -R + RJ$$

with a rate kernel defined by

$$R_{nn}(x, A) = \begin{cases} -(\lambda + n \cdot \mu) 1_A(x) & \text{for } i = \min\{1 \leq j \leq k : x_j = 0\} \\ -\lambda 1_A(x) & \text{if } x_j > 0 \text{ for all } j \end{cases}$$

for all $n \in \mathbb{N}_0$, $x \in (\mathbb{R}_0^+)^k$ and $A \in \sigma((\mathbb{R}_0^+)^k)$. From this definition it is obvious that the diagonal kernel R has an inverse R^{-1} .

Define for a kernel $K : E \times \sigma(E) \rightarrow \mathbb{R}$ and a function $f : E \rightarrow E$ the operation $K \circ f(x, A) := K(f(x), A)$. Further denote the identity kernel on E by Id . By remark 1, we then have the first statement

Theorem 1 *The transient distribution kernel of the queueing process of the M/G/k retrial queue is given by*

$$P(t) = \sum_{n=0}^{\infty} \int_0^t \int_0^{t_n} \dots \int_0^{t_2} e^{-\int_0^{t_1} R \circ \Phi_u du} (R \circ \Phi_{t_1}) J e^{-\int_{t_1}^{t_2} R \circ \Phi_u du} (R \circ \Phi_{t_2-t_1}) J \dots \\ \dots e^{-\int_{t_n}^{t_{n-1}} R \circ \Phi_u du} (R \circ \Phi_{t_n-t_{n-1}}) J e^{-\int_{t_n}^t R \circ \Phi_u du} (Id \circ \Phi_{t-t_n}) dt_1 \dots dt_n$$

the sum entry for $n = 0$ being the kernel $e^{-\int_0^t R \circ \Phi_u du} (Id \circ \Phi_t)$.

Proof: For the special case of piecewise-deterministic Markov processes that is applied here, the iteration in remark 1 always uses the first type of iteration step (for $t < t_*(x)$). By induction on n it is shown that the n -th sum entry in the above theorem 1 equals $P^{(n)}$ in remark 1.

☺

Theorem 2 *The transient distribution kernel of the queueing process can also be written as*

$$P(t) = \sum_{n=0}^{\infty} \int_0^t \int_0^{t_n} \dots \int_0^{t_2} (Q \circ \Phi_{t_1})(Q \circ \Phi_{t_2-t_1}) \dots (Q \circ \Phi_{t_n-t_{n-1}})(Id \circ \Phi_{t-t_n}) dt_1 \dots dt_n$$

the sum entry for $n = 0$ being the kernel $Id \circ \Phi_t$.

Proof: The proof of this theorem requires two lemmata, which are given after this proof. For a kernel K and a time duration t , abbreviate in this proof $K_t := K \circ \Phi_t$. We now have

$$\sum_{n=0}^{\infty} \int_0^t \int_0^{t_n} \dots \int_0^{t_2} (Q \circ \Phi_{t_1})(Q \circ \Phi_{t_2-t_1}) \dots (Q \circ \Phi_{t_n-t_{n-1}})(Id \circ \Phi_{t-t_n}) dt_1 \dots dt_n \\ = \sum_{n=0}^{\infty} \int_0^t \int_0^{t_n} \dots \int_0^{t_2} (-R_{t_1} + (RJ)_{t_1}) \dots (-R_{t_n} + (RJ)_{t_n}) Id_{t-t_n} dt_1 \dots dt_n$$

The integrand of every entry of this sum can be written as

$$(-R_{t_1} + (RJ)_{t_1}) \dots (-R_{t_n} + (RJ)_{t_n}) Id_{t-t_n} = \sum_{l \in \{0,1\}^n} K_{l_1} \dots K_{l_n} Id_{t-t_n}$$

with

$$K_{l_j} = \begin{cases} -R_{t_j} & \text{for } l_j = 0 \\ (RJ)_{t_j} & \text{for } l_j = 1 \end{cases}$$

for all $j = 1, \dots, n$. Summing up over the number $m(l) := |\{j : l_j = 1\}|$ of jump events (RJ) and rewriting the integration limits, we obtain

$$\sum_{n=0}^{\infty} \int_0^t \int_0^{t_n} \dots \int_0^{t_2} (Q \circ \Phi_{t_1})(Q \circ \Phi_{t_2-t_1}) \dots (Q \circ \Phi_{t_n-t_{n-1}})(Id \circ \Phi_{t-t_n}) dt_1 \dots dt_n \\ = \sum_{m=0}^{\infty} \int_0^t \int_0^{t_m} \dots \int_0^{t_2} H(0, t_1)(RJ)H(t_1, t_2)(RJ) \dots (RJ)H(t_m, t) dt_1 \dots dt_m$$

abbreviating

$$H(t_j, t_{j+1}) := \sum_{n=0}^{\infty} \int_{t_j}^{t_{j+1}} \int_{t_j}^{u_n} \cdots \int_{t_j}^{u_2} (-R_{u_1-t_j}) (-R_{u_2-u_1}) \cdots \\ \cdots (-R_{u_n-u_{n-1}}) Id_{t_{j+1}-u_n} du_1 \cdots du_n$$

for all $j \in \{0, \dots, m\}$, defining $t_0 := 0$, and using

$$(RJ)_{t_j} = (R \circ \Phi_{t_j})J = (Id \circ \Phi_{t_j})RJ = Id_{t_j}RJ \quad (6)$$

Now lemma 2 yields

$$\sum_{n=0}^{\infty} \int_0^t \int_0^{t_n} \cdots \int_0^{t_2} (Q \circ \Phi_{t_1})(Q \circ \Phi_{t_2-t_1}) \cdots (Q \circ \Phi_{t_n-t_{n-1}})(Id \circ \Phi_{t-t_n}) dt_1 \cdots dt_n \\ = \sum_{m=0}^{\infty} \int_0^t \int_0^{t_m} \cdots \int_0^{t_2} e^{-\int_0^{t-1} R_u du} Id_{t_1}(RJ) e^{-\int_0^{t_2-t_1} R_u du} Id_{t_2-t_1}(RJ) \cdots \\ \cdots e^{-\int_0^{t_m-t_{m-1}} R_u du} Id_{t_m-t_{m-1}}(RJ) e^{-\int_0^{t-t_m} R_u du} Id_{t-t_m} du_1 \cdots du_m$$

which together with relation (6) and theorem 1 proves the theorem.

☺

Lemma 1 Choose $n \in \mathbb{N}$ and let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ denote a symmetric function, which is integrable over every finite interval. Then the equation

$$\int_s^t \cdots \int_s^t f(t_1, \dots, t_n) dt_1 \cdots dt_n = n! \cdot \int_s^t \int_s^{t_n} \cdots \int_s^{t_2} f(t_1, \dots, t_n) dt_1 \cdots dt_n$$

holds for all $s < t \in \mathbb{R}^+$.

Proof: Since for every permutation $\sigma \in S_n$,

$$\int_s^t \int_s^{t_n} \cdots \int_s^{t_2} f(t_1, \dots, t_n) dt_1 \cdots dt_n = \int_{s < t_1 < \dots < t_n < t} f(t_1, \dots, t_n) dt_1 \cdots dt_n \\ = \int_{s < t_1 < \dots < t_n < t} f(t_{\sigma(1)}, \dots, t_{\sigma(n)}) dt_1 \cdots dt_n \\ = \int_{s < t_{\sigma^{-1}(1)} < \dots < t_{\sigma^{-1}(n)} < t} f(t_1, \dots, t_n) dt_1 \cdots dt_n$$

we conclude

$$\int_s^t \cdots \int_s^t f(t_1, \dots, t_n) dt_1 \cdots dt_n = \sum_{\sigma \in S_n} \int_{s < t_{\sigma(1)} < \dots < t_{\sigma(n)} < t} f(t_1, \dots, t_n) dt_1 \cdots dt_n \\ = n! \cdot \int_s^t \int_s^{t_n} \cdots \int_s^{t_2} f(t_1, \dots, t_n) dt_1 \cdots dt_n$$

which was to be proven.

☺

Lemma 2 For every $s < t \in \mathbb{R}^+$, we have

$$\begin{aligned} & \sum_{n=0}^{\infty} \int_s^t \int_s^{t_n} \dots \int_s^{t_2} (-R \circ \Phi_{t_1-s})(-R \circ \Phi_{t_2-t_1}) \dots (-R \circ \Phi_{t_n-t_{n-1}}) (Id \circ \Phi_{t-t_n}) dt_1 \dots dt_n \\ & = e^{-\int_0^{t-s} R \circ \Phi_u du} (Id \circ \Phi_{t-s}) \end{aligned}$$

Proof: For every $n \in \mathbb{N}$, $x \in E$ and $A \in \sigma(E)$, the function

$$\begin{aligned} f(t_1, \dots, t_n) & := (-R \circ \Phi_{t_1-s})(-R \circ \Phi_{t_2-t_1}) \dots (-R \circ \Phi_{t_n-t_{n-1}}) (Id \circ \Phi_{t-t_n})(x, A) \\ & = (-R(\Phi_{t_1-s}(x), \{\Phi_{t_1-s}(x)\})) \dots (-R(\Phi_{t_n-s}(x), \{\Phi_{t_n-s}(x)\})) 1_A(\Phi_{t-s}(x)) \end{aligned}$$

is symmetric in t_1, \dots, t_n , as can be shown by induction on $n \in \mathbb{N}$. Furthermore, for every $x = (m, y) \in E$ the above function is bounded by $(\lambda + m\mu)^n$, hence integrable over every finite interval.

Then lemma 1 yields for the n -th sum entry of the left side

$$\begin{aligned} & \int_s^t \int_s^{t_n} \dots \int_s^{t_2} (-R \circ \Phi_{t_1-s})(-R \circ \Phi_{t_2-t_1}) \dots (-R \circ \Phi_{t_n-t_{n-1}}) (Id \circ \Phi_{t-t_n}) dt_1 \dots dt_n \\ & = \frac{1}{n!} \left(\int_0^{t-s} -R \circ \Phi_u du \right)^n (Id \circ \Phi_{t-s}) \end{aligned}$$

which proves the statement.

☺

Remark 4 A mere restatement of the above theorem 2 in different terms of writing is the formula

$$\begin{aligned} P(t; x, A) & = \sum_{n=0}^{\infty} \int_0^t \int_0^{t_n} \dots \int_0^{t_2} \int_{(\mathbb{R}_0^+)^k} Q((x - t_1 1_k)^+, dy_1) \dots \\ & \dots \int_{(\mathbb{R}_0^+)^k} Q((y_n - (t_n - t_{n-1}) 1_k)^+, dz) (Id \circ \Phi_{t-t_n})(z, A) dt_1 \dots dt_n \end{aligned}$$

Remark 5 Expressions as in theorems 1 and 2 also hold for general piecewise-deterministic processes without jumps from a border set, i.e. with $t_*(x) = \infty$ for all $x \in E$. This can be seen from the fact that in the above derivations the special form of Q has not been exploited. Note that the letter Q in this section denotes an infinitesimal generator defined in terms of the transition measures and not, as in remark 1, the transition measure itself.

Remark 6 The formulae derived in theorem 2 and in remark 4 can be computed by the same kind of iteration as given in (3). Namely, for the result of theorem 2 we define the start values $P_0(t) := Id \circ \Phi_t$ for all $t \in \mathbb{R}_0^+$ and the iteration

$$P_{n+1}(t) := \int_0^t (Q \circ \Phi_s) P_n(t-s) ds + P_0(t)$$

for all $t \in \mathbb{R}_0^+$. Then we have the limit $P(t) = \lim_{n \rightarrow \infty} P_n(t)$.

Of course, this way of computing the transient distributions is valid for the generalizations of remark 5, too.

3.2 Stability and Stationary Distribution

For the M/G/k retrial queue, an analysis regarding stability and the computation of a stationary distribution can be performed along the lines of section 2.2. Again one obtains a QBD–matrix with kernel entries which is to be solved, although this time the matrix is level–dependent.

Analogous to section 2.2, we define the resolvents

$$L((n, x), \{m\} \times A) = \delta_{m,n} \cdot \int_0^\infty e^{-(s+\lambda s+n\cdot\mu)} 1_A((x - s \cdot 1_k)^+) ds$$

as well as

$$K((n, x), \{m\} \times A) = \int_0^\infty e^{-(s+\lambda s+n\cdot\mu)} \lambda Q_1((n, (x - s \cdot 1_k)^+), \{m\} \times A) \\ + n \cdot \mu Q_2((n, (x - s \cdot 1_k)^+), \{m\} \times A) ds$$

for $n, m \in \mathbb{N}$, $x \in (\mathbb{R}_0^+)^k$ and $A \in \sigma((\mathbb{R}_0^+)^k)$, as well as $G := K + L$.

Then the kernel G can be written as a block matrix

$$G(x, A) = \begin{pmatrix} G_{00}(x, A) & G_{01}(x, A) & 0 & 0 & 0 & \dots \\ G_{10}(x, A) & G_{11}(x, A) & G_{12}(x, A) & 0 & 0 & \dots \\ 0 & G_{21}(x, A) & G_{22}(x, A) & G_{23}(x, A) & 0 & \dots \\ 0 & 0 & G_{32}(x, A) & G_{33}(x, A) & G_{34}(x, A) & \ddots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots \end{pmatrix}$$

with kernel entries G_{nm} , i.e. G has the form of a level–dependent block tridiagonal matrix.

Set $A = [0, y_1[\times \dots \times [0, y_k[$ with $y_1, \dots, y_k \in \mathbb{R}^+$. As in section 2.2, the kernel entries of G are determined by the values

$$G_{n,n}(x, A) = \int_0^\infty e^{-(s+\lambda s+n\cdot\mu s)} 1_A((x - s \cdot 1_k)^+) ds \\ + \int_0^\infty e^{-(s+\lambda s+n\cdot\mu s)} \lambda \prod_{j=1, j \neq i}^k 1_{[0, y_j[}(x_j - s) B([0, y_i[) ds \\ = \frac{1}{1 + \lambda + n \cdot \mu} \left(e^{-(1+\lambda+n\cdot\mu)M(x-y)^+} + \lambda B([0, y_i[) e^{-(1+\lambda+n\cdot\mu)M(x-y)^+} \right)$$

for $i = \min\{1 \leq j \leq k : x_j = 0\}$ and

$$G_{n,n}(x, A) = \int_0^\infty e^{-(s+\lambda s+n\cdot\mu s)} 1_A((x - s \cdot 1_k)^+) ds \\ + \int_0^\infty e^{-(s+\lambda s+n\cdot\mu s)} n \cdot \mu 1_A((x - s \cdot 1_k)^+) ds \\ = \frac{1 + n \cdot \mu}{1 + \lambda + n \cdot \mu} e^{-(1+\lambda+n\cdot\mu)M(x-y)^+}$$

if $x_j > 0$ for all j , as well as

$$\begin{aligned} G_{n,n-1}(x, A) &= \int_0^\infty e^{-(s+\lambda s+n\cdot\mu s)} n \cdot \mu \prod_{j=1, j \neq i}^k 1_{[0, y_j[}(x_j - s) B([0, y_i]) ds \\ &= n \cdot \mu B([0, y_i]) e^{-(1+\lambda+n\cdot\mu)M(x-y)^+} \end{aligned}$$

for $i = \min\{1 \leq j \leq k : x_j = 0\}$ and

$$\begin{aligned} G_{n,n+1}(x, A) &= \int_0^\infty e^{-(s+\lambda s+n\cdot\mu s)} \lambda 1_A((x - s \cdot 1_k)^+) ds \\ &= \lambda e^{-(1+\lambda+n\cdot\mu)M(x-y)^+} \end{aligned}$$

if $x_j > 0$ for all j . As in section 2.2, we have defined $M(x - y)^+ := \max\{0, (x_1 - y_1), \dots, (x_k - y_k)\}$ for all $x, y \in (\mathbb{R}_0^+)^k$.

In the same way as Tweedie [16] generalized the analysis of $G/M/1$ -type block matrices towards $G/M/1$ -type block matrices with kernel entries, one can generalize the method in Hofmann [13, 14] for the analysis of level-dependent $M/G/1$ -type matrices, and hence for the above level-dependent QBD-type kernel G . This leads to stability conditions for G as well as to an algorithm for its solution. Then the solution of G leads to the stationary distribution via formula (5), as described at the end of section 2.2.

Remark 7 The generalization of results in Hofmann [13] towards $M/G/1$ -type matrices with kernel entries allows the inclusion of BMAP arrivals into the model. Then the kernel G would become a $(\mathbb{N}_0 \times \{1, \dots, m\})^2$ -block matrix with kernel entries if it is ordered according to the number of users in the system and the current phase of the arrival stream.

References

- [1] J. Azema, M. Kaplan-Duflo, and D. Revuz. Mesure invariante sur les classes recurrentes des processus de Markov. *Z. Wahrscheinlichkeitstheor. Verw. Geb.*, 8:157–181, 1967.
- [2] D. Baum. Convolution Algorithms for BMAP/G/1 Queues. Research Report No.96–22, Dept. of Mathematics and Computer Science, University of Trier, 1996.
- [3] D. Baum. A $BMAP|G|1$ -analysis based on convolution calculus. *J. Math. Sci., New York*, 92(4):3990–4002, 1998.
- [4] L. Breuer. *Spatial Queues*. PhD thesis, University of Trier, Germany, 2000.
- [5] L. Breuer. The Periodic BMAP/PH/c Queue. *Queueing Systems*, 38(1):67–76, 2001.
- [6] L. Breuer, A. Dudin, and V. Klimenok. A Retrial BMAP/PH/N System. *Queueing Systems*, 40(4):431–455, 2002.

- [7] O. Costa. Stationary distributions for piecewise-deterministic Markov processes. *J. Appl. Probab.*, 27(1):60–73, 1990.
- [8] D. Cox. The analysis of non-Markovian stochastic processes by the inclusion of supplementary variables. *Proc. Camb. Philos. Soc.*, 51:433–441, 1955.
- [9] M. Davis. Piecewise-deterministic Markov processes: A general class of non-diffusion stochastic models. *J. R. Stat. Soc., Ser. B*, 46:353–388, 1984.
- [10] M. Davis. *Markov models and optimization*. London: Chapman & Hall, 1993.
- [11] F. Dufour and O. L. Costa. Stability of piecewise-deterministic Markov processes. *SIAM J. Control Optimization*, 37(5):1483–1502, 1999.
- [12] B. Gnedenko and I. Kovalenko. *Introduction to queueing theory*. Boston, MA: Birkhaeuser, 1989.
- [13] J. Hofmann. *The BMAP/G/1 queue with level dependent arrivals: An extended queueing model for stations with nonrenewal and state dependent input traffic*. PhD thesis, University of Trier, Germany, 1998.
- [14] J. Hofmann. The BMAP/G/1 Queue with Level-Dependent Arrivals - An Overview. *Telecommunication Systems*, 16(3-4):347–359, 2001.
- [15] B. F. Nielsen and V. Ramaswami. A Computational Framework for a Quasi Birth and Death Process with a Continuous Phase Variable. In V. Ramaswami and P. Wirth, editors, *Teletraffic Contributions for the Information Age. Proceedings of the 15th International Teletraffic Congress - ITC 15*, pages 477–486, Amsterdam, 1997. Elsevier, North-Holland.
- [16] R. Tweedie. Operator-geometric stationary distributions for Markov chains, with application to queueing models. *Adv. Appl. Probab.*, 14:368–391, 1982.