Passage Times in Fluid models with Application to Risk Processes

V. Ramaswami
AT&T Labs-Research, 180 Park Avenue
Florham Park, NJ 07932, USA

Abstract

An efficient quadratically convergent algorithm has been derived earlier by Ahn & Ramaswami for computing the busy period distribution of the canonical fluid flow model. In this paper, we derive formulae for a variety of passage time distributions in the canonical fluid flow model in terms of its busy period distribution and that of its reflection about the time axis. These include several passage time distributions with taboo not only of the fluid level 0 but also of a set $[a, \infty)$ of levels. These are fundamental to the analysis of a large set of complex applied probability models, and their use is illustrated in the context of a general insurance risk model with Markovian arrival of claims and phase type distributed claim sizes, a context in which we have also introduced some new ideas that make the analysis very transparent.

Key words : Insurance Risk, Fluid-flow, Transient Results, matrix-geometric method.

1 Introduction

In this paper, we consider the canonical Markov modulated fluid flow model and derive a number of passage time distributions for it exploiting fully the previously obtained results of Ramaswami [16] and Ahn & Ramaswami [2], [3], [4]. (Our other related papers are [1], [5].) All passage time distributions considered here are characterized explicitly in terms of a matrix $\Psi(s)$ giving the Laplace-Stieltjes transforms of the first return time to fluid level 0 in each of two fluid flows that are reflections of each other, given that an upward trajectory of the concerned flow starts at level 0 at the time origin. (A matrix arises due to the consideration of the environmental states at the beginning and at the end of the interval under consideration.) In [3], we

have developed a powerful algorithm for computing $\Psi(s)$ which is similar to the Latouche-Ramaswami algorithm for QBDs [11] and has quadratic convergence. We have demonstrated in [3] its theoretical properties as well as computational advantages. For completeness, that algorithm is presented in Appendix 1 of this paper.

The passage time distributions for the fluid model derived here, particularly the passage times involving taboo positive levels (see Section 2), are fundamental quantities that can be used in a straightforward manner to analyze many complex applied probability models that may be modeled as fluid flows. While we exemplify this here (only) in the context of a simple insurance risk model, other examples in which these can be used are:

(i) Enhancements of the insurance risk model to such as those involving: (a) a cap on the reserve such that when the reserve attains the cap, all premiums have to be paid out as dividends until the reserve drops below the cap; (b) a threshold at which the reserve is replenished by a (possibly random) amount; (c) multiple thresholds and dividend payouts, and/or addition to reserves, based on the level of reserve quantized with respect to these thresholds.

(ii) Fluid queues with or without a finite fluid buffer operating under various threshold policies.

(iii) Financial risk models, where a capital amount is lent to borrowers, grows through interest receipts and suffers randomly occurring defaults.

(iv) Various continuous state space inventory models.

(v) General state dependent fluid flow models where dependency on the fluid level is quantized in terms of a discrete number of fluid levels. (The reader may wish to refer to [17], [8] for a fairly exhaustive survey on fluid flow models and their application to queues, risk models, etc.)

Note that many of the fluid models cited above involving upper boundaries and thresholds have hitherto resisted detailed time dependent analysis for lack of tractable formulae characterizing passage times with a taboo not only of fluid level 0 but also of some positive levels $[a, \infty)$. We have characterized such taboo passage times here, and that should make the analysis of these models now eminently possible. In that sense, this paper does indeed set the stage for the examination of a large variety of highly complex applied probability models using a unified approach. Given the success of the matrix-geometric methods of M.F. Neuts [14] in the context of discrete state space processes, this should not be surprising, given that our own starting point [16] (see also [4]) may be viewed as an extension of matrix-geometric type arguments for a continuous state space process.

The example we have used here is an insurance risk process subject

to a Markovian arrival process of claims and phase type claim sizes that allows for incorporation of various types of dependencies, such as among inter-claim intervals, among claim sizes, dependency of the claim size on the state of the environmental process generating claims, etc. We also construct a highly convenient fluid flow model closely associated with it. The fluid flow formulation we adopt is novel and is demonstrated to be such that translation of time dependent results from the fluid flow model to those for the insurance risk model is exceptionally straightforward.

1.1 Fluid Flow Models

The (unrestricted) fluid flow model under consideration

$$(\mathcal{F}, \mathcal{J}) = \{(F(t), J(t)) : \ t \geq 0\},$$

where $F(t)$ is the level of a fluid buffer and $J(t)$ the state of an environmental process is specified by the following conditions:

(a) $\mathcal{J} = \{J(t), t \geq 0\}$ is a continuous time Markov chain with state space $S = S_1 \cup S_2 \cup S_3$ and infinitesimal generator Q that, when partitioned according to the sets S_i, has the form

$$Q = \begin{pmatrix} Q_{11} & Q_{12} & Q_{13} \\ Q_{21} & Q_{22} & Q_{23} \\ Q_{31} & Q_{32} & Q_{33} \end{pmatrix}. \tag{1}$$

Specifically, the submatrix Q_{ij} contains the elements $Q(r, s)$ of the infinitesimal generator with $r \in S_i$ and $s \in S_j$. [Throughout, for any matrix A, we denote its (r, s)-th element as $A(r, s)$ or as $[A]_{rs}$ and reserve the notation A_{ij} to denote the submatrix of A formed by its elements $A(r, s)$ with $r \in S_i$ and $s \in S_j$.]

(b) The fluid level process $\{F(t) : t \geq 0\}$ is such that: during sojourn of $\mathcal{J}$ in state $i \in S_1$, the fluid level increases at rate $c(i) > 0$; during sojourn of $\mathcal{J}$ in state $i \in S_2$, the fluid level decreases at (absolute) rate $c(i) > 0$; during sojourn of $\mathcal{J}$ in S_3, the fluid level does not change. Note specifically that given $J(t) = i \in S_2$, the instantaneous rate of change of the fluid at time t is $-c(i) < 0$.

Denote by c the vector with elements $c(i)$, $i \in S_1 \cup S_2$, and note that following our convention the subvector $c_i = (c(j) : j \in S_i)$, $i = 1, 2$. For later purposes, we shall refer to the above canonical model as $\mathcal{F}(Q, c_1, -c_2)$.

Reflection: Associated closely with the fluid flow $\mathcal{F} = \mathcal{F}(Q, c_1, -c_2)$ is the fluid flow $\mathcal{F}^r = \mathcal{F}(Q, -c_1, c_2)$ obtained by reversing the roles of the up and

down environment states. The paths of each of these processes starting at fluid level x at time 0 is obtained by reflecting about a line parallel to the time axis drawn at height x of a corresponding path of the other, and that explains the choice of our notation. As we shall see later, for computing certain first passage times of interest, it becomes convenient to consider both these processes simultaneously.

Restricted Fluid Flow: Associated with the fluid flow model $\mathcal{F}(c_1, -c_2)$ is also the restricted fluid flow model with boundary at fluid level 0, say $\mathcal{F}^+ = \mathcal{F}^+(Q, c_1, -c_2)$ which is such that the fluid level is allowed to decrease only when it is positive. Thus, when $J(t) = i \in S_2$, the fluid level decreases at (absolute) rate $c(i)$ as long as it remains positive but otherwise remains constant at level zero until an upward change can occur by a transition of the environment into the set S_1. The restricted fluid flow model corresponding to $\mathcal{F}^r$ shall be denoted in what follows by $\mathcal{F}^{r+}$.

1.2 Insurance Risk Process

The canonical insurance risk process we consider is the process of risk reserve of an insurance company that starts with a reserve $u \geq 0$, collects premiums at a constant rate $\gamma > 0$ and pays out claims as they occur until it goes broke. Typical quantities of interest with respect to the insurance reserve process are the probability of ruin and the joint distribution of the various combinations of the time to ruin, the surplus immediately prior to ruin, the deficit immediately after ruin, and the size of the claim that causes ruin. Their determination is important in assessing required risk reserves, replenishments thereof, re-insurance levels, etc. These have been the subject of several past studies using specialized methods; see [9] and references therein; for a detailed set of references, the reader may also refer to Asmussen [8]. One of our goals here is to characterize all these in terms of a couple of fundamental quantities for a closely associated fluid model for which highly efficient algorithms exist.

To enable analysis using fluid flow models, we assume that the claims process is a general Markovian point process with claim sizes distributed as phase type distributions whose parameters may or may not depend on the environmental Markov process. Specifically, it is assumed that an underlying continuous time Markov chain (CTMC) on a finite state space $E = \{1, \cdots, m\}$ evolves such that the instantaneous rate of transition of the environment from state i to state $j \neq i$ in E without an accompanying claim is given by the quantity $D_0(i, j) \geq 0$, and the instantaneous rate of

transition from i to j (possibly with $j = i$) in E with a claim is given by $D_1(i,j) \geq 0$; the diagonal elements of D_0 are assumed to be negative and such that the row sums of the matrix $D_0 + D_1$ are all zero. The pair of matrices D_0 and D_1 thus define a general Markovian claim arrival process. A claim that occurs at an epoch of transition from state i to j will be assumed to have a phase type distribution [14], [12] with representation (α_{ij}, T_{ij}).

The general formulation above allows for unification of many special cases of claim processes modeled as Markov additive processes with phase type jumps. Our formulation allows for correlation among interarrival times of jumps as well as the possibility of the jump sizes depending on the environmental Markov chain. Some simple examples are given below for illustration:

Example 1: Exponentially distributed claim sizes occurring at Poisson epochs: Here $D_0 = -\lambda$ and $D_1 = \lambda$, where λ is the rate of the Poisson process of claim epochs and $\alpha = 1, T = -\mu$ where μ^{-1} is the average claim size.

Example 2: Phase type distributed claim sizes occurring at epochs of a phase type renewal process: In this model, the interarrival times of claims are assumed to have a continuous phase type distribution (see [14], [12]) with parameter (β, M) where β is a probability vector of size m, and M is an infinitesimal generator of a CTMC with m states. For this model, $D_0 = M$ and $D_1 = -M\mathbf{1}\beta$, where throughout $\mathbf{1}$ represents a column vector of 1's of appropriate dimension. The case of independent claim sizes corresponds to the case where α_{ij} and T_{ij} do not depend on i, j, while allowing these to vary according to i gives the ability to model dependency of the claim size on the state of the environmental CTMC just prior to each claim arrival epoch, while allowing them to vary according to j gives the ability to model dependence on the environmental CTMC just after the claim arrival; of course, both dependencies are incorporated by letting these parameters vary both over i and j simultaneously.

The general scenario described corresponds to the case where the claim arrival instants occur according to a MAP (Markovian arrival process) which is a special case of a process defined by Neuts [13] (see also [12]), and the claim sizes may be allowed to depend on the states of the MAP prior to and posterior to claim epochs as in Example 2 or is simply independent of either or both of them. An example where dependencies in the sizes of successive claims on the states of the MAP may become necessary is one where the claim process arises as the superposition of claims arising from different types of catastrophes, where each catastrophe has its own pattern of (possibly correlated) claim interarrival times and claim sizes generated by it. It is

well-known [6] that MAPs are dense in the class of all point processes and that phase type distributions are dense in the class of all distributions on $[0, \infty)$. Thus, our model class is indeed very general and allows for great flexibility in modeling.

Insurance Process modeled as a Fluid Flow

The key idea here is to pretend as though each claim of size x arrives continuously over (an unseen) time interval whose length is x/γ at rate γ per unit time. Thus, the effect of a claim distributed as $PH(\alpha, T)$ is treated as though a downward journey is made by the risk reserve process over a time interval of length distributed as $PH(\alpha, \gamma T)$ during which the reserve (as long as it exists) is depleted at rate γ. That gives rise to a fluid model (with S_3 being empty) where upward rates are all γ and downward rates are all $-\gamma$. The paths of the risk process before ruin can be obtained from segments of the fluid process before the fluid level becomes empty, and this is achieved by replacing downward linear paths in the fluid flow model by downward jumps of appropriate sizes. Thus, our artifice allows us to embed the risk process into a fluid flow model, and by incising out times spent in the downward states of the fluid model and gluing the rest together gives rise to the risk process. Our artifice, of course, changes clock time since we have introduced into the risk process spurious time intervals of downward linear descents (replacing instantaneous downward jumps), and one needs to make appropriate translation of clock times in the two processes. As we will see soon, the choice of a common constant γ for the absolute rates of ascent and descent makes the translation of results (with respect to clock time) from the fluid model to the risk model almost trivial.

With respect to the general risk model, it is convenient to assume, without loss of generality (see Appendix 2), that all phase type distributions $PH(\alpha_{ij}, T_{ij})$ have the same order, say n, which is the largest among those of the considered PH distributions. If $|E| = m$, this then allows for the consideration of a fluid model with just $m + m^2 n$ environmental states as described below, with further reductions of this number in special cases. This artifice not only reduces dimensionality of the model but also reduces much notational clutter arising by way of having to keep track of phases in separate phase type models.

It is now easy to verify that the artifice replacing downward jumps by linear segments and using a common size for the phase type distributions gives an instance of the fluid flow model, and furthermore, the various sets and parameters defining it are as given below.

(a) $Q_{11} = D_0$ and c_1 is an m-vector all whose elements are equal to γ. Clearly, any transition without a claim is a transition between two environmental states in the upward segment of the fluid with the rate of change γ per unit time (premiums are collected here at rate γ without claims being paid out.) To achieve this, we take the set S_1 as a copy of the state space, E, of the environmental Markov chain defining the claim arrival process which is defined as a MAP on the state space E.

(b) For the general case, the set S_2 is indexed by three indices (i, j, k) where $i \in E$ is the environmental state just prior to the claim epoch, $j \in E$ is the environmental state posterior to the claim epoch, and $k \in L$, where $L = \{1, \cdots, n\}$ is the set of transient states in the CTMC governing the phase type interval over which claims are viewed as occurring continuously at rate γ. This set is of size $m^2 n$ and is a copy of $E \times E \times L$. The matrix Q_{12} is such that its nonzero entries are all defined by the equation

$$Q_{12}(i, (i, j, k)) = D_1(i, j)\, \alpha_{ij}(k),$$

and this represents the rate at which a claim arrives marked by a transition of the environment state from i to j (no transition if $j = i$) and the initial state in the CTMC defining the resulting claim size being k.

(c) The matrix Q_{22} is such that its nonzero elements are all defined by the equation

$$Q_{22}((i, j, r), (i, j, s)) = \gamma\, T_{ij}(r, s), \quad 1 \leq r, s \leq n.$$

The introduction of the constant γ in the rate above is to scale the time interval over which claim occurs continuously at rate γ as is necessary in our construction.

(d) Finally, Q_{21} is such that all its nonzero elements are defined by the equation

$$Q_{21}((i, j, r), j) = \gamma\, [-\sum_{s \in L} T_{ij}(r, s)]$$

Note that at the end of an interval of descent in the constructed fluid model, the state of the environment process has to be the same as at its beginning, since the interval constructed is spurious and over that the environment process does not change.

(e) Of course c_2 is a $m^2 n$ vector, all whose elements are equal to γ.

The size $m^2 n$ for the set S_2 is of course an overkill in some instances. Specifically, if the phase type distributions of claim sizes do not depend on the environment prior to the claim arrival epoch, then the set S_2 can be

made smaller as the first component defining its states becomes redundant; we can get away with an S_2 of dimension mn only by keeping track just of the environmental state just after the claim epoch and the phase of the CTMC governing the PH-distribution.

In any case, the above discussion shows that even in the most general situation, a highly parsimonious fluid flow model can be derived from the risk reserve process, and our task is to analyze the risk reserve process through the resulting fluid flow model. This is done in Section 3 after deriving some required passage time distributions for the canonical fluid flow model in the next section.

2 Passage Times for Fluid Flows

In this section, we consider the fluid flow model $\mathcal{F} = \mathcal{F}(Q, c_1, -c_2)$ and consider a number of first passage time distributions of the associated restricted process $\mathcal{F}^+$ of interest. To that end, we introduce the following generic notations:

$\tau(x, y)$ = First passage time of $\mathcal{F}^+$ from (fluid) level x to level y

${}^a\tau(x, y)$ = First passage time of $\mathcal{F}^+$ from level x to level y avoiding a visit in the interim to the levels in the interval $[a, \infty)$

${}_a\tau(x, y)$ = First passage time of $\mathcal{F}^+$ from level x to level y avoiding a visit in the interim to the levels in the interval $[0, a]$

${}_a^b\tau(x, y)$ = First passage time of $\mathcal{F}^+$ from level x to level y avoiding a visit in the interim to the levels in $[0, a] \cup [b, \infty)$, $a < b$

Return time to level 0

We note that $B = {}_0\tau(0, 0)$ is the distribution of the first return time to level 0 in the restricted process $\mathcal{F}^+$, a quantity we called the busy period of the restricted fluid flow. In [3], we have provided an extremely efficient algorithm to compute the matrix transform $\Psi(s)$, $Re(s) \geq 0$ (throughout this paper, for transforms the argument s will be assumed to be such that $Re(s) \geq 0$) defined by the elements

$$\Psi(s; i, j) = \int_0^\infty e^{-st} d_t \mathcal{P}_{(0,i)}[B \leq t,\ J(B) = j], \quad i \in S_1,\ j \in S_2,$$

where $\mathcal{P}_{(x,i)}[A]$ denotes the conditional probability of the event A given that $F(0) = x$ and $J(0) = i$; for completeness, we have restated that algorithm in Appendix 1, and we refer to [3] for details of its derivation. Note that the

elements of $\Psi(\cdot)$ give the joint distribution of the busy period duration and the (downward) environmental state at the end of the busy period, given the (upward) environmental state at the start of the busy period. Note that since the rates of fluid change away from the boundary at 0 do not depend on the fluid level (spatial homogeneity property), the busy period distribution given above is also the distribution of first return times to any level $x > 0$ avoiding lower levels for the process $\mathcal{F}^+$; more precisely, for any $x \geq 0$ the distribution of ${}_x\tau(x,x)\chi\{J({}_x\tau(x,x)) = j\}$ given $(F(0), J(0)) = (x,i)$ is the same as that of ${}_0\tau(0,0)\chi\{J({}_0\tau(0,0)) = j\}$ given $(F(0), J(0)) = (0,i)$.

First Passage Time ${}_0\tau(x,0)$ from $x > 0$ to 0

We also recall one more matrix of transforms introduced earlier in [3], Theorem 3, namely,

$$H(s) = C_2^{-1}[Q_{22} - sI + Q_{23}(sI - Q_{33})^{-1}Q_{32} + \{Q_{21} + Q_{23}(sI - Q_{33})^{-1}Q_{31}\}\Psi(s)].$$

Note that for a model with $S_3 = \phi$, the empty set, the corresponding formula is obtained by dropping all terms in the above which involve the set S_3, and the formula simplifies in that case to

$$H(s) = C_2^{-1}[Q_{22} - sI + Q_{21}\Psi(s)].$$

From [3], the matrix $H(s)$ is such that for all $x \geq 0$, the matrix of transforms ${}_0\tilde{f}_{22}(x,0,s) = e^{H(s)x}$, which is a square matrix of order $|S_2|$, is the matrix of LSTs of a first passage time of $\mathcal{F}^+$ from level x to level 0 with paths starting with a downward segment. Specifically, for $i, j \in S_2$, the quantity $[{}_0\tilde{f}_{22}(x,0,s)]_{ij}$ gives the joint distribution of a first passage time to level 0 which hits level 0 in the environment state j, given that the process starts in (x,i); thus, it is the distribution of ${}_0\tau(x,0)\chi\{J({}_0\tau(x,0)) = j\}$ given the initial state (x,i), $i, j \in S_2$. Note that due to the spatial homogeneity of the model, for any $y > 0$, ${}_0\tilde{f}_{22}(x,0,s)$ also gives the first passage times to level y from level $y+x$. That is, ${}_0\tau(x,0)\chi\{J({}_0\tau(x,0)) = j\}$ given the initial state (x,i), $i \in S_2$ has the same distribution as that of ${}_y\tau(y+x,y)\chi\left(J({}_y\tau(y+x,y)) = j\right)$ given the initial state $(y+x,i)$.

First Passage Time ${}_0\tau(0,x)$ from 0 to $x > 0$ avoiding level 0

The quantity we wish to determine is the first passage time LST matrix to be denoted by ${}_0\tilde{f}_{11}(0,x,s)$, $x > 0$, which is of size $|S_1| \times |S_1|$ and whose (i,j)-th element gives the LST of the first passage time avoiding level 0 from

the state $(0,i)$, $i \in S_1$ to level x in the specific state (x,j). (With the self explanatory notations, note that due to spatial homogeneity of the model, ${}_0\tilde{f}_{11}(0,x,s) = {}_y\tilde{f}_{11}(y,y+x,s)$.) This involves several steps which we take below.

Explicitly in terms of $\Psi(s)$, recall the matrix of transforms $K(s)$ introduced in [2] such that

$$K(s) = C_1^{-1}(Q_{11} - sI) + \Psi(s)C_2^{-1}Q_{21} + \Theta(s)Q_{31},$$

where

$$\Theta(s) = C_1^{-1}Q_{13}(sI - Q_{33})^{-1} + \Psi(s)C_2^{-1}Q_{23}(sI - Q_{33})^{-1},$$

and $C_i = diag(c_i)$, $i = 1,2$ are diagonal matrices with positive elements on the diagonal. Note also that when $S_3 = \phi$, the formula for $K(s)$ simplifies to

$$K(s) = C_1^{-1}(Q_{11} - sI) + \Psi(s)C_2^{-1}Q_{21}.$$

We demonstrated (see [2], [4]) that for fixed $x > 0$ and $i,j \in S_1$, the matrix $\tilde{N}_{11}(x,s) = e^{K(s)x}$ is such that its (i,j)-th element $\tilde{N}_{11}(x,s;i,j)$ is the Laplace-Stieltjes transform (LST) of the distribution function (over time t) of the taboo expected number of visits in $[0,t]$ by the process $\mathcal{F}$ to the state (x,j) avoiding level 0 in the time interval $(0,t]$, given that the process starts in $(0,i)$.

Given that the matrix $\tilde{N}_{11}(x,s)$ is the LST of a Markov renewal kernel, note that for $x > 0$, (by the standard interpretation of a renewal density) we may interpret the quantity $\tilde{N}_{11}(x,s;i,j)$ as the Laplace transform (with respect to time) of the density of the fluid flow $\mathcal{F}^+$ crossing level x in the state (x,j) at time $t > 0$ avoiding level 0 in the interval $(0,t]$, given that the process starts in $(0,i)$ at time 0. This interpretation will be used in what follows in several places.

We first establish the following lemma.

Lemma 1 *Let $x > 0$ and $i,j \in S_1$. Given that the fluid flow starts in (x,i), the LST (with respect to time and transform variable s) of the expected number of visits to (x,j) avoiding level 0 in the time interval $(0,t]$ is the (i,j)-th element of the matrix*

$$\Xi(x,s) = \Psi(s)\int_{[0,x)} e^{H(s)y}[C_2^{-1}Q_{21} + C_2^{-1}Q_{23}(sI - Q_{33})^{-1}Q_{31}]e^{K(s)y}\,dy.$$

Proof: Assume that the fluid process starts in (x, i), $x > 0$, $i \in S_1$. Consider the Laplace transform of the elementary probability that the fluid process makes an upcrossing at time $t > 0$ of level x in the state (x, j), $j \in S_1$, avoiding level 0 in its path up to that point. During this interval, there is a unique fluid level $0 < x - y \leq x$ which is the lowest level visited by the process, and we can, by conditioning on the value of y, write

$$[\Xi(x,s)]_{ij} = \int_{[0,x)} \sum_{k \in S_2} \sum_{r \in S_1} [\Psi(s)]_{ik} \times [e^{H(s)y}\{C_2^{-1}Q_{21} + C_2^{-1}Q_{23}(sI - Q_{33})^{-1}Q_{31}\}]_{kr} [e^{K(s)y}]_{rj}\, dy.$$

(Remark: The above equation conditions also on the environmental state (k) at the first return to x and also the environmental state (r) immediately after the lowest point $x - y$ in the path; note that we need to account for the fact that after hittling the lowest level $x - y$, the process could spend some time there inthe set S_3 before starting another upward journey; the premultiplier C_2^{-1} before the terms Q_{21} and Q_{23} appears due to the fact that Q_{21} Q_{23} are rates with respect to infinitesimal time intervals and what is considered in the integrand is an infinitesimal interval over the fluid level; see [4], Remark 5.) Hence the lemma.

The following result is helpful in evaluating the matrix $\Xi(x, s)$. In this, given a matrix A, we denote by $vec(A)$ a column vector obtained by writing the successive columns of A one below the other. A useful result we wish to recall [10] is that for conformable matrices,

$$vec(AXB) = (B^t \otimes A)\, vec(X),$$

where B^t is the transpose of B, $\otimes$ denotes the Kronecker product of matrices.

Lemma 2 *Let*

$$U(x,s) = \int_{[0,x)} e^{H(s)y}\{C_2^{-1}Q_{21} + C_2^{-1}Q_{23}(sI - Q_{33})^{-1}Q_{31}\}e^{K(s)y}\, dy.$$

For all $Re(s) > 0$, we have

$$vec\,(U(x,s)) = [(K^t(s) \otimes I) + (I \otimes H(s))]^{-1} vec\left(e^{H(s)x}D(s)e^{K(s)x} - D(s)\right),$$

where

$$D(s) = C_2^{-1}Q_{21} + C_2^{-1}Q_{23}(sI - Q_{33})^{-1}Q_{31}.$$

Proof: By multiplying the integral to be evaluated on the right by $K(s)$ and integrating by parts gives

$$U(x,s)K(s) + H(s)U(x,s) = e^{H(s)x}D(s)e^{K(s)x} - D(s),$$

which is the same as the set of linear equations

$$[(K^t(s) \otimes I) + (I \otimes H(s))]vec\,(U(x,s)) = vec\left(e^{H(s)x}D(s)e^{K(s)x} - D(s)\right).$$

The proof would be complete if we show that the coefficient matrix in the above set of equations is nonsingular for all $Re(s) > 0$ which we set out to do now.

We have shown in [2], Lemma 1, that for all $Re(s) > 0$, the matrix $K(s)$ is invertible (and therefore each of its eigenvalues has strictly negative real part; see the proof of Lemma 1 in [2].) Therefore, for $Re(s) > 0$, we can rewrite the coefficient matrix of the above linear system as

$$\left(K^t(s) \otimes I\right)\left[(I \otimes I) + (\{K^t(s)\}^{-1} \otimes H(s))\right].$$

Noting that each eigenvalue of a Kronecker product is obtained as the product of a pair of eigenvalues of the two matrices constituting the Kronecker product ([10]), clearly, the inverse fails to exist at s only if there exist eigenvalues $\lambda(s)$ of $K(s)$ and $\mu(s)$ of $H(s)$ such that $\mu(s)/\lambda(s) = -1$. But then $\mu(s) = -\lambda(s)$ would have positive real part implying that the matrix $H(s)$ has an eigenvalue with positive real part. That would contradict the fact that for all $x > 0$ all eigenvalues of $e^{H(s)x}$ (which is the matrix of first passage time LSTs to level 0 from level x) are within the unit disk. This contradiction shows that the coefficient matrix of the linear system is indeed invertible, and the proof is complete.

Theorem 1 *For $x > 0$, the LST $[{}_0\tilde{f}_{11}(0,x,s)]_{ij}$ of the first passage avoiding level 0 from state $(0,i)$ into level x through entry into the specific state (x,j), $x > 0$, $i,j \in S_1$ is determined by the equation*

$$[{}_0\tilde{f}_{11}(0,x,s)] = [e^{K(s)x}][I + [\Xi(x,s)]^{-1}.$$

Proof: Let the fluid flow process start in $(0,i)$, with $i \in S_1$. Also, let $j \in S_1$. By noting that an upcrossing of x at the specific state (x,j) avoiding level zero occurs either at the first such upcrossing or as a subsequent return to level x avoiding level zero, we have

$$[e^{K(s)x}]_{ij} = [{}_0\tilde{f}_{11}(0,x,s)]_{ij} + \sum_{k \in S_1} [{}_0\tilde{f}_{11}(0,x,s)]_{ik}\,[\Xi(x,s)]_{kj}\ ;$$

hence the result.

Distribution of busy period ${}_0^x\tau(0,0)$ avoiding level $x > 0$

We can now write a simple formula for a busy period distribution of the flow that also avoids the levels in the interval $[x, \infty)$, $x > 0$.

Theorem 2 *For $x > 0$, the matrix*

$$ {}^x\Psi(s) = {}^x\tilde{f}_{12}(0,0,s) = \Psi(s) - {}_0\tilde{f}_{11}(0,x,s)\Psi(s)e^{H(s)x} $$

is the Laplace-Stieltjes transform matrix of a busy period of the flow $\mathcal{F}^+$ that does not touch the level x.

Proof: The term subtracted in the formula of the theorem characterizes the busy period that touches level x; in such busy periods, the flow must make a first passage into x, return to x again, and then proceed down to 0.

At this time, we note that we can introduce quantities similar to those introduced above for the restricted fluid flow $\mathcal{F}^{r+}$ associated with the reflection as well. Such quantities for that flow will be denoted by attaching a post-superscript r to the appropriate symbol. Thus, for instance $K^r(s)$ is the $K(s)$ matrix for the restricted flow associated with the reflected flow, $\Psi^r(s)$ its busy period, $\tau^r(x,y)$ the first passage time in it from level x to level y, etc.

The most important point to note is that all the passage times introduced for a flow thus far, are determined explicitly in terms of its busy period transform. Except for $\Xi(x,s)$, the computation of all other quantities is trivial once the busy period transform has been determined.

First Passage ${}_0^x\tau(x,0)$ from $x > 0$ to 0 without a return to x

Assume that the fluid model $\mathcal{F}^+$ starts in $x > 0$ with a downward segment, i.e., in a state (x,i) with $i \in S_2$. Denote by ${}^x\tilde{f}_{22}(x,0,s)$ the matrix of LSTs of first passage times to level 0 avoiding a return to level x.

Theorem 3 *We have,*

$$ {}^x\tilde{f}_{22}(x,0,s) = {}_0\tilde{f}^r_{22}(0,x,s), $$

where by the choice of notations, the right side of the above equation gives the first passage time from level 0 to level x avoiding 0 for the reflected flow model.

Proof: Consider a path of $\mathcal{F}$ that goes from x to 0 without visiting x in between. Reflecting this about a line parallel to the time axis drawn at fluid level x gives a path for the reflected flow that starts with an upward segment and goes to $2x$ without returning to x. Due to spatial homogeneity, the distribution of the time to go from x to $2x$ without touching x is the same as that of going from 0 to x without touching zero.

Return times to level x avoiding level 0

Let $x > 0$. If the fluid process $\mathcal{F}$ starts at level x with an upward segment, then the distribution of its return to fluid level x is governed by $\Psi(s)$ and such returns end in states of S_2; such returns avoid level 0, of course. The following result provides the formula for a return to x avoiding level 0 when the starting segment is a downward segment.

Theorem 4 *Assume that the fluid flow starts in level $x > 0$ with environmental phase in the set S_2 (i.e., a downward segment). The distribution of the return time to level x avoiding level 0 is given by the transform matrix*

$$ {}_0\tilde{f}_{21}(x,x,s) = {}^x\Psi^r(s) = \Psi^r(s) - {}_0\tilde{f}^r_{22}(0,x,s)\Psi^r(s)e^{H^r(s)x}. $$

Proof: The result is trivial by considering a reflection of a path returning to x avoiding level 0 about a line drawn parallel to the time axis at height x. The reflection is clearly a return time to x in the reflection avoiding level $2x$ and, due to spatial homogeneity, is distributed the same as the return time to 0 avoiding x. Thus, the result follows by applying Theorem 2 to the reflected flow.

3 Back to the Risk Process

We now return to the insurance risk process and determine a bunch of interesting quantities related to it. We assume as before that the initial risk reserve is u and the premium rate is γ.

The following observations are all trivial and follow from the fact that for the fluid flow associated with the risk process, upward and downward rates are all equal in absolute value to a common constant:

Observations:

(a) Let T denote the time to ruin in the insurance risk process and τ denote the first passage time to hit level 0 in the associated fluid flow model. T is the sum of the lengths of the intervals in $[0, \tau]$ during which the fluid level is increasing.

(b) For any interval of time during which the fluid flow returns to a level $u \geq 0$ without touching 0 in between, the length of time of that interval is twice the time the fluid flow moves upward within that interval.

(c) Let $0 \leq v < u$. Consider an interval of first passage of the fluid flow $\mathcal{F}$ from u to v. Denote by t the total length of that interval and by τ the length of time the fluid moves upward during that interval. Then clearly, $\tau = (t - (u-v)/\gamma)/2$; this is so, once again due to the fact that upward and downward rates are both equal to γ in absolute value.

(d) Consider a first passage interval for the fluid flow from a level $u \geq 0$ to $v > u$ without visiting 0 in between. Denote by t the total length of that interval and by τ the length of time the fluid moves upward during that interval. Then clearly,

$$\tau = \frac{v-u}{\gamma} + \frac{1}{2}(t - \frac{v-u}{\gamma}) = t/2 + (v-u)/(2\gamma);$$

this again is a consequence of of the fact that all rates for the fluid flow are equal in absolute value. Note that it follows that the total total length of time during which the fluid flow is going downward in the considered first passage time is $t/2 - (v-u)/(2\gamma)$.

These considerations immediately allow us to write down a variety of results for the insurance risk model from the fluid flow model in terms of quantities established in the previous section. We start with the simplest, the distribution of the time to ruin.

Theorem 5 *Assume that the insurance risk process starts in the state* (u, i), $u \geq 0$, $i \in S_1$. *Denote the LST of the distribution of the time to ruin by* $[R(u,s)]_i$. *The vector* $R(u,s)$ *with elements* $[R(u,s)]_i$, $i \in S_1$ *is given by*

$$R(u,s) = \Psi(s/2)\,[e^{\frac{su}{2\gamma}}\, e^{H(s/2)u}]\,\mathbf{1}, \quad u \geq 0.$$

Proof: In the case $u = 0$, by our Observations (a),(b) given above, the time to ruin of the insurance risk process is exactly one half the busy period duration of the associated fluid flow model, and that gives the formula $R(0,s) = \Psi(s/2)\mathbf{1}$. For $u > 0$, since the initial segment is upward, first the

flow has to return to u in a downward state, and then make a first passage to 0. The result now follows by using Observation (b) on the return time to u and Observation (c) on the first passage time to 0 that follows.

Corollary 1 *For the risk model starting with risk reserve u, the vector of ruin probabilities is given by $R(u,0) = \Psi(0)e^{H(0)u}\mathbf{1}$.*

Joint Distribution of time to ruin, surplus X_1 before, and deficit X_2 after ruin.

In each case considered below, we consider different starting states $i \in S_1$ and compute the vector of joint distributions whose i-th component gives the required joint distribution given the starting state $i \in S_1$.

Case 1: $u = 0$

When the initial risk reserve is 0, we can write the following formula immediately for the vector of joint distributions for the ruin time, surplus before ruin and deficit immediately after ruin (LST with respect to ruin time and probability differential for X_1 and X_2):

$$e^{-sx_1/(2\gamma)}e^{K(s/2)x_1}[\gamma^{-1}Q_{12}]e^{Q_{22}(x_1+x_2)/\gamma}[\gamma^{-1}Q_{21}]\mathbf{1}\; dx_1\, dx_2;$$

of these, due to Observation (d), the term $e^{-sx_1/(2\gamma)}e^{K(s/2)x_1}$ characterizes the total length of upward intervals up to reaching the surplus x_1 prior to ruin (without getting ruined earlier); at such a point a change of environment in the fluid model occurs into a downward state and lasts precisely for a length $(x_1 + x_2)/\gamma$ wiping out the reserve x_1 and leaving a deficit x_2; the final multiplication by $\mathbf{1}$ is so that we don't care for the environmental state immediately after ruin in the insurance risk process; the appearance of γ^{-1} before the terms Q_{12} and Q_{21} is since we are considering a differential over level as opposed to over time and γ^{-1} gives the required Jacobian (see Remark 1 in [4].)

Case 2: $u > 0$ **and the reserve just before ruin is** $X_1 = u + x_1 > u$

By considering that several returns to u avoiding level 0 may occur before ruin, we can write the required vector containing the LST of ruin time and pdf of X_1 and X_2 as:

$$[I + \Xi(u, s/2)][e^{-sx_1/(2\gamma)}e^{K(s/2)x_1}][\gamma^{-1}Q_{12}]e^{Q_{22}(u+x_1+x_2)/\gamma}[\gamma^{-1}Q_{21}]\mathbf{1}.$$

Here the term $[I + \Xi(u, s/2)]$ takes care of all the returns to u avoiding 0, and the term immediately after that takes care of the interval of final ascent from u to $u + x_1$ avoiding u; we have used both Observations (b) and (d) in the above.

Case 3: $u > 0$ and the reserve just before ruin is $X_1 = u - x_1 < u$.

It is easy to write the final formula for the required vector and then explain what goes on:

$$\begin{aligned}\Psi(s/2)[e^{sx_1/(2\gamma)}\, e^{H(s/2)x_1}] \\ \times\ {}^{(u-x_1)}\Psi^r(s/2)\,[I + \Xi(u - x_1, s/2)] \\ \times\ [\gamma^{-1}Q_{12}]e^{Q_{22}(u-x_1+x_2)/\gamma}[\gamma^{-1}Q_{21}]\mathbf{1}.\end{aligned}$$

The sketch of the derivation of the above formula goes as follows: The initial segment at level u being upward, the flow should first return to u (this gives the first term $\Psi(s/2)$ due to Observation (b)). Then it needs to make a first passage into $u - x_1$; that gives the second term $[e^{sx_1/(2\gamma)}\, e^{H(s/2)x_1}]$ by Observation (c). Then the flow has to make an upcrossing of $u - x_1$ without visiting level 0 (no ruin); and this term after a reflection about the line parallel to the time axis at $u - x_1$ (draw a picture!) gives the term ${}^{(u-x_1)}\Psi^r(s/2)$ since the reflected path is a busy period with a taboo upper level for the reflected flow; we have of course used Observation (b) in this. Then the flow may make many more returns to $u - x_1$ in a state of S_1 avoiding level 0 (which are all returns before ruin time for the insurance process) and then change environment into S_2; this is what gives us the term $[I + \Xi(u - x_1, s/2)]\gamma^{-1}Q_{12}$; the factor $s/2$ is again due to the fact that in each of these return intervals, exactly half the time is spent by the fluid flow going upward. Finally, the final claim size should be, of course, $u - x_1 + x_2$ or correspondingly the fluid flow should spend precisely $(u - x_1 + x_2)/\gamma$ amount of time in S_2 which gives the remaining terms.

Theorem 6 *Assume that the risk process starts in state (u, i), $u \geq 0$, $i \in S_1$. Then the marginal distribution of the deficit at ruin is a phase type random variable with distribution $PH(e_i\Psi(0)e^{H(0)u}, \gamma^{-1}Q_{22})$, where e_i is a unit coordinate vector with 1 for its i-th component.*

Proof: Note that

$$\int_0^\infty e^{K(0)x_1}[\gamma^{-1}Q_{12}]e^{\gamma^{-1}Q_{22}x_1}\, dx_1 = \Psi(0).$$

Given that the process starts in level 0, the left side is an expression for the matrix of return probabilities to zero by conditioning on the value x_1 just prior to the last linear downward segment to level 0. Thus, setting $s = 0$ and integrating over x_1 in the joint density given for Case 1 results in the expression

$$\Psi(0)e^{\gamma^{-1}Q_{22}x_2}[-\gamma^{-1}Q_{22}]\mathbf{1},$$

since $Q_{21}\mathbf{1} = -Q_{22}\mathbf{1}$. Premultiplying the above by e_i shows this to be the density of the asserted PH distribution. Thus, the result is proved in the case $u = 0$.

The proof for the case $u > 0$ is similar and follows by noting that the distribution of the phase at the entrance to level 0 for the fluid model is given by $e_i\Psi(0)e^{H(0)u}$.

4 Concluding Remarks

A key result we obtained is the first passage time (Theorem 1) from level 0 to level x avoiding level 0 for the fluid flow $\mathcal{F}^+$, and this we did by a careful use of Markov renewal arguments (Lemma 1). This simple result gave a sufficient breakthrough to obtain several passage time distributions with taboo not only of 0 but also of certain levels $[a, \infty)$. Note that the probabilistic interpretation of the kernel $K(s)$ obtained by us earlier in [2] has been a key to all these.

For the insurance model itself, we introduced two simplifying facts: (a) We used a common size for all phase type distributions for a clean representation of the model; (b) we used a fluid flow construction related to it with all absolute rates of fluid changes being equal to a constant which substantially simplified translating fluid results to insurance risk model results.

The basic quantities we have to compute in analyzing the insurance risk model thus turn out to be simply the two busy period matrix transforms $\Psi(s)$ and $\Psi^r(s)$ for which we have powerful algorithms in [3]. Of course, in addition to these we need to compute the transforms $\Xi(x, s)$ and $\Xi^r(x, s)$ of Lemma 1.

As for the transform ${}^x\Psi(s)$, note that it is the LST of a nonnegative kernel over time t and thus very well behaved. Also once $\Psi(s)$ is computed there is no recursive computations here. Thus, despite the subtraction in the formula of Theorem 2, there is no reason to expect any numerical difficulties in its use in the formulae.

As noted earlier, this paper provides not only a clean, systematic approach to the insurance risk problem, but the passage times derived here

provide a means to analyze the transient behavior of many complex applied probability models of which the risk model is but a very simple example.

Appendix 1: Algorithm for $\Psi(\cdot)$

In this appendix, we briefly recall the algorithm developed in [3] for the computation of the matrix $\Psi(s)$ associated with the fluid flow $\mathcal{F} = \mathcal{F}(Q, c_1, -c_2)$. The computation of $\Psi^r(s)$ is achieved by applying the same algorithm to the flow $\mathcal{F}^r$.

We assume the conventions for the partitioning of the state space and various matrices as given in this paper. We also define the matrix $C = diag(C_1, C_2, I)$, where the identity matrix used is of order $|S_3|$.

For $\lambda > 0$, let

$$P_\lambda = \frac{1}{\lambda} C^{-1} Q + I.$$

Choose (fixed) positive numbers λ and δ such that

$$\begin{aligned} \lambda &\geq \max_{i \in S}\{-[C^{-1}Q]_{ii}\} \\ \max_{i \in S} \left[\frac{Re(s)}{\lambda} C^{-1}\right]_{ii} &\leq \delta < 1, \text{ and} \\ \max_{i \in S} \left[P_\lambda - \frac{Re(s)}{\lambda} C^{-1}\right]_{ii} &> 0. \end{aligned}$$

Define the matrices

$$A_2(s,\lambda) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \lambda C_2 (sI + 2\lambda C_2)^{-1} & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$A_1(s,\lambda) = \Lambda C (sI + \Lambda C)^{-1} \begin{pmatrix} 0 & 0 & 0 \\ \frac{1}{2}P_{21} & \frac{1}{2}P_{22} & \frac{1}{2}P_{23} \\ P_{31} & P_{32} & P_{33} \end{pmatrix},$$

$$A_0(s,\lambda) = \begin{pmatrix} P_{11} - \frac{s}{\lambda} C_1^{-1} & P_{12} & P_{13} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

where $\Lambda = diag(\lambda I, 2\lambda I, \lambda I)$ and $P = P_\lambda$.

Consider now the following algorithm.

Algorithm

Fix $\epsilon > 0$ and set diff $= 100$;

$H^{**}(1,s,\lambda) = (I - A_1(s,\lambda))^{-1} A_0(s,\lambda)$;
$L^{**}(1,s,\lambda) = (I - A_1(s,\lambda))^{-1} A_2(s,\lambda)$;
$G^{**}(1,s,\lambda) = L^{**}(1,s,\lambda)$;
$T(1) = H^{**}(1,s,\lambda)$;
Do while (diff $> \epsilon$)
$k = k+1$;
$U^{**}(k,s,\lambda) = H^{**}(k-1,s,\lambda)L^{**}(k-1,s,\lambda)$
$+ L^{**}(k-1,s,\lambda)H^{**}(k-1,s,\lambda)$;
$M = (H^{**}(k-1,s,\lambda))^2$;
$H^{**}(k,s,\lambda) = (I - U^{**}(k,s,\lambda))^{-1} M$;
$M = (L^{**}(k-1,s,\lambda))^2$;
$L^{**}(k,s,\lambda) = (I - U^{**}(k,s,\lambda))^{-1} M$;
$G^{**}(k,s,\lambda) = G^{**}(k-1,s,\lambda) + T(k-1)L^{**}(k,s,\lambda)$;
$T(k) = T(k-1)H^{**}(k,s,\lambda)$;
diff $= \max_{j,k \in S} \mid [G^{**}(k,s,\lambda)]_{j,k} - [G^{**}(k-1,s,\lambda)]_{j,k} \mid$;
end

$\Psi(s) \approxeq G_{12}^{**}(k,s,\lambda)[G_{22}^{**}(k,s,\lambda)]^{-1}$.

We have established in [3] that at its termination, the above algorithm yields the matrix $\Psi(s)$ with error at most ϵ in its entries, and furthermore that the error in the k-th iterate is $O[\{\eta(s)\}^{2^k}]$ for a constant $0 < \eta(s) < 1$ so much so that the iterates converge quadratically to the required limit. For details, refer to [3].

Appendix 2: Phase Type Distributions

In this appendix, we briefly review PH distributions and show how one may use a common order for the phase type distributions arising in the general risk model.

The distribution $PH(\alpha, T)$ is the distribution of the absorption time in the CTMC (whose states are called phases) with initial probability vector $(\alpha, 1 - \alpha\mathbf{1})$ and infinitesimal generator

$$\begin{pmatrix} T & -T\mathbf{1} \\ 0 & 0 \end{pmatrix}.$$

We assume that T^{-1} exists which is equivalent to assuming that starting in any of the phases absorption into the last phase does indeed occur a.s. The order of the matrix T, which is the number of transient phases, is also called

the order of the PH-distribution. Also, in most applications, $\alpha\mathbf{1} = 1$ so that the distribution has no atom at zero.

Phase type distributions include as special cases the exponential distributions and mixtures and convolutions of finitely many of them. Such distributions are known to be dense in the set of all distributions on $[0, \infty)$.

For details concerning PH distributions, we refer the reader to [14] and [12] restricting ourselves here to showing how given two PH-distributions, we may, without loss of generality assume that they both have the same order. To this end, it is enough to show that given a PH-distribution of order m, there is a representation of that also as a PH-distribution of order $m + n$.

To this end, suppose we are given $PH(\alpha, T)$ of order m. Take any $PH(\beta, S)$ of order n, and consider the PH-distribution $PH(\delta, U)$ of order $m + n$ defined by

$$\delta = (\alpha, 0_n),$$

where 0_n is an n-vector of zeros, and U is the infinitesimal generator given by

$$U = \begin{pmatrix} T & 0_{mn} \\ -S\mathbf{1}\beta & S \end{pmatrix}$$

Clearly, $PH(\delta, U) = PH(\alpha, T)$, since the states $m + 1, \cdots, m + n$ are never visited by the CTMC governing the former. (Note: This is not the unique way to convert a PH distribution to one of a higher order.)

Acknowledgment: I am thankful to Andrei Badescu and Soohan Ahn for a careful reading of the manuscript and for their comments.

References

[1] Ahn, S & Ramaswami, V (2003). Fluid flow models & queues - A connection by stochastic coupling. *Stochastic Models,* **19(3)**, 325–348.

[2] Ahn, S. & Ramaswami, V. (2004). Transient analysis of fluid flow models via stochastic coupling to a queue. *Stochastic Models,* **20(1)**, 71-101.

[3] Ahn, S. & Ramaswami, V. (2005). Efficient algorithms for transient analysis of stochastic fluid flow models. *J. Appl. Prob.*, to appear.

[4] Ahn, S. & Ramaswami, V. (2004). Transient analysis of fluid models via elementary level crossing arguments. *Submitted.*

[5] Ahn, S., Jeon, J. & Ramaswami, V. (2005). Steady state analysis of finite fluid flow models using finite QBDs. *QUESTA*, to appear.

[6] Asmussen, S & Koole, G (1993) Marked point processes as limits of Markovian arrival streams, *J. Appl. Prob.,* 30, 365-372.

[7] Asmussen, S (1994). Busy period analysis, rare events and transient behavior in fluid flow models. *J. Appl. Math. and Stoch. Anal.*, **7(3)**, 269-299.

[8] Asmussen, S (2004). *Applied Probability & Queues*, Second Edition, John Wiley, NY.

[9] Badescu, A, Breuer, L, Soares da Silva, Latouche, G, Remich, M.A. & Stanford, D (2004) Risk processes analyzed as fluid queues. *Scand. Actuarial J.*, to appear.

[10] Graham, A (1981). *Kronecker Products and Matrix Calculus with Applications.* John Wiley & Sons.

[11] Latouche, G & Ramaswami, V. (1993). A logarithmic reduction algorithm for Quasi-Birth-and-Death processes, *J. Appl. Prob.,* **30**, 650-674.

[12] Latouche, G & Ramaswami, V. (1999). *Introduction to Matrix Analytic Methods in Stochastic Modeling.* SIAM & ASA, Philadelphia.

[13] Neuts, M.F. (1979). A versatile Markovian point process. *J. Appl. Prob.,* 16, 764-779.

[14] Neuts, M.F. (1981). *Matrix-Geometric Solutions in Stochastic Models – An Algorithmic Approach.* The Johns Hopkins University Press, Baltimore, MD.

[15] Neuts, M.F. (1989). *Structured stochastic matrices of M/G/1 type and their applications.* Marcel Dekker, NY.

[16] Ramaswami, V. (1999). Matrix analytic methods for stochastic fluid flows. in *Teletraffic Engineering in a Competitive World - Proc. of the 16th International Teletraffic Congress.* D. Smith and P. Key (eds.), 1019–1030, Elsevier, NY.

[17] Scheinhardt, W. (1998) *Markov-modulated and feedback fluid queues*, Thesis, University of Twente, Enscheide, The Netherlands.