Søren Asmussen

On Ruin Problems and Queues of Markov-Modulated M/G/1 Type

Preprint
October
1986
4

Institute of Mathematical Statistics
University of Copenhagen

Søren Asmussen

ON RUIN PROBLEMS AND QUEUES OF MARKOV-MODULATED M/G/1 TYPE

Preprint 1986 No. 4

INSTITUTE OF MATHEMATICAL STATISTICS
UNIVERSITY OF COPENHAGEN

October 1986

On Ruin Problems and Queues of Markov-Modulated M/G/1 Type

By Søren Asmussen, University of Copenhagen

Abstract

Exact solutions and approximations are derived for risk processes / queues where the arrival intensity and the distribution of claim sizes / service times at time t depend on the state Z_t of a underlying finite Markov jump process. The main mathematical tool is random walks on Markov chains, and in particular Wiener-Hopf factorisation problems and conjugate distributions (Esscher transforms) are involved.

Contents

1. Introduction

Ruin probabilities in risk theory and waiting time distributions for queues reduce in some basic cases to just the same random walk first passage time probabilities and are then conveniently studied within the same framework. In particular, this is so for compound Poisson risk processes with unit premium rate and M/G/1 queues where

$$\psi(u,T) = \mathbb{P}(V_t > u \mid V_0 = 0). \tag{1.1}$$

Here $\psi(u,T)$ is the probability of ruin before time T with initial risk reserve u, and V_t is the virtual waiting time at time t. Nevertheless, when formulating more realistic models or more specific questions, risk theory and queueing theory (even in the M/G/1 setting) may of course lead to different problems. For example, premium rates which depend on the current risk reserve are of main interest in risk theory but the ruin probabilities do not correspond to any reasonable queueing model. Similarly, the study of say other queue disciplines than the FIFO one comes up in a number of queueing applications but can hardly be given a risk theoretic interpretation.

The present paper is concerned with a particular type of generalisation which, however, seems equally well motivated from the point of view of risk theory and queues. This is *Markov-modulation* : the rate β of the Poisson arrival process $\{N_t\}_{t \geq 0}$ and the distribution B of the claim sizes / service times U_1, U_2, ... are not fixed in time but depend on the state of a underlying Markov jump process $\{Z_t\}_{t \geq 0}$ such that $\beta = \beta_i$ and $B = B_i$ when Z_t = i. A sample path of the corresponding risk process

$$R_t \quad = \sum_{n=1}^{N_t} U_t - t \tag{1.2}$$

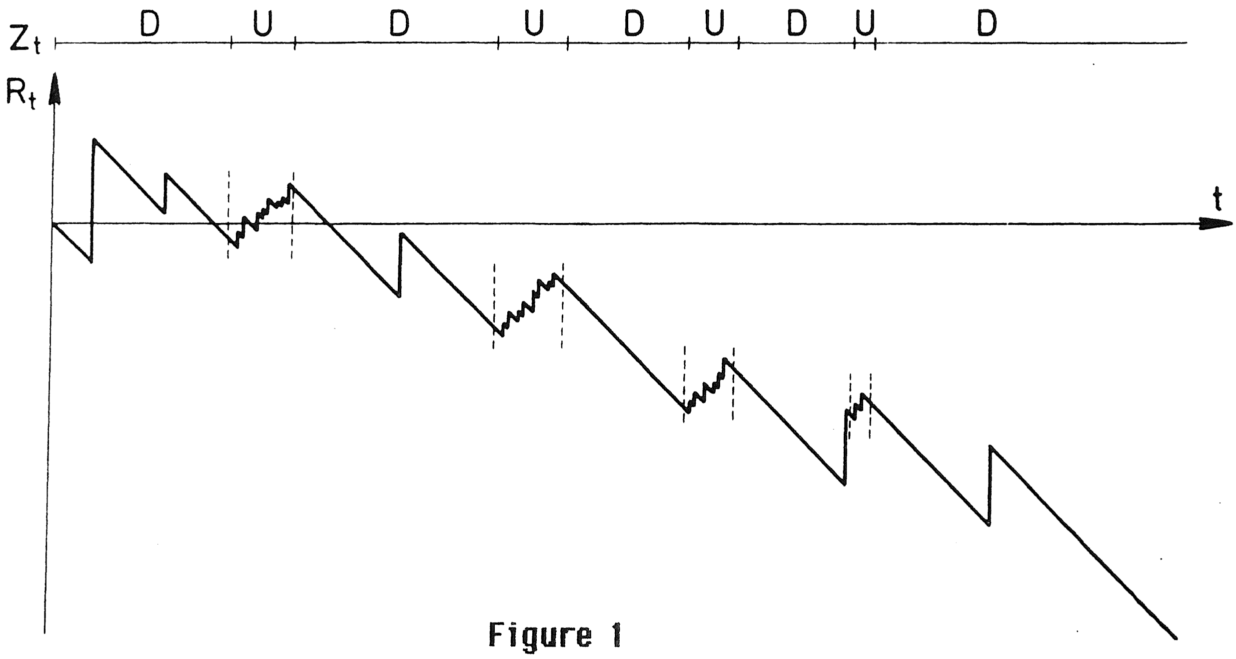


Figure 1

is depicted in Fig. 1, corresponding to two states U,D (Up and Down) such that the process has many but small claims and an upwards drift in the Up state, and rare but large claims and a downwards drift in the Down state (the overall average drift is negative corresponding to a positive safety loading). In particular, the arrival process is more bursty than the Poisson process in the sense that periods with very frequent arrivals and periods with very few arrivals alternate. In health insurance, sojourns of $\{Z_t\}$ in certain states could correspond to certain types of epidemics, and in automobile insurance, Z_t could be the weather type at time t. The Markov property may admittedly be questionnable in some cases, but at least it should be noted that it does not require exponential distribution of say periods of cold weather: using phase-type representations (see e.g. Section III.6 of Asmussen [9], henceforth referred to as APQ) any given distribution may be approximated arbitrarily close. Anyway, the possibility of allowing the parameters of the process to vary in time seems a major step towards more realistic models and better motivated than many other extensions like renewal arrival processes. Furthermore it leads, as we shall see, to mathematical problems which are tractable or at least ameneable to numerical computations.

The relevance of Markov-modulation may have been noticed in risk theory, but to our knowledge no substantial mathematical results have been produced within the setting of compound Poisson risk processes. For general random walks some results on ruin problems can, however, be found in [25], [4], [16], and also for queues, the subject has received some recent attention, see Regterschot and de Smit [35], [40] and references there. Briefly speaking, the state is that algorithms exist which provide numerical values of expected waiting times $\mathbb{E}W_\infty$ and queue lengths $\mathbb{E}Q_\infty$, of $\mathbb{V}\mathrm{ar}W_\infty$, $\mathbb{V}\mathrm{ar}Q_\infty$ and so on in the steady state ($T = \infty$). Just as in the one-dimensional case it seems, however, somewhat more difficult to obtain the waiting time probabilities $\mathbb{P}(W_\infty > u)$ themselves and also the time-dependent solutions can hardly be evaluated at all (this is a problem even in the standard M/M/1 case, cf. [36], [7], and the methods of these papers do certainly not apply in the present generality). In risk theory, these quantities are, however, the ones of main importance in view of an extension of (1.1) to be given in Section 7, and one main purpose of the present paper is to look into approximations (a main recent reference in this area is Höglund [16] which, however, does not seem to substantially overlap with the present paper).

The paper is organised as follows. We start in Section 2 by a brief look at Wiener-Hopf factorisation identities for general random walks on Markov chains. The literature on this subject is extensive ([33], [1], [2], [3], [4], [6], [39]) but also somewhat bewildering, and the results are not always easy to compare neither mutually nor with the standard random walk case. The formulas presented here are close analogues of those of Feller [13] Ch. XII, formulating the results in terms of measures rather than transforms or operators, and also even for the standard case the proof may be slighty easier than the usual ones ([13], APQ VII.3). As example, some apparently new remarks on a Wiener-Hopf interpretation of the rate matrix in matrix-geometric models ([27]) are given as well as some applications related to [34]. In Section 3, we specialise to the M/G/1 setting. Parts of the material is parallel to but also simpler than [6], [35] though a direct comparison is not straightforward. A main new idea here is the introduction of a uniformisation (randomisation) procedure which substantially simplifies the proofs as well as the form of the results. The material is of later relevance in connection with algorithmic solutions of the queueing problems (Sections 7,8) as well as the constants in the approximations in Sections 5,6 involve the Wiener-Hopf factors. Section 4 contains some auxiliary material on moments and conjugate distributions (Esscher transforms), extending and simplifying [18], [20] (see also [15]). In Section 5, we derive computable expressions for the constants of the Cramér-Lundberg approximation earlier obtained in [25], [4] as well as we show the finite horizon version (first derived by Segerdahl [37] in the standard case) and note some versions of Lundberg's inequality. Section 6 then contains some of the main results of the paper, diffusion approximations with correction terms for the finite as well as the infinite horizon case. The approximations are of the same form as in Siegmund [38], Asmussen [7], and were documented in these papers to have an outstanding fit at least in the standard case. For example, the relative error is typically $< 0.1\%$ when the safety loading is $> 18\%$ (in queueing terms, when the traffic intensity ρ is > 0.85). In Section 7 we then give some of the relevant translation to queues, and finally Section 8 contains some remarks on GI/M/1 and M/M/1 models.

Though the paper does not contain numerical illustrations, the point of view is nevertheless largely algorithmical. The aim is to present the results (exact solutions and approximations) in a form which is ready for numerical implementation on a computer. From a computational point of view, some of the main ingredients are

a) evaluation of complex moment generating functions $\hat{B}(\theta) = \int e^{\theta x} B(dx)$, $\theta \in \mathbb{C}$, and their (real) derivatives $\hat{B}^{(k)}(\theta) = \int x^k e^{\theta x} B(dx)$, $\theta \in \mathbb{R}$.
b) rootfinding in the complex plane, i.e. the solution of equations of the form $g(\theta) = 0$, $\theta \in \mathbb{C}$.
c) matrix manipulation: determinants, inverses, eigenvalues, eigenvectors.

Here standard routines are available for c), and in presumably most examples closed expressions can be found in a). Thus the main difficulties seem to be inherent in step b). One should note here, however, that some relevant software has been developed by Regterschot and de Smit in connection with [35], and that (private communication) the rootfinding is not considered a major obstacle.

2. Wiener-Hopf factorisation for the general random walk case

Consider a random walk $\{S_n\}$ on a Markov chain $\{J_n\}$ (or Markov-modulated random walk, cf. APQ X.4). Assuming that $\{J_n\}$ has a finite state space E (say with p elements) and letting $X_n = S_n - S_{n-1}$, this means that $\{(J_n, X_n)\}$ is a Markov chain on $E \times \mathbb{R}$ with transition function depending only on the first coordinate . Thus the process is completely specified by the measures $F(i,j) = F(i,j;\cdot) = F(i,j;dx)$ given by

$$F(i,j;A) = \mathbb{P}_i(J_1 = j, X_1 \in A).$$

and by the initial conditions (we consider only the case $S_0 = X_0 = 0$ and let $\mathbb{P}_i$ correspond to $J_0 = i$). We use notation like $\mathbf{F}$ for the matrix which has the measure $F(i,j)$ as its ij^{th} element, $\mathbf{F}*\mathbf{G}$ for the matrix with ij^{th} element $\Sigma_{k \in E} F(i,k)*G(k,j)$ and $\mathbf{F}^{*n}$ for the n^{th} convolution power of $\mathbf{F}$ (we identify $\mathbf{F}^{*0}$ with the identity matrix $\mathbf{I}$). The total mass of a measure H is denoted by $\|H\|$, and $\|\mathbf{F}\|$ stands for the matrix $(\|F(i,j)\|)_{i,j \in E}$. Thus $\|\mathbf{F}\|$ reduces to the transition matrix $\mathbf{P} = (p_{ij})$ for $\{J_n\}$ which we assume irreducible. In particular, a stationary distribution $\pi = (\pi(j))_{j \in E}$ exists. Also $\tilde{}$ refers to the time-reversed (or *dual*) process $\{(\tilde{J}_n, \tilde{X}_n)\}$ given by the transition function

$$\tilde{F}(i,j;A) = \mathbb{P}_\pi(J_0 = j, X_1 \in A \mid J_1 = i) = \pi(j)\, F(j,i;A) / \pi(i).$$

Note that this corresponds to the usual time-reversed transition probabilities $\tilde{p}_{ij} = \pi(j) p_{ji} / \pi(i)$ when looking at $\{\tilde{J}_n\}$ alone whereas $\tilde{F}(i,j)/\|\tilde{F}(i,j)\|$, the conditional distribution of $\tilde{X}_1$ given $\tilde{J}_0 = i$, $\tilde{J}_1 = j$, is the same as $F(j,i)/\|F(j,i)\|$, the conditional distribution of X_1 given $J_0 = j$, $J_1 = i$. Let finally

$$\tau_+ = \inf\{n \geq 1 : S_n > 0\}, \quad G_+(i,j;A) = \mathbb{P}_i(S_{\tau_+} \in A, J_{\tau_+} = j, \tau_+ < \infty),$$
$$\tau_- = \inf\{n \geq 1 : S_n \leq 0\}, \quad G_-(i,j;A) = \mathbb{P}_i(S_{\tau_-} \in A, J_{\tau_-} = j, \tau_- < \infty).$$

Our goal is to obtain analogues of the formula $F = G_+ + G_- - G_- * G_+$ used as the basic Wiener-Hopf identity in [13] Ch. XII (this point of view is also followed in APQ).

To this end, define

$$R_+(i,j;A) = \mathbb{E}_i \sum_{n=0}^{\tau_+-1} I(J_n = j,\, S_n \in A),$$

$$G_\ominus(i,j) = \pi(j)\widetilde{G}_-(j,i)/\pi(i), \quad \mathbf{U}_\ominus = \sum_{n=0}^{\infty} \mathbf{G}_\ominus^{*n}, \quad \mathbf{U}_+ = \sum_{n=0}^{\infty} \mathbf{G}_+^{*n}.$$

Thus $\mathbf{R}_+$ is the pre-τ_+-occupation measure and $\mathbf{U}_\ominus$ is the renewal measure corresponding to $\mathbf{G}_\ominus$.

Proposition 2.1 $\mathbf{R}_+ = \mathbf{U}_\ominus$

Proof Let n be fixed and write $i = i_0$, $j = i_n$. Then

$$\mathbb{P}_i(\tau_+ > n, J_n = j, S_n \in A) =$$

$$\sum_{i_1 \ldots i_{n-1}} p_{i_0 i_1} \cdots p_{i_{n-1} i_n} \mathbb{P}(S_k \le 0, k \le n, S_n \in A \mid J_0 = i_0, J_1 = i_1, \ldots, J_n = i_n) =$$

$$\sum_{i_1 \ldots i_{n-1}} p_{i_n i_{n-1}} \cdots p_{i_1 i_0} \pi(i_n)\ \mathbb{P}(\widetilde{S}_n \le \widetilde{S}_k, k \le n, \widetilde{S}_n \in A \mid \widetilde{J}_0 = i_n, \widetilde{J}_1 = i_{n-1}, \ldots, \widetilde{J}_n = i_0)/\pi(i_0) =$$

$$\pi(j)\, \mathbb{P}_j(\widetilde{J}_n = i, \widetilde{S}_n \in A; \widetilde{F}_n)/\pi(i)$$

where F_n is the event that n is a descending ladder epoch for $\{S_n\}$. Summing over n, we get

$$R_+(i,j;A) = \pi(j) \sum_{n=0}^{\infty} \widetilde{G}_-^{*k}(j,i;A) / \pi(i)$$

and since it easily follows by induction that $G_\ominus^{*k}(i,j) = \pi(j)\, \widetilde{G}_\ominus^{*k}(j,i) / \pi(i)$, the proof is complete. □

Lemma 2.1 $\mathbf{R}_+ + \mathbf{G}_+ = \mathbf{I} + \mathbf{R}_+ * \mathbf{F}$

Proof This is just a special case of Prop. 3.2 of Pitman [31], but for the sake of completeness we reproduce the proof. Integrating the identity

$$\sum_{n=0}^{\tau_+-1} I(J_n = j, S_n \in A) + I(J_{\tau_+} = j, S_{\tau_+} \in A) =$$

$$I(J_0 = j, S_0 \in A) + \sum_{n=0}^{\tau_+-1} I(J_{n+1} = j, S_{n+1} \in A)$$

w.r.t. $\mathbb{P}_i$, the first three terms become $R_+(i,j)$, $G_+(i,j)$, $I(i,j)$, and an easy conditioning argument shows that the $\mathbb{P}_i$-expectation of the last term is $(R_+*F)(i,j;A)$. □

Theorem 2.1 $F = G_\ominus + G_+ - G_\ominus * G_+$. *Equivalently,*

$$I - F = (I - G_\ominus)*(I - G_+) \tag{2.1}$$

Proof By Prop.2.1 and Lemma 2.1,

$$U_\ominus + G_+ = I + U_\ominus * F.$$

Convolving with $G_\ominus$ to the left we get

$$U_\ominus - I + G_\ominus * G_+ = G_\ominus + U_\ominus * F - F,$$

and subtracting, the result follows. □

Now define

$$\mu = \mathbb{E}_\pi X_1 = \sum_{i,j \in E} \pi(i) \int_{-\infty}^{\infty} x\, F(i,j;dx)$$

Then:

Lemma 2.2 (a) $S_n/n \to \mu$ a.s.;

(b) *If* $\mu<0$ *then* $\|G_+\|$ *is substochastic* ($\mathrm{spr}(\|G_+\|) < 1$) *whereas* $\mathrm{spr}(\|G_\ominus\|) = 1$ *and* $\pi\|G_\ominus\| = \pi$;

(c) *If* $\mu>0$ *then* $\mathrm{spr}(\|G_\ominus\|) < 1$ *whereas* $\|G_+\|$ *is stochastic* ($\mathrm{spr}(\|G_+\|) = 1$) *with* $\pi_+ = \pi(I - \|G_\ominus\|)$ *as positive left eigenvector;*

(d) *If* $\mu=0$ *then* $\mathrm{spr}(\|G_+\|) = \mathrm{spr}(\|G_\ominus\|) = 1$.

Proof (a) is well-known and easily proved (APQ X.4). Similarly in (b) it follows from $S_n \to \infty$ just as in the one-dimensional case that $\mathrm{spr}(\|G_+\|) < 1 = \mathrm{spr}(\|\tilde{G}_-\|)$. The way G_- is constructed from G_- then ensures that also $\mathrm{spr}(\|G_\ominus\|) = 1$, and $\pi\|G_\ominus\| = \pi$ follows from

$$\sum_{i \in E} \pi(i) G_\ominus(i,j) = \pi(j) \sum_{i \in E} G_-(j,i)) = \pi(j).$$

Also (d) and the first part of (c) is similar as (b). For the last claim in (c), note that $\mathrm{spr}(\|G_\ominus\|) < 1$ implies that $\pi_+ \neq 0$. Also Theorem 2.1 yields

$$I - \|F\| = (I - \|G_\ominus\|)(I - \|G_+\|)$$

and multiplying by $\pi = \pi \|F\|$ to the left we get $\pi_+(\mathbf{I} - \|\mathbf{G}_+\|) = 0$. □

An interesting interpretation of the Wiener-Hopf factorisation can be given for Markovs chains of the GI/M/1 type having a matrix-geometric stationary distribution ([27] or APQ X.4-5). Here one is interested in the occupation measure $\mathbf{R}_-$ (defined the obvious way) in the case of a Markov-modulated right-continuous random walk given by matrices $\mathbf{F}(1)$, $\mathbf{F}(0)$, $\mathbf{F}(-1)$, ... with elements $F(i,j;l) = \mathbb{P}_i(X_1 = l, J_1 = j)$ (the state $l \in \mathbb{Z}$ of S_n is denoted as the *level* and the state $j \in E$ of J_n as the *phase*). It is well-known and easy to see that $\mathbf{R}_-$ has matrix-geometric form, i.e. the restriction of $\mathbf{R}_-$ to level $k>1$ is $\mathbf{R}^k$ where $\mathbf{R}$ is the restriction of $\mathbf{R}_-$ to level $k=1$. To interpret $\mathbf{R}$, we need the variant

$$\mathbf{I} - \mathbf{F} = (\mathbf{I} - \mathbf{G}_{\oplus})*(\mathbf{I} - \mathbf{G}_-) \tag{2.2}$$

of (2.1) which follows by simply interchanging the role of τ_+ and τ_- in the proof. The similar variant of Prop. 2.1 states that $\mathbf{U}_{\oplus} = \mathbf{R}_-$. Taking the restriction to level 1 and noting that $\mathbf{G}_{\oplus}$ is concentrated at level 1, we get $\mathbf{G}_{\oplus} = \mathbf{R}$. That is:

Corollary 2.1 *The rate matrix* $\mathbf{R}$ *in Markov chains having a matrix-geometric stationary distribution is related to the ascending ladder height distribution* $\widetilde{\mathbf{G}}_+$ *of the time-reversed Markov-modulated random walk by means of* $\mathbf{R} = \mathbf{G}_{\oplus}$. *That is,* $R(i,j) = \pi(j)\widetilde{G}_+(j,i)/\pi(j)$ *and* (by the generating function version of (2.2))

$$\mathbf{I} - s\mathbf{F}(1) - \mathbf{F}(0) - s^{-1}\mathbf{F}(-1) - \ldots = (\mathbf{I} - s\mathbf{R})(\mathbf{I} - \mathbf{G}_-(0) - s^{-1}\mathbf{G}_-(-1) - \ldots) \tag{2.3}$$

where $\mathbf{G}_-(k)$ *is the matrix with* ij^{th} *element* $G_-(i,j;\{k\})$.

To see that the Wiener-Hopf interpretation of $\mathbf{R}$ is more than a curiosity, we shall give a short and transparent proof of the results of Ramaswami and Latouche [34], covering the known cases where the rate matrix $\mathbf{R}$ can be found explicitly. Here $\mathbf{F}(k) = \mathbf{0}$, $k=-2,-3, \ldots$, and equating coefficients in (2.3) we get $\mathbf{G}_-(-1) = \mathbf{F}(-1)$,

$$\mathbf{I} - \mathbf{F}(0) = \mathbf{I} - \mathbf{G}_-(0) + \mathbf{R}\mathbf{G}_-(-1) = \mathbf{I} - \mathbf{G}_-(0) + \mathbf{R}\mathbf{F}(-1),$$

$$\mathbf{F}(1) = \mathbf{R}(\mathbf{I} - \mathbf{G}_-(0)) = \mathbf{R}(\mathbf{I} - \mathbf{F}(0) - \mathbf{R}\mathbf{F}(-1)) \tag{2.4}$$

Case 1° $F(1) = w\beta$ *where* w *is a column vector and* β *a row vector satisfying* $\beta e = 1$ (here e is the column vector with all components equal to one). This means that an upwards jump from phase i occurs w.p. $w(i)$ and that the new phase then is chosen according to β. The occupation measure interpretation of $\mathbf{R}$ therefore shows that $R(i,j) = w(i)\xi'(j)$ for some row vector ξ'. Normalising such that $\xi' = \eta\xi$ where $\eta > 0$ is the spectral radius of $\mathbf{R}$ and $\xi w = 1$, we get $\mathbf{R}^2 = \eta\mathbf{R}$ and hence by (2.4)

$$\mathbf{R} = \mathbf{F}(1)(\mathbf{I} - \mathbf{F}(0) - \eta\mathbf{F}(-1))^{-1} \tag{2.5}$$

which is the desired explicit formula for $\mathbf{R}$ (given that η has been computed which is possible without knowing $\mathbf{R}$, cf. [27] or [34]).

Case 2° $F(-1) = v\alpha$ *where* v *is a column vector and* α *a row vector satisfying* $\alpha e = 1$. This means that the phase after a downwards jump always has distribution α. Let

$$N(1,2) = \sum_{n=0}^{\tau_- - 1} I(S_n = 1, S_{n+1} = 2)$$

be the number of upcrossings from 1 to 2 before τ_- and $N(2,1)$ the similar number of downcrossings. Then $N(1,2) = N(2,1)$ and hence

$$\mathbb{E}_i(N(1,2); J_\tau = j) = \mathbb{E}_i(N(2,1); J_\tau = j) \tag{2.6}$$

Recalling that $R^1(i,k)$ is the expected number of sojourns in phase k and level 1 before τ_- given $J_0 = i$, (2.6) can be rewritten as

$$\sum_{k,k' \in E} R(i,k)F(k,k';1)\alpha(j) = \sum_{k,k' \in E} R^2(i,k)F(k,k';-1)\alpha(j).$$

I.e., $\mathbf{RF}(1)e\alpha = \mathbf{R}^2\mathbf{F}(-1)e\alpha = \mathbf{R}^2\mathbf{F}(-1)$ and hence by (2.4)

$$\mathbf{R} = \mathbf{F}(1)(\mathbf{I} - \mathbf{F}(0) - \mathbf{F}(1)e\alpha)^{-1} \tag{2.7}$$

giving an explicit formula of similar form as (2.5) for $\mathbf{R}$.

3. Wiener-Hopf factorisation for the M/G/1 case

Recalling the basic set-up and notation in Section 1, and in particular the definition (1.2) of the risk process $\{R_t\}$, we let

$$M_T = \sup_{0 \le t \le T} R_t, \quad M = \sup_{0 \le t < \infty} R_t, \quad \tau(u) = \inf\{t \ge 0 : R_t > u\},$$

$$\psi_{ij}(u,T) = \mathbb{P}_i(\tau(u) \le T, J_{\tau(u)} = j), \quad \psi_{ij}(u) = \mathbb{P}_i(\tau(u) < \infty, J_{\tau(u)} = j).$$

Then $\Sigma_j\, \psi_{ij}(u,T)$ is the probability of ruin before time T which may alternatively be expressed as $\mathbb{P}_i(M_T > u)$, and $\Sigma_j\, \psi_{ij}(u) = \mathbb{P}_i(M > u) = \mathbb{P}_i(\tau(u) < \infty)$ is the probability of ultimate ruin. Further we assume that the state space E of $\{Z_t\}$ is finite, say with p elements, and that $\{Z_t\}$ is ergodic (irreducibility suffices for this). Then a limiting stationary distribution π exists and the average drift of $\{R_t\}$ is

$$\mathbb{E}_\pi \frac{R_t}{t} = \rho - 1 \qquad \text{where } \rho = \sum_{i \in E} \pi(i)\beta_i \mathbb{E}_i U.$$

In queueing terms, ρ is the traffic intensity, cf. [35] and in risk theory, $\rho^{-1}-1$ is the safety loading. We assume throughout that $\rho < 1$.

For tecnical purposes, it now turns out to be convenient to introduce uniformisation, cf. e.g. [17]. To this end, let $\mathbf{\Lambda} = (\lambda_{ij})$ be the intensity matrix for $\{Z_t\}$ and choose an η satisfying $\eta > \beta_i - \lambda_{ii}$ for all i and a Poisson process $\{N_t^*\}$ with intensity η. We then construct $\{Z_t\}$, $\{N_t\}$ the following way: if $\{N_t^*\}$ has an arrival at a given time t where $Z_t = i$, then a coin is tossed to give an arrival for $\{N_t\}$ w.p. β_i/η, a jump of Z_t to state j w.p. λ_{ij}/η and a dummy event w.p. $(\eta+\lambda_{ii}-\beta_i)/\eta$. We let $\sigma(n)$ be the n^{th} arrival epoch for $\{N_t^*\}$ and $\sigma(0) = 0$, $S_n = R_{\sigma(n)}$, $X_n = S_n - S_{n-1}$, $J_n = Z_{\sigma(n)}$. By general results on uniformisation, the stationary distribution for $\{J_n\}$ and $\{Z_t\}$ are the same, viz. π, and we have a Markov-modulated random walk as in Section 2 with

$$\mu = \sum_{i \in E} \pi_i \left(\beta_i/\eta \int_{-\infty}^{\infty} x\, B_i(dx) - 1/\eta\right) = (\rho-1)/\eta < 0.$$

Note that the traditional way of imbedding a random walk $\{S_n^*\}$ in the risk process corresponds to observing $\{R_t\}$ at the times of claims (loosely speaking, we have added some dummy claims of size zero). However, since $\{R_t\}$ can increase only at the times of claims, $\{R_t\}$, $\{S_n\}$ and $\{S_n\}$ have the same maximum and the same ascending ladder height distribution $\mathbf{G}_+$. Since these quantities are what is important for the ruin problem, and $\{S_n^*\}$ is an auxiliary quantity rather than of intrinsic interest, one may therefore as well work with $\{S_n\}$. The gain is that the descending ladder height distribution has a particular simple form.

To see this, the basic observation is that the X_n (and hence also the X_n) are of the form U_n-T_n where the T_n are exponential with intensity η. This property being preserved by time-reversal, it follows that $G\_(i,j)/\|G\_(i,j)\|$ is the same distribution as that of the $-T_n$ (no matter i,j) so that for $\mathbf{G}\_$ it only remains to evaluate the matrix $\mathbf{Q} = \|\mathbf{G}_\ominus\|$ with elements $q_{ij} = \|G_\ominus(i,j)\|$. Let $\hat{\mathbf{F}}(\theta)$ be the matrix with elements

$$F(i,j;\theta) = \int_{\infty}^{\infty} e^{\theta x} F(i,j;dx) = \mathbb{E}_i e^{-\theta T}\, \mathbb{E}_i[e^{\theta U}; J_1=j]$$

let $\mathbf{H}(\theta) = (\eta+\theta)(\mathbf{I}-\hat{\mathbf{F}}(\theta))$, $\mathbf{H}_+ = \mathbf{I}-\hat{\mathbf{G}}_+$ and

$$\mathbf{H}\_(\theta) = (\eta+\theta)(\mathbf{I} - \hat{\mathbf{G}}_\ominus(\theta)) = (\eta+\theta)\mathbf{I} - \eta\mathbf{Q} \tag{3.1}$$

Note that (2.1) implies

$$\mathbf{I}-\hat{\mathbf{F}} = (\mathbf{I}-\hat{\mathbf{G}}_\ominus)(\mathbf{I}-\hat{\mathbf{G}}_+) \tag{3.2}$$

and hence

$$\mathbf{H}(\theta) = \mathbf{H}\_(\theta)\, \mathbf{H}_+(\theta) \tag{3.3}$$

for $0 \geq \mathrm{Re}\,\theta > -\eta$ (the truth of (3.3) for all θ with $\mathrm{Re}\,\theta \leq 0$ then follows by analytic continuation, using the explicit forms of $\mathbf{H}\_,\mathbf{H}$ given in (3.1) and Lemma 3.1 below). We shall need:

Condition 3.1 *There exist* p-1 *distinct solutions* $\lambda_2, \dots \lambda_p$ *with* $\mathrm{Re}\,\lambda_i < 0$ *to the equation* $\det \mathbf{H}(\theta) = 0$.

Note that when $p=1$, then $H(\theta) = 0$ is simply the usual Lundberg equation (however, the solution γ occuring in the Cramer-Lundberg approximation and Lundberg's inequality has $\gamma>0$). We discuss Condition 3.1 somewhat further below and proceed to state and prove the main result.

Theorem 3.1 *Suppose that Condition 3.1 holds and that* $\rho\leq 1$, *and let* $\pi^{(i)}$ *be a non-zero row vector with* $\pi^{(i)}\mathbf{H}(\lambda_i) = 0$, $i=2,\dots p$, $\pi^{(1)} = \pi$, $\lambda_1=0$. *Then* $\mathbf{Q} = \mathbf{I} + \mathbf{Q}_0/\eta$ *where*

$$\mathbf{Q}_0 = \mathbf{\Pi}^{-1}\begin{pmatrix} \lambda_1 \pi^{(1)} \\ \lambda_2 \pi^{(2)} \\ \\ \lambda_p \pi^{(p)} \end{pmatrix}, \quad \mathbf{\Pi} = \begin{pmatrix} \pi^{(1)} \\ \pi^{(2)} \\ \\ \pi^{(p)} \end{pmatrix}$$

Proof We first note that the elements of $\mathbf{G}_+(\lambda_i)$ are effectively smaller than those of $\|\mathbf{G}_+\|$ when Re $\lambda_i < 0$. Since this last matrix is substochastic when $\rho<1$ and stochastic when $\rho=1$, $\hat{\mathbf{G}}_+(\lambda_i)$ must be substochastic which implies that $\mathbf{I} - \hat{\mathbf{G}}_+(\lambda_i)$ is non-singular. Hence by (3.3) $\pi^{(i)}\mathbf{H}(\lambda_i) = 0$ implies $\pi^{(i)}\mathbf{H}_-(\lambda_i) = 0$ which in terms of $\mathbf{Q}$ means that $\pi^{(i)}\mathbf{Q} = \omega_i\pi^{(i)}$ where $\omega_i = 1+\lambda_i/\eta$. Hence we have found p different eigenvalues $\omega_1,\dots,\omega_p$ for $\mathbf{Q}$ and the corresponding eigenvectors $\pi^{(1)},\dots,\pi^{(p)}$ which immediately implies that $\mathbf{\Pi}^{-1}$ exists and that $\mathbf{\Pi Q}$ is the matrix with rows $\omega_1\pi^{(1)},\dots, \omega_p\pi^{(p)}$. This is equivalent to the assertion of the Theorem. □

Remark As the proof of Th. 3.1 shows, then Condition 3.1 implies that $\mathbf{Q}$ can be written on diagonal form (the eigenvalues of the proof are different but all that really is required is the existence of linearly independent eigenvectors). Reversion of the proof shows immediately that the converse is also true. That is, the set-up is equivalent to the matrix $\mathbf{Q}$ to be of a spectral form which one intuitively feels is the typical case. When $\rho<1$, it is shown in [35] that det $\mathbf{H}(\theta) = 0$ has exactly p-1 roots with Re $\rho_i < 0$, zero as simple root and all other roots have Re $\rho > 0$. Some examples seem to indicate that when $\rho=1$, then typically det $\mathbf{H}(\theta) = 0$ has exactly p-1 roots with Re $\rho_i < 0$, zero as double root and all other roots have Re $\rho > 0$.

We define the moment matrices $\mathbf{M}^{(k)}$, $\mathbf{M}_+^{(k)}$, $\mathbf{M}_\ominus^{(k)}$ by

$$M^{(k)}(i,j) = \int_{-\infty}^{\infty} x^k F(i,j;dx) = \hat{F}^{(k)}(i,j;0),$$

$$M_{+}^{(k)}(i,j) = \int_{-\infty}^{\infty} x^k G_{+}(i,j;dx) = \hat{G}_{+}^{(k)}(i,j;0),$$

$$M_{\ominus}^{(k)}(i,j) = \int_{-\infty}^{\infty} x^k G_{\ominus}(i,j;dx) = \hat{G}_{\ominus}^{(k)}(i,j;0) = (-1)^k k! \eta^k.$$

Lemma 3.1 $\mathbf{H}(\theta) = \theta\mathbf{I} - \mathbf{S}(\theta) - \mathbf{\Lambda}$ *where* $\mathbf{S}(\theta)$ *is the diagonal matrix with the* $\beta_i(\hat{B}_i(\theta)-1)$ *in the diagonal. Furthermore the* $\mathbf{M}^{(k)}$ *are determined by* $\mathbf{S}'(0) = \mathbf{I} + \mathbf{\Lambda}/\eta + \eta\mathbf{M}^{(1)}$, $\mathbf{S}^{(k)}(0) = k\mathbf{M}^{(k-1)} + \eta\mathbf{M}^{(k)}$, $k = 2,3, \ldots$

Proof Obviously $\mathbb{E}_i e^{-\theta T} = \eta/(\eta+\theta)$ and $e^{\theta U} = 1$ unless an arrival occurs. Hence

$$\mathbb{E}_i[e^{\theta U}; J_1=j] = \begin{cases} \beta_i B_i(\theta)/\eta + (\eta+\lambda_{ii}-\beta_i)/\eta & i=j \\ \lambda_{ij}/\eta & i\neq j \end{cases}$$

$$(\eta+\theta)\hat{F}(\theta) = \eta\{\mathbf{S}(\theta)/\eta + \mathbf{I} + \mathbf{\Lambda}/\eta\}$$

from which the asserted expression for $\mathbf{H}(\theta)$ follows. Differentiating, we get

$$\mathbf{I} - \mathbf{S}'(\theta) = \mathbf{I} - \hat{\mathbf{F}}(\theta) - (\eta+\theta)\hat{\mathbf{F}}'(\theta),$$
$$\mathbf{S}^{(k)}(\theta) = k\hat{\mathbf{F}}^{(k-1)}(\theta) + (\eta+\theta)\hat{\mathbf{F}}^{(k)}(\theta), \quad k = 2,3, \ldots$$

(by induction). Let $\theta=0$ and note that $\hat{\mathbf{F}}(0) = \mathbf{I} + \mathbf{\Lambda}/\eta$.

Having found the fundamental matrices $\mathbf{Q}$, $\mathbf{Q}_0$ and thereby $\mathbf{G}_-$, the next step is to derive expressions for the relevant functionals of $\mathbf{G}_+$, in particular $\|\mathbf{G}_+\|$ and the $\mathbf{M}_+^{(k)}$.

Theorem 3.2 $\mathbf{\Pi}(\mathbf{I} - \|\mathbf{G}_+\|)$ *is the matrix with rows* $\pi^{(1)}(\mathbf{I} - \mathbf{S}'(0))$, $\pi^{(2)}\mathbf{\Lambda}/\lambda_2, \ldots, \pi^{(p)}\mathbf{\Lambda}/\lambda_p$. *Similarly, the rows of* $\mathbf{\Pi}\mathbf{M}_+^{(1)}$ *are* $\pi^{(1)}\mathbf{S}''(0)/2$,

$$\pi^{(i)}(\mathbf{I} - \mathbf{S}'(0) - \mathbf{\Lambda}/\lambda_i)/\lambda_i, \quad i = 2, \ldots, p,$$

and those of $\mathbf{\Pi}\mathbf{M}_+^{(2)}$ *are* $\pi^{(1)}\mathbf{S}'''(0)/3$,

$$2\pi^{(i)}(\mathbf{I} - \mathbf{S}'(0) - \lambda_i\mathbf{S}''(0) - \mathbf{\Lambda}/\lambda_i)/\lambda_i.$$

Proof It follows from (3.2) that

$$\mathbf{I} - \|\mathbf{F}\| - \theta\mathbf{M}^{(1)} - \theta^2\mathbf{M}^{(2)}/2 - \theta^3\mathbf{M}^{(3)}/3! - \ldots =$$
$$(\mathbf{I} - \|\mathbf{G}_\ominus\| - \theta\mathbf{M}_\ominus^{(1)} - \theta^2\mathbf{M}_\ominus^{(2)}/2 - \ldots)(\mathbf{I} - \|\mathbf{G}_+\| - \theta\mathbf{M}_+^{(1)} - \theta^2\mathbf{M}_+^{(2)}/2 - \ldots).$$

Equating coefficients and recalling that $\mathbf{I} - \|\mathbf{F}\| = -\mathbf{\Delta}/\eta$, $\|\mathbf{G}_\ominus\| = \mathbf{Q}$, $\mathbf{M}_\ominus^{(k)} = k!\,(-1)^k\,\mathbf{Q}/\eta^k$, we get

$$-\mathbf{\Delta}/\eta = (\mathbf{I} - \mathbf{Q})\,(\mathbf{I} - \|\mathbf{G}_+\|) \tag{3.4}$$
$$\mathbf{M}^{(1)} = -\eta^{-1}\mathbf{Q}(\mathbf{I} - \|\mathbf{G}_+\|) + (\mathbf{I} - \mathbf{Q})\,\mathbf{M}_+^{(1)} \tag{3.5}$$
$$\mathbf{M}^{(2)} = 2\eta^{-2}\mathbf{Q}(\mathbf{I} - \|\mathbf{G}_+\|) + 2\eta^{-1}\mathbf{Q}\mathbf{M}_+^{(1)} + (\mathbf{I} - \mathbf{Q})\mathbf{M}_+^{(2)} \tag{3.6}$$

The idea is now simply to solve recursively for the matrices $\|\mathbf{G}_+\|$, $\mathbf{M}^{(1)}$, $\mathbf{M}^{(2)}$ by (in a similar manner as in the proof of Th.3.1) determining the action on the basis vectors $\pi^{(1)},\ldots,\pi^{(p)}$. For i=2,...,p it follows from (3.4) that

$$\pi^{(i)}\,(\mathbf{I} - \|\mathbf{G}_+\|) = -\,\pi^{(i)}\,\mathbf{\Delta}/\eta(1-\omega_i) \;=\; \pi^{(i)}\,\mathbf{\Delta}/\lambda_i \tag{3.7}$$

whereas for i=1 (3.5) and Lemma 3.1 yield

$$\pi^{(1)}\,(\mathbf{I} - \|\mathbf{G}_+\|) = -\eta\,\pi^{(1)}\mathbf{M}^{(1)} = \pi^{(1)}(\mathbf{I} - \mathbf{S}'(0)) \tag{3.8}$$

This shows the assertion concerning $\|\mathbf{G}_+\|$ (note also that since $\mathbf{\Pi}$ is invertible, $\mathbf{I} - \|\mathbf{G}_+\|$ can be computed once $\mathbf{\Pi}(\mathbf{I} - \|\mathbf{G}_+\|)$ is known. Similarly, (3.5) and (3.7) yield

$$\begin{aligned}\pi^{(i)}\mathbf{M}_+^{(1)} &= \pi^{(i)}\mathbf{M}^{(1)}/\,(1-\omega_i) + \eta^{-1}\omega_i\,\pi^{(i)}\,(\mathbf{I} - \|\mathbf{G}_+\|)/(1-\omega_i)\\ &= \pi^{(i)}(\mathbf{I} + \mathbf{\Delta}/\eta - \mathbf{S}'(0))/\lambda_i - (\lambda_i^{-2} + 1/\eta\lambda_i)\pi^{(i)}\mathbf{\Delta}\\ &= \pi^{(i)}(\mathbf{I} - \mathbf{S}'(0))/\lambda_i - \pi^{(i)}\mathbf{\Delta}/\lambda_i^2\end{aligned}$$

whereas from (3.6), (3.8) and Lemma 3.1

$$\begin{aligned}\pi^{(1)}\mathbf{M}_+^{(1)} &= -\eta^{-1}\pi^{(1)}(\mathbf{I} - \|\mathbf{G}_+\|) + \pi^{(1)}\eta\mathbf{M}^{(2)}/2\\ &= \pi^{(1)}\mathbf{M}^{(1)} + \pi^{(1)}(\mathbf{S}''(0) - 2\mathbf{M}^{(1)})/2 = \pi^{(1)}\mathbf{S}''(0)/2.\end{aligned}$$

This shows the assertion on $\mathbf{\Pi}\mathbf{M}_+^{(1)}$ and the calculation in the case of $\mathbf{\Pi}\mathbf{M}_+^{(2)}$ is similar though more lengthy. □

4. Moments and conjugation

We let $\hat{F}_t(\alpha)$ be the matrix with elements

$$\hat{F}_t(i,j;\alpha) = \mathbb{E}_i[e^{\alpha R_t}; Z_t = j].$$

Then simple calculations along the lines of the proof of Lemma 3.1 yield

$$(d/dt)\hat{F}_t(\alpha) = \hat{F}_t(\alpha)(S(\alpha) + \mathbf{\Lambda} - \alpha I) \qquad \text{and hence}$$
$$\hat{F}_t(\alpha) = e^{t[S(\alpha) + \mathbf{\Lambda} - \alpha I]} \tag{4.1}$$

Further $e^{\kappa(\alpha)}$ denotes the spectral radius (Perron-Frobenius root) of $F_1(\alpha)$ and $\nu^{(\alpha)} = (\nu^{(\alpha)}(i))_{i\in E}$ the corresponding positive left (row) eigenvector normalised by $\nu^{(\alpha)}e = 1$ (here as before e is the column vector with all components equal to one). In particular, $\nu^{(0)} = \pi$.

Remark 4.1 It follows by general spectral theory that κ can alternatively be characterised as $(\log \text{spr}(F_\delta))/\delta$ or even simply $\text{spr}(S(\alpha) + \mathbf{\Lambda} - \alpha I)$.The reason that we have given the definition in terms of F_1 is to facilitate comparison with and translation to discrete time Markov-modulated random walks, where the basic governing parameter F of Sections 2-3 plays the role of F_1. In the same manner say the proof of Prop. 4.1 below has a slightly more direct continuous time version as well as similar remarks apply at a number of other places.

Proposition 4.1 *The function* $\kappa(\alpha)$ *is strictly convex with*

$$\kappa'(0) = \lim_{t\to\infty} \mathbb{E}R_t/t = \pi S'e - 1 \tag{4.2}$$

$$\kappa''(0) = \lim_{t\to\infty} \mathbb{V}\text{ar}\, R_t/t = \pi S''e + 2\pi S'DS'e \tag{4.3}$$

$$\kappa'''(0) = \pi S'''e + 3\pi S'DS''e + 3\pi S''DS'e + 6\pi S'DS'DS'e \tag{4.4}$$

Here $D = (e\pi - \mathbf{\Lambda})^{-1} - e\pi$ *and* $S^{(k)} = S^{(k)}(0)$ *is the diagonal matrix with the* $\beta_i \mathbb{E}_i U^k = \beta_i \hat{B}_i^{(k)}(0)$ *in the diagonal.*

Proof The strict convexity is proved in [21], and discrete time versions of (4.2), (4.3) are in [18], [20] (see also [7] p. 140). However, it is not apparent how to generalise to κ''' and even the formula for κ'' comes out in a rather indirect way. We shall therefore give the proof in full detail. Differentiating $e^{\kappa}v = v\hat{F}_1$ w.r.t. α we get

$$e^{\kappa}(\kappa'v+v') = v\hat{F}'_1 + v'\hat{F}_1 \tag{4.5}$$

$$e^{\kappa}(\kappa''v+(\kappa')^2v + 2\kappa'v'+v'') = v\hat{F}''_1 + 2v'\hat{F}'_1 + v''\hat{F}_1 \tag{4.6}$$

Noting that $ve = 1$ implies $0 = v'e = v''e = \ldots$, letting $\alpha = 0$ in (4.5) and using $\hat{F}_1(0)e = e$ we get

$$\kappa'(0) = \pi M_1^1 e \tag{4.7}$$

where $M_\delta^k = F_\delta^{(k)}(0)$. Using the same method for (4.6) the v'' terms vanish on both sides, but the v' term on the r.h.s. not. However, from (4.5) we have

$$\pi(M_1^1 - \kappa'(0)I) = v'(0)(I - \hat{F}_1(0)) = v'(0)(I + e\pi - \hat{F}_1(0)),$$

$$v'(0) = \pi(M_1^1 - \kappa'(0)I)D_1 \qquad \text{where } D_\delta = (I + e\pi - \hat{F}_\delta(0))^{-1}$$

(that the inverse exists follows since $\hat{F}_\delta(0) = e^{\delta A}$ is an ergodic transition matrix with stationary distribution π). Letting $\alpha = 0$ and multiplying by e to the right in (4.6) we get

$$\kappa''(0) + \kappa'(0)^2 = \pi M_1^2 e + 2\pi(M_1^1 - \kappa'(0)I)D_1 M_1^1 e,$$

$$\kappa''(0) = \pi M_1^2 e - 3\kappa'(0)^2 + 2\pi M_1^1 D_1 M_1^1 e \tag{4.8}$$

(using $\pi D_1 = \pi$), and (4.8) is indeed the expression of [20]. Similarly, (4.6) yields

$$v''(0) = [2v'(0)(M_1^1 - \kappa'(0)I) + \pi(M_1^2 - \kappa'(0)^2 I - \kappa''(0)I)]D_1,$$

$$\kappa'''(0) + 3\kappa'(0)^2\kappa''(0) + \kappa'(0)^3 = \pi M_1^3 e + 3v'(0)\pi M_1^2 e + 3v''(0)M_1^1 e \tag{4.9}$$

which can be solved for $\kappa'''(0)$. To arrive at the continuous time versions, note that $e^{\delta\kappa(\alpha)}$ is the spectral radius of $\hat{F}_\delta(\alpha)$ so that (4.7) yields $\delta\kappa'(0) = \pi M_\delta^1 e$. Since $M_\delta^1 = \delta(S' - I) + O(\delta^2)$, (4.2) follows and (4.3), (4.4) are derived by similar methods, using $M_\delta^k = \delta S'' + O(\delta^2)$, $k = 2,3,\ldots$, $De = \pi D = 0$ and

$$D_\delta = (I + e\pi - e^{\delta A})^{-1} = \delta^{-1}D + O(1).$$

□

Let now $h^{(\alpha)}$ be the positive right eigenvector of $\hat{F}_t(\alpha)$, $\hat{F}_t(\alpha)h^{(\alpha)} = e^{\kappa(\alpha)}h^{(\alpha)}$, and define $F_t^{(\theta)}$ by

$$F_t^{(\theta)}(i,j;dx) = \frac{h^{(\theta-\theta_0)}(j)}{h^{(\theta-\theta_0)}(i)} e^{(\theta-\theta_0)x - t\kappa(\theta-\theta_0)} F_t(i,j;dx) \qquad (4.10)$$

where θ_0 is some arbitrary location parameter. Following [38], [7] we assume that a γ_0 with $\kappa'(\gamma_0) = 0$ exists, and let $\theta_0 = -\gamma_0$. Also the existence of a $\gamma > 0$ with $\kappa(\gamma) = 0$, $\kappa''(\gamma) < \infty$ is needed (the discussion of such conditions is much as in the one-dimensional case and can be found in [18]).

It follows immediately from (4.10) that

$$\hat{F}_t^{(\theta)}(i,j;\alpha) = \frac{h^{(\theta-\theta_0)}(j)}{h^{(\theta-\theta_0)}(i)} e^{-t\kappa(\theta-\theta_0)} \hat{F}_t(i,j;\alpha+\theta-\theta_0)$$

That is, if $\mathbf{\Delta}_\theta$ is the diagonal matrix with the $h^{(\theta-\theta_0)}(i)$ in the diagonal, then

$$\hat{F}_t^{(\theta)}(\alpha) = e^{-t\kappa(\theta-\theta_0)} \mathbf{\Delta}_\theta^{-1} \hat{F}_t(\alpha+\theta-\theta_0)\, \mathbf{\Delta}_\theta \qquad (4.11)$$

$$= e^{-t\kappa(\theta-\theta_0)} \exp\{t[\mathbf{S}(\alpha+\theta-\theta_0) + \mathbf{\Delta}_\theta^{-1}\mathbf{\Lambda}\mathbf{\Delta}_\theta - (\alpha+\theta-\theta_0)\mathbf{I}]\}.$$

From this it follows by simple calculations that the rows of $\|F_t^{(\theta)}\| = \hat{F}_t^{(\theta)}(0)$ sum to one and that $\hat{F}_{t+s}^{(\theta)} = \hat{F}_t^{(\theta)}\hat{F}_s^{(\theta)}$. Thus we have a new Markov-modulated continuous time random walk, which for the present case can even be interpreted as a risk process. The changed parameters correspond to arrival intensities and claim size distributions given by

$$\beta_{\theta;i} = \beta_i \hat{B}_i(\theta-\theta_0), \quad \hat{B}_{\theta;i}(\alpha) = \frac{\hat{B}_i(\alpha+\theta-\theta_0)}{\hat{B}_i(\theta-\theta_0)},$$

and the intensity matrix $\mathbf{\Lambda}_\theta$ with ij th off-diagonal element

$$\lambda_{ij} \frac{h^{(\theta-\theta_0)}(j)}{h^{(\theta-\theta_0)}(i)}.$$

The basis for this interpretation is (4.1) and the formula

$$\hat{F}_t^{(\theta)}(\alpha) = e^{t[\,\mathbf{S}_\theta(\alpha) + \mathbf{\Lambda}_\theta - \alpha\mathbf{I}]} \qquad (4.12)$$

(here $S_\theta(\alpha) = S(\theta+\alpha) - S(\theta)$ is the diagonal matrix with the $\beta_{\theta;i}(B_\theta(\alpha)-1)$ in the diagonal) which can be obtained from (4.11) after some rather tedious calculations using the formula

$$\mathbf{\Lambda}_\theta = \mathbf{\Delta}_\theta^{-1}\mathbf{\Lambda}\mathbf{\Delta}_\theta + S(\theta-\theta_0) - (\kappa(\theta-\theta_0) + (\theta-\theta_0)\mathbf{I}).$$

We write $\mathbb{P}_{\theta;i}$ instead of $\mathbb{P}_i$ when the process is governed by $\{F^{(\theta)}\}_{t\geq 0}$ rather than $\{F_t\}_{t\geq 0}$, and $\pi^{(\theta)}$ denotes the corresponding stationary distribution for $\{Z_t\}$ or equivalently the positive left eigenvector for $\|\hat{F}_1^{(\theta)}\|$. It is immediately checked from (4.11) that $\pi^{(\theta)}(i) = h^{(\theta-\theta_0)}(i)\, \nu^{(\theta-\theta_0)}(i)$. Special notation like h^L, π^L, $\beta_{L;i}$ etc. are used for the Lundberg case $\theta_L = \theta_0 + \gamma$. Note that moments for the $\mathbb{P}_{\theta;i}$-process can easily be obtained in terms of the given κ-function in the same way as for a standard exponential family. In fact, by (4.11)

$$\mathrm{spr}(\hat{F}_t^{(\theta)}(\alpha)) = e^{-\kappa(\theta-\theta_0)}\,\mathrm{spr}(\hat{F}_t(\alpha+\theta-\theta_0)).$$

Thus $\kappa_\theta(\alpha) = \kappa(a+\theta-\theta_0) - \kappa(\theta-\theta_0)$,

$$\kappa_\theta'(0) = \lim_{t\to\infty} \mathbb{E}_\theta R_t/t = \kappa'(\theta-\theta_0) = \pi^{(\theta)} S_\theta e - 1 \tag{4.13}$$

and similarly for the analogues of (4.3), (4.4). Note that for $\theta = 0$ (4.13) becomes $\kappa'(-\theta_0) = 0$ while (by convexity) the expression is > 0 for $\theta > 0$ (in particular $\theta = \theta_L$) and < 0 for $\theta < 0$.

Lemma 4.1 *Let* T *be a stopping time w.r.t. the filtration* $\mathcal{F}_n = \sigma(J_k, S_k : k\leq n)$ *and* $F \in \mathcal{F}_n$ *an event satisfying* $F \subseteq \{T<\infty\}$. *Then for any* i, θ

$$\mathbb{P}_i F = \mathbb{P}_{\theta_0;i} F$$

$$= h^{(\theta-\theta_0)}(i)\,\mathbb{E}_{\theta;i}[\, h^{(\theta-\theta_0)}(J_T)^{-1} \exp\{(\theta_0-\theta)R_T + T\kappa(\theta-\theta_0)\}\, ; F]$$

$$(= h^L(i)\,\mathbb{E}_{L;i}[\, h^L(J_T)^{-1} \exp\{-\gamma R_T\}\, ; F] \qquad \text{when } \theta = \theta_L).$$

This likelihood ratio identity plays a crucial role in Sections 5-6 (for a proof, see [11], [41], [22]). We mention at this point one further application which will, however, not be spelled out in the present paper. This is importance sampling in the simulation evaluation of ruin probabilities, cf. [8], where $T = \tau(u)$. For example, one may simulate from $\mathbb{P}_{L;i}$, obtain i.i.d replicates of $(J_{\tau(u)}, R_{\tau(u)})$ and give estimates of the $\psi_{ij}(u)$ based on Lemma 4.1. The details follow [8] in a rather straightforward manner, and we would feel that the set-up of the present paper adds a further main example to [8] of models which are non-trivial to handle analytically but can be simulated with great advantage using this particular technique.

5. The overshoot and the Cramér-Lundberg approximation

If in Lemma 4.1 we let $T = \tau(u)$, $\theta = \theta_L$, $F = \tau(u)$ and define $B(u) = R_{\tau(u)} - u$ as the overshoot, we get $\psi_{ij}(u) = e^{-\gamma u}C(i,j;u)$ where

$$C(i,j;u) = \frac{h^L(i)}{h^L(j)}\mathbb{E}_{L;i}[e^{-\gamma B(u)}; J_{\tau(u)} = j] \tag{5.1}$$

Therefore the study of the distribution of $B(u)$ (or rather of the joint distribution of $J_{\tau(u)}, B(u)$) becomes of basic importance.

Proposition 5.1 *Suppose* $\theta \geq 0$. *Then a limit* $(J_{\tau(\infty)}, B(\infty))$ *of* $(J_{\tau(u)}, B(u))$ *as* $u \to \infty$ *exists in the sense of convergence of distributions. The distribution of the limit is given by the density*

$$b_k(x) = m(\theta)^{-1} \sum_{j \in E} \pi_+^{(\theta)}(j)\, G_+^{(\theta)}(j,k;(x,\infty)) \tag{5.2}$$

on the set $\{J_{\tau(\infty)} = k\}$. *Here* $\pi^{(\theta)}$ *is the stationary distribution for* $\|\mathbf{G}_+\|$ *and* $m(\theta) = \pi_+^{(\theta)}\mathbf{M}_+^{(1)}(\theta)e$.

Proof Obviously $(J_{\tau(u)}, B(u))$ is a semi-regenerative process (APQ X.3) with first semi-regeneration point $(J_{\tau(0)}, B(0)) = (J_{\tau_+}, S_{\tau_+})$, and a closer study shows that the given formulae are simply a translation of standard results for that setting (the non-lattice property being obvious). □

Corollary 5.1 $\psi_{ij}(u) \cong C(i,j)e^{-\gamma u}$, $u \to \infty$, *where the matrix* $\mathbf{C}$ *is given by*

$$\mathbf{C} = \kappa'(\gamma)^{-1}\, h^L \nu^L\, (\gamma\mathbf{I} - \mathbf{Q}_0)(\mathbf{I} - \|\mathbf{G}_+\|) \tag{5.3}$$

Proof By Prop. 5.1 and general results on weak convergence, the assertion holds with

$$C(i,j) = \lim_{u \to \infty} C(i,j;u) = \frac{h^L(i)}{h^L(j)}\mathbb{E}_{L;i}[e^{-\gamma B(\infty)}; J_{\tau(\infty)} = j]$$

and it only remains to check that $\mathbf{C}$ has the form (5.3). By Prop. 5.1,

$$\mathbb{E}_{L;i}[e^{-\gamma B(\infty)}; J_{\tau(\infty)} = j] = \int_0^\infty e^{-\gamma x} b_j(x)\, dx =$$

∞

$$m(\theta_L)^{-1} \sum_{l \in E} \pi_+^L(l) \gamma^{-1} \int_0 (1 - e^{-\gamma x}) \, G^L(l,j;dx) =$$

$$(\gamma m(\theta_L))^{-1} \sum_{l \in E} \pi_+^L(l) \{ \|G_+^L(l,j)\| - \hat{G}_+^L(l,j;-\gamma) \} \qquad (5.5)$$

In matrix formulation, this means that $\mathbf{C} = \mathbf{\Delta}_L e d \mathbf{\Delta}_L^{-1}$ where

$$d = (\gamma m(\theta_L))^{-1} \pi^L (\|\mathbf{G}_+^L\| - \hat{\mathbf{G}}_+^L(-\gamma)) = m(\theta_L)^{-1} \pi_+^L (\mathbf{I} - \hat{\mathbf{G}}_+^L(-\gamma))$$

To reduce this expression further, we need to involve also descending ladder heights, and here some caution is needed since these involve the uniformisation parameter η, whereas the ascending ones and the exponential family construction do not. One way to overcome this difficulty is to first fix the uniformisation parameter η for the given process, consider the discrete time Markov-modulated random walk with transform $\mathbf{F}(\alpha)$ given by Lemma 3.1, and form the corresponding exponential family $\{\mathbf{F}^{(\theta)}\}$. A rather tedious calculation (which we omit) then shows that the Lundberg conjugate is given by

$$\hat{\mathbf{F}}^L(\alpha) = \{\alpha \mathbf{I} - \mathbf{S}_L(\alpha) - \mathbf{\Delta}_L\}/(\eta_L + \alpha)$$

where $\mathbf{S}_L$, $\mathbf{\Delta}_L$ are the same as for the continuous-time exponential family and $\eta_L = \eta + \gamma$. That is, $\mathbf{F}^L$ can be related to the $\mathbb{P}_L$- distribution of $\{R_t\}$ in the same way as $\mathbf{F}$ to the $\mathbb{P}$-distribution of $\{R_t\}$ (this is not in general the case for $\theta \neq \theta_L$!). In particular, relating the means by an obvious time-average consideration, we get $\eta_L \pi^L \mathbf{M}^{(1)}(\theta_L) e = \kappa'(\gamma)$, cf. (4.13), and by Lemma 2.2(c) we may take

$$\pi^L = \pi_+^L (\mathbf{I} - \|\mathbf{G}_\ominus^L\|) = \nu^L \Delta^L (\mathbf{I} - \|\mathbf{G}_\ominus^L\|)$$

(since $\|\mathbf{G}^L\| e = e$). Then by (3.5)

$$m(\theta_L) = \pi_+^L \mathbf{M}_+^{(1)}(\theta_L) e = \pi^L \mathbf{M}^{(1)}(\theta_L) e = \eta_L^{-1} \kappa'(\gamma) = (\eta + \gamma)^{-1} \kappa'(\gamma).$$

Also it follows easily from Lemma 4.1 that

$$\|\mathbf{G}_\ominus^L\| = \mathbf{\Delta}_L^{-1} \hat{\mathbf{G}}_\ominus(\gamma) \, \mathbf{\Delta}_L, \quad \hat{\mathbf{G}}_+^L(-\gamma) = \mathbf{\Delta}_L^{-1} \|\mathbf{G}_+\| \, \mathbf{\Delta}_L$$

and hence

$$\mathbf{C} = \frac{\eta+\gamma}{\gamma\kappa'(\gamma)} \Delta_L e \nu^L \Delta_L (\mathbf{I} - \|\mathbf{G}^L_\ominus\|) (\mathbf{I} - \hat{\mathbf{G}}^L_+(-\gamma))\Delta_L$$

$$= \frac{\eta+\gamma}{\gamma\kappa'(\gamma)} h^L \nu^L (\mathbf{I} - \hat{\mathbf{G}}_\ominus(\gamma)) (\mathbf{I} - \|\mathbf{G}_+\|)$$

which is the same as the asserted expression. □

Except for some of the last constant manipulations, a result of the same form as Corollary 5.1 can be found in [25], [4], but no algorithms like those of Section 3 (and (4.2) for $\kappa(\gamma)$) were given for the numerical evaluation of $\mathbf{C}$. The present approach is somewhat different and leads also to certain related results, for example Segerdahls [37] time-dependent version of the Cramér-Lundberg approximation:

Corollary 5.2 $\psi_{ij}(u,T) \equiv C(i,j)e^{-\gamma u}\, \Phi\left(\dfrac{T - u/\kappa'(\gamma)}{(u\kappa''(\gamma)/\kappa'(\gamma)^3)^{1/2}}\right), u\to\infty.$

Proof By Lemma 4.1,

$$\psi_{ij}(u,T) = \frac{h^L(j)}{h^L(i)} e^{-\gamma u}\, \mathbb{E}_{L;i}[e^{-\gamma B(u)}; J_{\tau(u)} = j, \tau(u) \le T] .$$

It is not difficult to see (the details are in [10]) that $\tau(u)$ is asymptotically normal with mean $u/\kappa'(\gamma)$ and variance $u\kappa''(\gamma)/\kappa'(\gamma)^3$ (cf. Prop. 4.1) and that $\tau(u)$ is asymptotically independent of $(J_{\tau(u)}, B(u))$. From this the Corollary follows immediately. □

We finally remark that also various versions of Lundberg's inequality easily come out from (5.1) by obtaining suitable bounds on $C(i,j;u)$. For example, obviously

$$\psi_{ij}(u) \le \frac{h^L(i)}{h^L(j)} e^{-\gamma u}, \quad \mathbb{P}_i(\tau(u) < \infty) = \sum_{j\in E} \psi_{ij}(u) \le \frac{h^L(i)}{\min_j h^L(j)} e^{-\gamma u}.$$

6. Corrected diffusion approximations

We now think of the $\mathbb{P}_{0;i}$ as fixed and consider a limit where $\theta_0 \uparrow 0$, $u \to \infty$ in such a way that $\xi = u\theta_0 < 0$ remains fixed, and shall derive an inverse Gaussian approximation with correction terms (of order u^{-1}) for the $\psi_{ij}(u,T)$. The treatment is an extension of [38], [7] and for the steps which are essentially the same those papers may be consulted for more detail.

We let

$$G(T;\xi,c) = 1 - \Phi(cT^{-1/2} - \xi T^{1/2}) + e^{2\xi c}\,\Phi(-cT^{-1/2} - \xi T^{1/2})$$

denote the inverse Gaussian distribution corresponding to the first passage time of a Brownian motion with drift ξ to level c, and let $h(\lambda,\xi) = (2\lambda+\xi^2)^{1/2}-\xi$. Then the Laplace transform of $G(\cdot;\xi,c)$ is $e^{-ch(\lambda,\xi)}$, and a suitable version of the functional central limit theorem for continuous-time Markov-modulated random walks yields easily the existence of a standard Brownian limit for

$$\{(u^2\kappa''(0))^{-1/2}(R(tu^2) - \kappa'(0)tu^2)\}_{t\ge 0}$$

and thereby as in [38], [7] that

$$\mathbb{E}_{\theta_0;i}[e^{-\lambda\tau\kappa_0''/u^2};\tau<\infty] \;\to\; e^{-h(\lambda,\xi)} \tag{6.1}$$

Here and in the following $\tau = \tau(u)$ and κ_0'' means $\kappa''(0) = \kappa''(-\theta_0)$ etc. The idea is now to invoke also the $O(u^{-1})$ terms in (6.1) and to perform a formal inversion. Thus in the following $\cong$ means up to $o(u^{-1})$ terms. Define $\tilde{\theta} = (2\lambda+\xi^2)^{1/2}/u = (h(\lambda,\xi)+\xi)/u$. Then Lemma 4.1 yields

$$e^{-h(\lambda,\xi)}\;\frac{h^{(\tilde{\theta}-\theta_0)}(i)}{h^{(\tilde{\theta}-\theta_0)}(j)}\;\mathbb{P}_{\tilde{\theta};i}(J_\tau = j) \quad = \tag{6.2}$$

$$\mathbb{E}_{\theta\,;i}[\exp\{h(\lambda,\xi)B(u)/u - \tau\kappa(\tilde{\theta}-\theta_0)\};\,J_\tau = j,\,\tau<\infty] \tag{6.3}$$

In (6.2) it is easily seen that $\mathbb{P}_{\tilde{\theta};i}(J_\tau = j) \cong \mathbb{P}_{\tilde{\theta};i}(J_{\tau(\infty)} = j) = \pi_+^{(\tilde{\theta})}(j)$ where $\pi_+^{(\theta)}$ is as in Prop. 5.1 (in fact, inspection of the standard proof of the exponential ergodicity of finite Markov chains, APQ XI.1 or VI.2, shows that the remainder term is even exponentially small because of $\theta \to 0$). If

c_1, c_2 denote the derivatives of $\pi_+^{(\beta)}(j)$, resp. $h^{(\beta)}(i)/\ h^{(\beta)}(j)$, at $\beta=0$, we therefore have up to an $o(u^{-1})$ term that (6.2) is

$$e^{-h(\lambda,\xi)}\,(\pi_+^{(0)}(j) + c_1(h(\lambda,\xi) + \xi)/u)\,(1 + c_2\,h(\lambda,\xi)/u) \quad =$$
$$e^{-h(\lambda,\xi)}\,(\pi_+^{(0)}(j) + c_3/u + c_4 h(\lambda,\xi)/u) \tag{6.4}$$

where $c_3 = c_1\xi$, $c_4 = c_1 + c_2\pi_+^{(0)}(j)$. Taylor expansion next gives

$$\kappa(\tilde{\theta}-\theta_0) \;=\; \kappa_0(\tilde{\theta}) - \kappa_0(\theta_0) \;=\; (\tilde{\theta}^2-\theta_0^2)\kappa''/2 + (\tilde{\theta}^3-\theta_0^3)\kappa_0'''/6 + o(u^{-3})$$
$$= \lambda\kappa_0''/u^2 + h_1(\lambda,\xi)\kappa_0'''/6u^3 + o(u^{-3})$$

where $h_1(\lambda,\xi) = (h(\lambda,\xi) + \xi)^3 - \xi^3$. Hence up to an $o(u^{-1})$ term (6.3) is

$$\mathbb{E}_{\theta_0,i}[e^{-\lambda\tau\kappa_0''/u^2}\{1 + h(\lambda,\xi)B(u)/u - c_5/2u\,h_1(\lambda,\xi)\tau\kappa_0''/u^2\};\ J_\tau = j]$$

where $c_5 = \kappa_0'''/3\kappa_0''$. Using (6.1) and similar asymptotic independence arguments as in [38], [7] and Section 5 shows that this can be written as

$$\mathbb{E}_{\theta_0,i}[e^{-\lambda\tau\kappa_0''/u^2};\ J_\tau = j] + c_6/u\;e^{-h(\lambda,\xi)}\,h(\lambda,\xi) - c_5/2u\;\pi_+^{(0)}(j)\;h_2(\lambda,\xi) \tag{6.5}$$

where $c_6 = \mathbb{E}_0[B(\infty)\ ;\ J_{\tau(\infty)} = j]$ and

$$h_2(\lambda,\xi) = \;-\,h_1(\lambda,\xi)\,(\partial/\partial\lambda)e^{-h(\lambda,\xi)} = e^{-h(\lambda,\xi)}\,[2\lambda + \xi^2 - \xi^3/(2\lambda+\xi^2)^{1/2}].$$

Before equating (6.4) and (6.5) we perform one more manipulation. Taylor expansion of $\kappa_0(\theta_0) = \kappa_0(\theta_L)$ shows easily that $-\gamma u/2 \cong \xi + c_5\xi^2/2u$ where c_5 is the same as above. In terms of order u^{-1} we can therefore replace ξ by $\tilde{\xi} = -\gamma u/2$, and the only $O(1)$ term comes from (6.4),

$$e^{-h(\lambda,\xi)} \cong \; e^{-h(\lambda,\tilde{\xi})} + c_5 h_3(\lambda,\tilde{\xi})/2u \;\cong e^{-h(\lambda,\tilde{\xi})}(1 + c_5 h_3(\lambda,\xi)/2u)$$

where

$$h_3(\lambda,\xi) = \xi^2\,(\partial/\partial\xi)\,h(\lambda,\xi) = \xi^3/(2\lambda+\xi^2)^{1/2} - \xi^2 = 2\lambda e^{-h(\lambda,\xi)} - h_2(\lambda,\xi).$$

Combining, we get

7. Queueing reformulations

For the Markov-modulated M/G/1 queue, it will be convenient to assume that the underlying Markov jump process is the time-reversed (dual) version $\{\tilde{Z}_t\}$ of $\{Z_t\}$ rather than $\{Z_t\}$ itself (this is no restriction since $\{Z_t\}$ and $\{\tilde{Z}_t\}$ are in one-one correspondance). That is, the intensities of $\{\tilde{Z}_t\}$ are $\tilde{\lambda}_{ij} = \pi(j)\lambda_{ji}/\pi(i)$. We let $\tilde{Y}_n$ be the state of $\{\tilde{Z}_t\}$ at the n^{th} arrival and denote by W_n the waiting time of the n^{th} customer, by V_t the virtual waiting time at time t and let $W, V, \tilde{Y}, \tilde{Z}$ etc. refer to the steady state (which is well-defined by standard regeneration arguments). As in Sections 2-3, M_T is the maximum of $\{R_t\}$ over $[0,T]$ and $\tilde{M}_T$, $\{\tilde{R}_t\}$ etc. refers to the process governed by $\{\tilde{Z}_t\}$.

Theorem 7.1 *Define* $\pi\beta = \sum_{i\in E} \pi(i)\beta_i$. *Then*

$$\mathbb{P}(W > u, \tilde{Y} = i) = \frac{\pi(i)\beta_i}{\pi\beta}\mathbb{P}_i(M > u) \tag{7.1}$$

$$\mathbb{P}(V > u, \tilde{Z} = i) = \pi(i)\,\mathbb{P}_i(M > u) \tag{7.2}$$

Proof Just the same sample path argument as in the standard case (APQ III.7-8) shows that (taking $V_0 = 0$ for simplicity)

$$V_t = \max_{0\le s\le t} \{\tilde{R}_t - \tilde{R}_s\}.$$

Therefore by a time reversion argument,

$$\mathbb{P}_\pi(V_t > u, \tilde{Z}_0 = j, \tilde{Z}_t = i) = \mathbb{P}_\pi(M_t > u, Z_0 = i, Z_t = j), \quad \text{i.e.}$$

$$\pi(j)\,\mathbb{P}_j(V_t > u, Z_t = i) = \pi(i)\,\mathbb{P}_i(M_t > u, Z_t = j). \tag{7.3}$$

Since $M_t = M$ eventually, it is obvious that M_t and Z_t are asymptotically independent. Hence in the limit (7.3) becomes

$$\pi(j)\,\mathbb{P}(V > u, \tilde{Z} = i) = \pi(i)\,\mathbb{P}_i(M > u)\,\pi(j)$$

and (7.2) follows. For (7.1), it suffices according to Th. X.4.3 of APQ to show that $\mathbb{P}(Y = i) = \pi(i)\beta_i/\pi\beta$. This follows for example by a standard time-average consideration, identifying $\mathbb{P}(Y = i)$ by the asymptotic proportion of arrivals in state i and noting that in a period of length T the present paper certainly substantiates this belief even though a direct comparison is not straightforward.

Relations (7.1), (7.2) can be rewritten as

$$\mathbb{P}(W > u, \tilde{Y} = i) = \frac{\pi(i)\beta_i}{\pi\beta} \sum_{j\in E} \psi_{ij}(u) \tag{7.6}$$

$$\mathbb{P}(V > u, \tilde{Z} = i) = \pi(i) \sum_{j\in E} \psi_{ij}(u) \tag{7.7}$$

Inserting the Cramér-Lundberg approximation (Corollary 5.1) we therefore immediately get an approximation for the tails of the waiting times, and inserting the corrected diffusion approximation (6.8) yields heavy trafic approximations. Also the time-dependent case can be handled:

Corollary 7.1 *Subject to the limiting procedure of Section 6* ($u\to\infty$, $u\theta_0 = \xi$, $T\kappa_0''/u^2 = T_0$)

$$\mathbb{P}_j(V_T > u, \tilde{Z}_T = i) \;=\; \pi(i)\, \mathbb{P}_i(M_T > u) + o(u^{-1}) \tag{7.8}$$

Proof This follows by an extension of the proof of (7.2). Let $T' = T - T^{1/4}$. Then

$$\mathbb{P}_i(M_{T'} \le u, M_T > u) \;=\; \mathbb{P}_i(T' < \tau(u) \le T) =$$
$$\kappa_0^2(T - T')/u^2\, G'(T_0 + c_5/u; \tilde{\xi}, 1 + c_8/u) + o(u^{-1}) \;=\; o(u^{-1}),$$

using a formal inversion of (6.6) in the third step. Combining with (7.3) and uniform geometrical ergodicity, cf. the remarks following (6.2), we get

$$\pi(j)\, \mathbb{P}_j(V_T > u, \tilde{Z}_T = i) = \pi(i)\, \mathbb{P}_i(M_T > u,\, Z_T = j) =$$
$$\pi(i)\, \mathbb{P}_i(M_{T'} > u,\, Z_T = j) + o(u^{-1}) = \pi(i)\pi(j)\, \mathbb{P}_i(M_{T'} > u) + o(u^{-1}) =$$
$$\pi(i)\pi(j)\, \mathbb{P}_i(M_T > u) + o(u^{-1}).$$

An approximation of similar form as (6.7) now follows by replacing $\mathbb{P}_i(M_T > u)$ by $\Sigma_j\, \hat{\psi}_{ij}(u,T)$ where $\hat{\psi}_{ij}(u,T)$ is the approximation (6.7) for $\psi_{ij}(u,T)$. The case of the Cramér-Lundberg-Segerdahl approximation ($u\to\infty$) is similar but easier, and Corollary 5.2 yields

$$\mathbb{P}_j(V_T > u, Z_T = i) \cong \pi(i)\, C(i,j) e^{-\gamma u}\, \Phi\Big(\frac{T - u/\kappa'(\gamma)}{(u\kappa''(\gamma)/\kappa'(\gamma)^3)^{1/2}}\Big) \qquad (7.10)$$

Also the analogous expression for the $\mathbb{P}_j(W_N > u, \widetilde{Y}_N = i)$ can be given but some calculations are necessary for identifying the constants. In particular, one needs to replace the continuous time Markov-modulated exponential family by the discrete time one generated by the random walk with generic increments U^*-T^* where U^* is a service time and T^* an interarrival time. The details are a matter of routine and therefore omitted.

8. Remarks on M/M/1 and GI/M/1 type models

From the standard one-dimensional case (see e.g. APQ Ch. IX) one expects that the GI/M/1 and M/M/1 cases where all B_i are exponential, $B_i(dx) = \delta_i e^{-\delta_i x}dx$, are not only simpler but also that descending ladder heights ($\mathbf{G}_-$) form a detour and that the approach via ascending ones ($\mathbf{G}_+$) is more direct.

We shall not here give all details but only indicate some main steps. The crux is to determine the distribution of the maximum M^* of a Markov-modulated random walk $\{S_n^*\}$ where the increments are of the form $X_n^* = U_n - T_n$ where given $J_0^* = i$, $J_1^* = j$ the distribution of U_1 is exponential with rate δ_j (not δ_i!), $A_{ij}(t) = \mathbb{P}_i(T_1 \le t, J_1^* = j)$ is arbitrary, and U_1, T_1 are independent. Again, we suggest to use an uniformisation procedure to relate M^* to the maximum M of a Markov-modulated random walk $\{(J_n, S_n)\}$ with a simpler $\mathbf{G}_+$. To this end, we choose $\eta > \max_j \delta_j$ and given $J_0 = i$, we toss a coin w.p. δ_i/η for heads. If heads come up, $(J_1, -X_1)$ are chosen according to A_{ij}, and if tails come up, we let $J_1 = i$ and X_1 be exponential with rate η. This means simply that the U_n are split up into geometric sums of exponential variables with rate η, but adds also the complication that $\{(J_n, S_n)\}$ starts differently from $\{(J_n^*, S_n^*)\}$ if the first coin tossing yields a tail. That is, M^* is distributed as M given the event F of an initial head. With $G_+^{\#}(i,j;A) = \mathbb{P}_i(S_{\tau_+} \in A, J_1 = j \mid F)$, it can then be seen in analogy with Prop. 7.1 that

Proposition 8.1 *The* $\mathbb{P}_i$ *- distribution of* M^* *is the* i *th component of the vector*

$$(\mathbf{I} - \|\mathbf{G}_+^{\#}\| + \mathbf{G}_+^{\#} * \mathbf{U}_+ * (\mathbf{I} - \|\mathbf{G}_+\|)\mathbf{e}. \qquad (8.1)$$

Obviously, $G_+(i,j)/\|G_+(i,j)\|$ and $G_+^{\#}(i,j)/\|G_+^{\#}(i,j)\|$ are both exponential with rate η, and we thus have to determine $\mathbf{Q} = \|\mathbf{G}_+\|$ and $\mathbf{Q}^{\#} = \|\mathbf{G}_+^{\#}\|$. First $\mathbf{Q}^{\#}$ can easily be determined in terms of $\mathbf{Q}$ since removing F^c corresponds to removing mass $1 - \delta_i/\eta$ from $G_+(i,i)$. That is,

$$q_{ij}^{\#} = q_{ij}/\mathbb{P}_i F = q_{ij}\,\eta/\delta_i,\ i \neq j, \quad q_{ii}^{\#} = (q_{ii} - 1 + \delta_i/\eta)\,\eta/\delta_i.$$

Finally to get $\mathbf{Q}$, the Wiener-Hopf identity (3.2) yields

$$\mathbf{I} - \hat{F}(\alpha) = (\mathbf{I} - \hat{\mathbf{G}}_{\ominus}(\alpha))(\mathbf{I} - \eta\mathbf{Q}/(\eta-\alpha))$$

and if $\rho < 1$ so that $\mathbf{Q}$ is substochastic (and $\mathbf{I} - \hat{\mathbf{G}}_{\ominus}(\alpha)$ non-singular for Re $\alpha > 0$), arguments of just the same type as in Section 3 yield

Theorem 8.1 $\mathbf{H}(\theta) = (\eta-\theta)(\mathbf{I} - \hat{F}(\theta))$ *is given by*

$$H(i,i;\theta) = -\delta_i A_{ii}(-\theta) + \delta_i - \theta, \quad H(i,j;\theta) = -\delta_i A_{ij}(-\theta),\ i \neq j.$$

Furthermore the matrix $\mathbf{Q}$ *has diagonal form if and only if* det $\mathbf{H}(\theta) = 0$ *has* p *solutions* $\beta_1,\dots,\beta_p$ *with* Re $\beta_i > 0$ *and corresponding linear independent right eigenvectors* $e_1, \dots, e_p$. *Then also* e_i *is eigenvector of* $\mathbf{Q}$ *corresponding to the eigenvalue* $\rho_i = 1 - \beta_i/\eta$, *and*

$$\mathbf{Q} = \|\mathbf{G}_+\| = (\rho_1 e_1 \dots \rho_p e_p)(e_1 \dots e_p)^{-1}.$$

Remark 8.1 The above results are related to [40] in much the same way as Sections 2,3,7 to [35], cf. the remarks following the proof of Prop. 7.1. In particular we refind the observation of [40] that the density of M^* on $(0,\infty)$ is a linear combination of exponential terms. This may be seen, e.g., by noting that

$$\hat{U}(\theta) = \sum_{n=0}^{\infty} \eta^n/(\eta-\theta)^n (\rho_1 e_1 \dots \rho_p e_p)(e_1 \dots e_p)^{-1}$$

$$= (\eta-\theta)\left(((1-\rho_1)\eta - \theta)^{-1} e_1 \dots ((1-\rho_p)\eta - \theta)^{-1} e_p \right)(e_1 \dots e_p)^{-1},$$

$\mathbf{G}_+^{\#}(\theta) = \eta\mathbf{Q}^{\#}/(\eta-\theta)$, so that the m.g.f. of (8.1) is a linear combination of terms of the form $((1-\rho_i)\eta - \theta)^{-1}$.

A different example is a Markov-modulated storage process $\{V_t\}_{t\geq 0}$ considered for an interesting special case in Gaver and Lehoczky [14]. Here V_t moves linearly at rate $\lambda(i)$ when $Z_t = i$ and 0 acts as reflecting barrier, cf. Fig. 2. Letting $R_t = \int^t \lambda(Z_s)ds$ and taking $V_0 = 0$ for simplicity, it is easy to see along the lines of Section 7 that

$$V_T = \max_{0\leq t\leq T} \{\tilde{R}_T - \tilde{R}_t\},$$

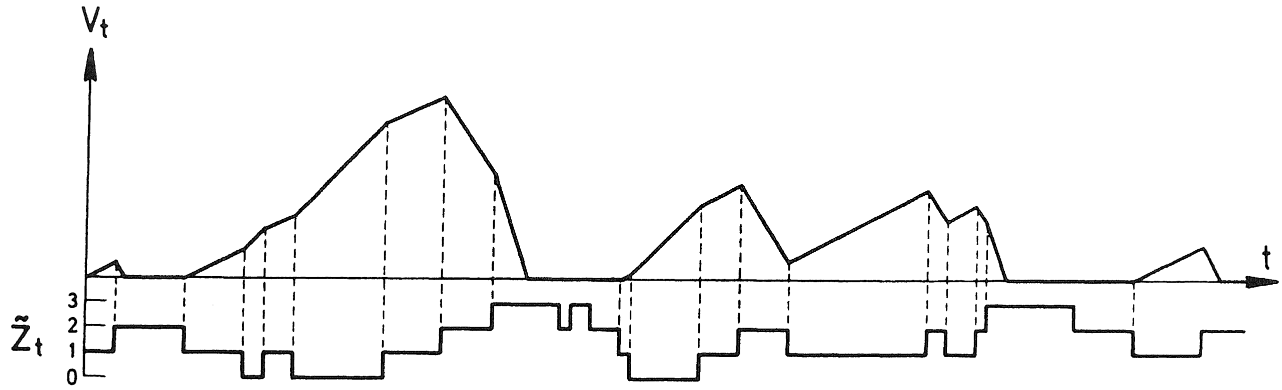


Figure 2

that a limit V exists if and only if $\sum \pi(i)\lambda(i) < 0$ and that then

$$\mathbb{P}(V > u, \tilde{Z} = i) = \pi(i)\, \mathbb{P}_i(M^* > u) \quad \text{where } M^* = \max_{0 \le t < \infty} R_t$$

(with the usual notational conventions for time-reversion). Letting $\{J_n^*\}$ be the imbedded jump chain of $\{Z_t\}$ (the transition matrix and stationary distribution can be obtained in a standard manner from $\mathbf{\Lambda}, \pi$), we have $M^* = \max\{0, X_1^*, X_1^*+X_2^*, \ldots\}$ where $X_n^* = \lambda(J_{n-1}^*)Y_n$ with $Y_1, Y_2, \ldots$ i.i.d. exponential with unit rate. Thus this model leads to a random walk problem of M/M/1 type. A minor variant of the uniformisation procedure discussed above for the GI/M/1 case applies here as well: choose $\eta > \max\{\lambda(i)^{-1}: \lambda(i)>0\}$ and split X_n^* up into a geometric sum of exponential variables with rate η whenever $\lambda(J_{n-1}^*) > 0$.

Acknowledgement The present work was done mainly while the author was visiting Stanford University. I gratefully acknowledge the hospitality of the Department of Operations Research and a grant from the Danish Natural Science Research Council. Thanks also to Ida Jørgensen and the Department of Computer Science at the University of Copenhagen, to AB Consult in Fredericia, to Niels Bach Pedersen and A/S Regnecentralen, and to Mrs. Mac. All were most helpful in the production of the manuscript.

References

The aim has been to give a comparatively complete bibliography, and therefore the list contains a few numbers not cited in the text.

[1] Arjas, E. (1972). On the use of a fundamental identity in the theory of semi-Markov processes. *Adv. Appl. Probab.* **4**, 271-284.

[2] Arjas, E, and Speed, T.P. (1973a). Symmetric Wiener-Hopf factorisations in Markov additive processes. *Z. Wahrscheinlichkeitsth. verw. Geb.* **26**, 105-118.

[3] Arjas, E. and Speed, T.P. (1973b). Topics in Markov additive processes. *Math. Scand.* **33**, 171-192.

[4] Arndt, K. (1980). Asymptotic properties of the distribution of the supremum of a random walk on a Markov chain. *Th. Probab. Appl.* **25**, 309-324.

[5] Arndt, K. (1981). On properties of boundary functionals of a random walk on a Markov chain. *Math. Operationsforschung Statist.* **12**, 85-100.

[6] Arndt, K. (1984). On the distribution of the supremum of a random walk on a Markov chain. *Limit Theorems and Related Problems* (A.A. Borovkov ed.), 253-267. Optimizations Software Inc., New York.

[7] Asmussen, S. (1984). Approximations for the probability of ruin within finite time. *Scand. Actuarial J.*, 31-57; *ibid.* 1985, 64.

[8] Asmussen, S. (1985). Conjugate processes and the simulation of ruin problems. *Stoch. Proc. Appl.* **20**, 213-229.

[9] Asmussen, S. (1987). *Applied Probability and Queues.* John Wiley & Sons, Chichester New York Brisbane Toronto Singapore.

[10] Asmussen, S. and Thorisson, H. (1986). Large deviation results for time-dependent queue lengt probabilities. *Submitted for publication.*

[11] Bellman, R. (1957). On a generalization of the fundamental identity of Wald. *Proc. Camb. Philos. Soc.* **53**, 257-259.

[12] Borovkov, K.A. (1980). Stability theorems and estimates of the rate of convergence of the components of factorizations for walks defined on Markov chains. *Th. Probab. Appl.* **25**, 325-334.

[13] Feller, W. (1971). *An Introduction to Probability Theory and Its Applications*, II (2nd ed.). Wiley, New York.

[14] Gaver, D.P. and Lehoczky, J.P. (1982). Channels that cooperatively service a data stream and voice messages. *IEEE Trans. Com*-**30**, 1153-1161.

[15] Höglund, T. (1974). Central limit theorems and statistical inference for finite Markov chains. *Z. Wahrscheinlichkeitsth. verw. Geb.* **29**,123-151.

[16] Höglund, T. (1986). The ruin problem for finite Markov chains. Technical report.

[17] Keilson, J. (1979). *Markov Chain Models - Rarity and Exponentiality.* Springer-Verlag, New York Heidelberg Berlin.

[18] Keilson, J. and Wishart, D.M.G. (1964). A central limit theorem for processes defined on a finite Markov chain. *Proc. Camb. Philos. Soc.* **60**, 547-567.

[19] Keilson, J. and Wishart, D.M.G. (1965). Boundary problems for additive processes defined on a finite Markov chain. *Proc. Camb. Philos. Soc* **61**, 173-190.

[20] Keilson, J. and Wishart, D.M.G. (1967).Addenda to processes defined on a finite Markov chain. *Proc. Camb. Philos. Soc.* **63**, 187-193.

[21] Kingman, J.F.C. (1961). A convexity property of positive matrices. *Quart. J. Math. Oxford* **12**, 283-284.

[22] Miller, H.D. (1961a). A generalization of Wald's identity with applications to random walks. *Ann. Math. Statist.* **32**, 549-560.

[23] Miller, H.D. (1961b). A convexity property in the theory of random variables defined on a finite Markov chain. *Ann. Math. Statist.* **32**, 1260-1270.

[24] Miller, H.D. (1962a). A matrix factorization problem in the theory of random variables defined on a finite Markov chain. *Proc. Camb. Philos. Soc.* **58**,268-285.

[25] Miller, H.D. (1962b). Absorption probabilities for sums of random variables defined on a finite Markov chain. *Proc. Camb. Philos. Soc.* **58**, 286-298.

[26] Mogul'skii, A.A. (1974). Factorization identities for processes with independent increments given on a Markov chain. *Th. Probab. Math. Statist.* **11**

[27] Neuts, M.F. (1981). *Matrix-Geometric Solutions in Stochastic Models.* Johns Hopkins University Press, Baltimore.

[28] Ney, P. and Nummelin, E. (1986). Markov additive processes. I. Eigenvalues and limit theorems; II. Large deviations. *Adv. Appl. Math.* (to appear).

[30] Newbould, M. (1973). A classification of a random walk defined on a finite Markov chain. *Z. Wahrscheinlichkeitsth. verw. Geb.* **26**, 95-104.

[31] Pitman, J.W. (1974). An identity for stopping times of a Markov process. *Studies in Probability and Statistics* (E.J. Williams ed.), 41-57. Jerusalem Academic Press, University of Melbourne.

[32] Presman, E.L. (1967). A boundary value problem for sums of lattice random variables given on a finite regular Markov chain. *Th. Probab. Appl.* **12,** 323-328.

[33] Presman, E.L. (1969). Factorization methods and boundary problems for sums of random variables given on a Markov chain. *Math. USSR Izv.* **3,** 815-852. (English translation).

[34] Ramaswami,V. and Latouche, G. (1985). A general class of Markov processes with explicit matrix-geometric solutions. *Opns. Res. Spectrum*

[35] Regterschot, G.J.K and de Smit, J.H.A. (1986). The queue M/G/1 with Markov-modulated arrivals and services. *Math. Opns. Res.*

[36] Seal, H.L. (1974). The numerical calculation of U(w,t), the probability of non-ruin in an interval (0,t). *Scand. Actuarial J.*, 121-139.

[37] Segerdahl, C.-O. (1955). When does ruin occur in the collective theory of risk? *Skand.Aktuar.Tidsskr.*

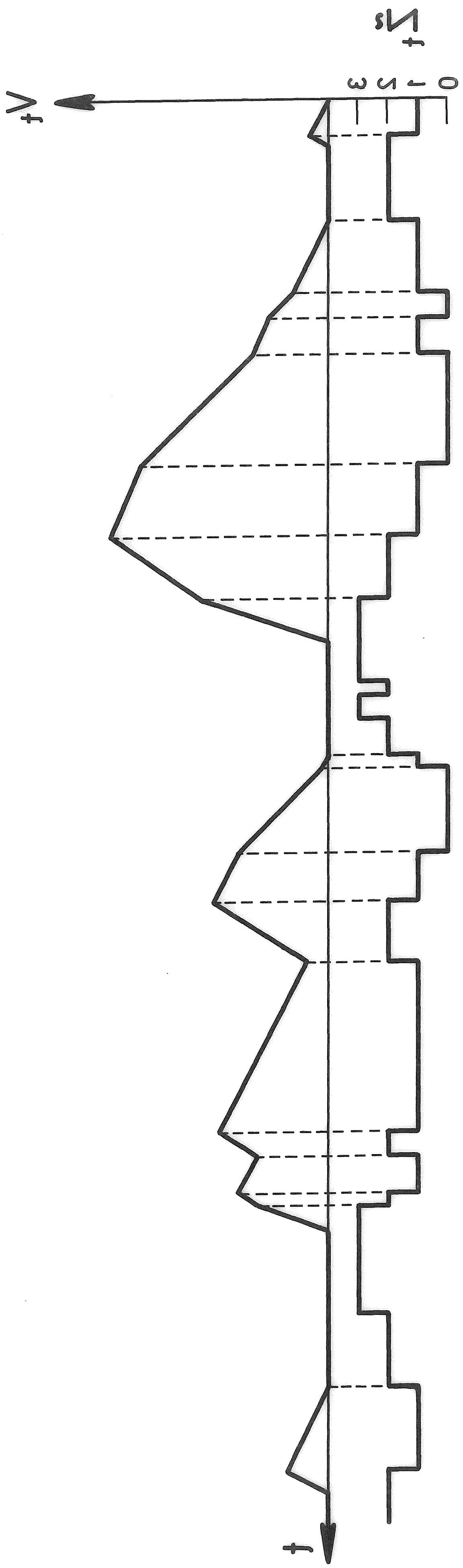
[38] Siegmund, D. (1979). Corrected diffusion approximations in certain random walk problems. *Adv. Appl. Probab.* **11**, 701-719.

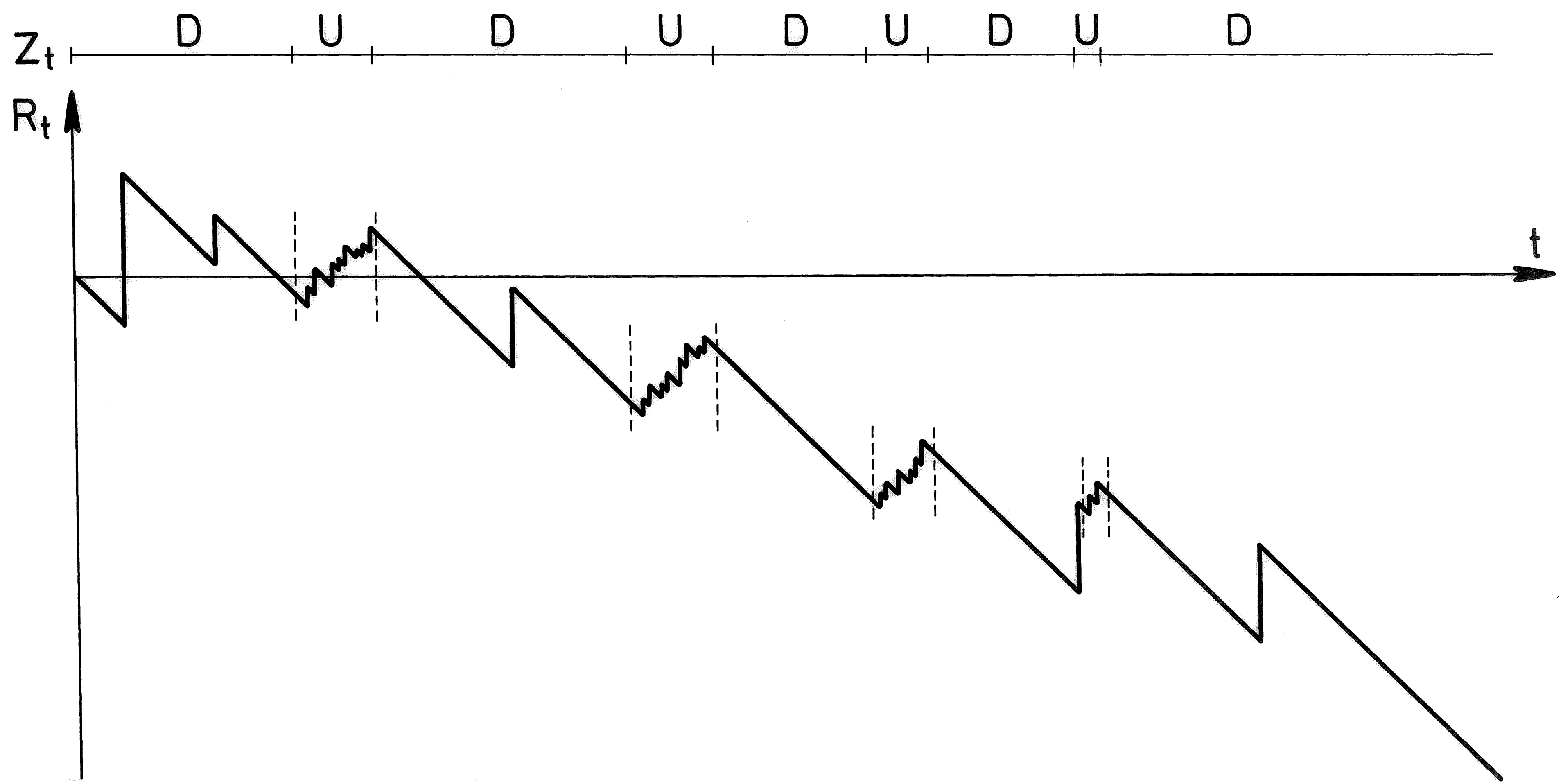
[39] de Smit, J.H.A. (1986). The single server semi-Markov queue. *Stoch. Proc. Appl.* **22**,37-50.

[40] de Smit, J.H.A and Regterschot, G.J.K. (1986). A semi-Markov queue with exponential service times. *Semi-Markov Processes. Theory and Applications* (J. Janssen ed.), 369-382 (1986). Plenum, New York.

[41] Tweedie, M.C.K. (1960). Generalization of Wald's fundamental identity of sequential analysis to Markov chains. *Proc. Camb. Philos. Soc.* **56,** 205-214.

[42] Volkov, I.S. (1958). On the distribution of sums of random variables defined on a homogeneous Markov chain. *Th. Probab. Appl.* **3**


0
1
2
3
t



Z_t
D
U
D
U
D
U
D
U
D
R_t
t

PREPRINTS 1985

COPIES OF PREPRINTS ARE OBTAINABLE FROM THE AUTHOR OR FROM THE INSTITUTE OF MATHEMATICAL STATISTICS, UNIVERSITETSPARKEN 5, 2100 COPENHAGEN Ø, DENMARK.

No. 1 Cohn, Harry: Almost Sure Convergence for Stochastically Monotone Temporally Homogeneous Markov Processes and Applications.

No. 2 Rootzén, Holger: Maxima and Exceedances of Stationary Markov Chains.

No. 3 Asmussen, Søren und Johansen, Helle: Über eine Stetigkeitsfrage betreffend das Bedienungssystem GI/GI/s.

No. 4 Asmussen, Søren and Thorisson, Hermann: A Markov Chain Approach to Periodic Queues.

No. 5 Hald, Anders: Galileo as Statistician.

No. 6 Johansen, Søren: The Mathematical Structure of Error Correction Models.

No. 7 Johansen, Søren and Johnstone, Iain: Some Uses of Spherical Geometry in Simultaneous Inference and Data Analysis.

PREPRINTS 1986

COPIES OF PREPRINTS ARE OBTAINABLE FROM THE AUTHOR OR FROM THE INSTITUTE OF MATHEMATICAL STATISTICS, UNIVERSITETSPARKEN 5, 2100 COPENHAGEN Ø, DENMARK.

No. 1 Jespersen, N.C.B.: On the Structure of Simple Transformation Models.

No. 2 Dalgaard, Peter and Johansen, Søren: The Asymptotic Properties of the Cornish-Bowden-Eisenthal Median Estimator.

No. 3 Jespersen, N.C.B.: Dichotomizing a Continuous Covariate in the Cox Regression Model.

No. 4 Asmussen, Søren: On Ruin Problems and Queues of Markov-Modulated M/G/1 Type.