

THE TAX IDENTITY FOR MARKOV ADDITIVE RISK PROCESSES

Hansjörg Albrecher ^{a,b} Florin Avram ^c Corina Constantinescu ^{a,d,*}
Jevgenijs Ivanovs ^{a,*}

^a *Department of Actuarial Science, Faculty of Business and Economics, University of Lausanne, Quartier UNIL-Dorigny, Batiment Extranef, CH-1015 Lausanne, Switzerland*

^b *Swiss Finance Institute*

^c *Department de Mathematiques, Université de Pau, France 64000*

^d *Institute for Financial and Actuarial Mathematics, Department of Mathematical Sciences, University of Liverpool, L69 7ZL, Liverpool UK*

Abstract

Taxed risk processes, i.e. processes which change their drift when reaching new maxima, represent a certain type of generalizations of Lévy and of Markov additive processes (MAP), since the times at which their Markovian mechanism changes are allowed to depend on the current position. In this paper we study generalizations of the tax identity of Albrecher and Hipp [3] from the classical risk model to more general risk processes driven by spectrally-negative MAPs. We use the Sparre Andersen risk processes with phase-type interarrivals to illustrate the ideas in their simplest form.

Keywords First-passage time, taxed Sparre Andersen risk process, spectrally-negative Markov additive processes

1 Introduction

In this work we reconsider the tax identities of Albrecher and Hipp [3], which relate the survival probabilities of the Cramér-Lundberg risk process and its taxed version. The initial motivation of this work was to extend these identities to risk models driven by spectrally-negative Markov additive processes (MAP), such as Sparre Andersen risk processes with phase-type (PH) interarrivals. As we will show in this paper, most of the theory, including the linear differential equation defining the survival probabilities, can indeed be

*Supported by the Swiss National Science Foundation Project 200021-124635/1.

generalized. In one dimension, the solution of this equation has an exponential form, which immediately leads to a power relation. This is not the case in the setting of MAP-driven risk processes. Nevertheless, the solution can be expressed using so-called product integrals.

The taxed version of a risk process $X(t)$ is defined by

$$X_\gamma(t) := X(t) - \int_0^t \gamma(\bar{X}(s)) d\bar{X}(s), \quad (1.1)$$

where $\bar{X}(t) := \sup_{0 \leq s \leq t} X(s)$ denotes the running maximum of $X(t)$. When $\gamma(x) \equiv \gamma \in [0, 1)$ is a constant, this means reducing the drift c to $c(1 - \gamma)$ whenever the process $X(t)$ (or, equivalently, $X_\gamma(t)$) coincides with its running maximum. In the presence of a modulating environment it is of interest to allow the taxation rate γ to depend on the state of this environment. We define the taxed version of a MAP-driven risk process in Section 6.

In Section 2 we review and reinterpret the derivation of the tax identity for the Cramér-Lundberg process. In Section 3 we show that the Sparre Andersen model with PH interarrival times can be analyzed in essentially the same way. The general case of a spectrally-negative MAP requires some new ideas, which nevertheless are of a similar spirit. Some basic theory of MAPs is given in Section 4, whereas Section 5 contains a short review of exit problems including the definition of a so-called scale matrix. The taxed MAP-driven risk process is analyzed in Section 6, see in particular Theorem 6.1, which extends the results of Ji et al. [19] and of Wei et al. [25]. Finally, in Section 7 we briefly discuss a special case, where certain formulas become more explicit.

For practical computations there are essentially two difficulties. One is that of calculating the scale matrix of the original MAP, i.e. of solving the two-sided exit problem, and the other is to numerically solve a system of linear differential equations with variable coefficients. For an asymptotic result on the tax identity for renewal models see Albrecher et al. [2].

2 Review of the tax identity for the Cramér-Lundberg model

The classical Cramér-Lundberg process of collective risk theory (see e.g. Gerber [15]) models the surplus of an insurance company at time t by

$$X(t) = X(0) + ct - \sum_{k=1}^{N(t)} C_k, \quad (2.1)$$

where C_k are i.i.d. positive random variables representing claims, $N(t)$ is an independent Poisson process with intensity λ modeling the times at which the claims occur, and c represents the premium rate. Define the probability of survival $\phi_\gamma(x) := P(X_\gamma(t) > 0 \text{ for all } t > 0 | X_\gamma(0) = x)$ for the taxed process $X_\gamma(t)$ whose dynamics are defined in (1.1), and assume that the taxation rate γ is constant.

As proposed in [1], it is beneficial to view the taxed process in the following way. While at the maximum it evolves as a linear drift of rate $c(1 - \gamma)$. Upon arrival of a claim, the linear drift process is stopped until the present level is regained, and then it is restarted again. Note that the excursions from the maximum initiated by an arrival of a claim do not depend on the taxation rate - these are excursions from the maximum of the original process. Furthermore, an excursion may lead to ruin, in which case we kill the linear drift process. According to this view, survival of the surplus process coincides with survival of the killed linear drift, but the latter satisfies the following differential equation

$$c(1 - \gamma)\phi'_\gamma(x) - \lambda q(x)\phi_\gamma(x) = 0, \quad (2.2)$$

where $q(x) = \mathbb{P}(H > x)$ is the probability that the height of an excursion from the maximum exceeds x . Indeed, the analysis of jumps of a Poisson process in the interval $[0, \frac{h}{c(1-\gamma)}]$ shows that

$$\phi_\gamma(x) = \left(1 - \lambda q(x) \frac{h}{c(1-\gamma)}\right) \phi_\gamma(x+h) + o(h),$$

because $\mathbb{P}(x < H \leq x+h) = o(1)$, as $h \downarrow 0$. Roughly, the main idea of the above derivation is to observe the risk process along the value axis instead of the time axis. This is exactly the viewpoint of Itô's excursion theory. Furthermore, we make the following observations:

1. The "killing event" is the first point of a nonhomogeneous Poisson process with rate $\frac{\lambda}{c(1-\gamma)}q(x)$, therefore we can write

$$\phi_\gamma(x) = e^{-\int_x^\infty \frac{\lambda}{c(1-\gamma)}q(u)du}, \quad (2.3)$$

which can also be obtained directly by solving (2.2).

2. The exponential form of the solution (2.3) immediately yields the following power relation between taxed and non-taxed survival probabilities:

$$\phi_\gamma(x) = \phi_0(x)^{\frac{1}{1-\gamma}}.$$

3. Formally, the derivative in Equation (2.2) is a right derivative $\phi_{\gamma'_+}(x)$. In fact, the left derivative is given by

$$c(1-\gamma)\phi_{\gamma'_-}(x) - \lambda q(x-)\phi_\gamma(x) = 0,$$

where $q(x-) = \mathbb{P}(H \geq x)$ is the left limit.

4. The Cramér-Lundberg risk model can be generalized to a spectrally-negative Lévy process, that is a process with stationary and independent increments with only negative jumps. In this more general setting one can use excursion theory to derive the counterparts of (2.2) and (2.3), see [4, 20]. In the theory of spectrally-negative Lévy processes there exists a so-called scale function $W(x)$, $x \geq 0$, which can be used to express $\phi_0(x)$ in the following way

$$\phi_0(x) = W(x)/W(\infty).$$

Our differential Equation (2.2) for a Cramér-Lundberg process without tax shows that

$$\frac{\lambda}{c}q(x) = \frac{W'_+(x)}{W(x)}.$$

Indeed, the right-hand side has a well-known interpretation in excursion theory: it is the rate of arrivals of excursions exceeding height x . So $q(x)$ is available explicitly whenever $W(x)$ is¹.

¹It may happen that $q(x)$ is more convenient to work with than the scale function; note that the "Riccati substitution" $\frac{W'_+(x)}{W(x)}$ is used in computer programs for solving ODE's for detecting "Liouvillian solutions", which are precisely those solutions for which $q(x)$ is rational, but for which $W(x)$ is not.

5. It is not immediately clear how to generalize the approach based on excursion theory to the MAP setting. Our approach, presented in Section 6 resembles excursion theory, but relies only on some basic tools from the theory of Markov chains.
6. It is not difficult to see that Equation (2.2) still holds if the taxation rate $\gamma(x)$ is a function of x , and we write $c(1 - \gamma(x))$ instead of $c(1 - \gamma)$. Hence the essential difficulty in obtaining $\phi_\gamma(x)$ is that of calculating the scale function corresponding to the original process.
7. Consider the general Gerber-Shiu discounted penalty function

$$m_0(x) = \mathbb{E}_x[e^{-\delta\tau_0^-} w(X(\tau_0^-)^-, X(\tau_0^-)) 1_{\tau_0^- < \infty}],$$

where τ_0^- is the time of ruin, w is a penalty as a function of surplus prior to ruin and deficit at ruin, and $\delta \geq 0$ is a force of interest. Let $m_\gamma(x)$ denote the corresponding quantity in the same risk model with tax. Assume $\delta = 0$ for now (the general case $\delta \geq 0$ can be regained by exponential killing, see also Remark 6.4). Equation (2.2) then extends to

$$c(1 - \gamma)m'_\gamma(x) - \lambda q(x)m_\gamma(x) + \lambda q(x)r_w(x) = 0,$$

where $r_w(x)$ is the penalty resulting from an excursion starting at the level x . This penalty function does not depend on taxation, hence we obtain

$$c(1 - \gamma)m'_\gamma(x) - \lambda q(x)m_\gamma(x) + [-cm'_0(x) + \lambda q(x)m_0(x)] = 0,$$

which can be rewritten using scale functions as

$$(1 - \gamma)m'_\gamma(x) - \frac{W'_+(x)}{W(x)}m_\gamma(x) + \left(-m'_0(x) + \frac{W'_+(x)}{W(x)}m_0(x)\right) = 0.$$

Assuming that the process $X(t)$ drifts to infinity (or that the force of interest is positive) we see that $m_\gamma(\infty) = 0$, which then yields

$$m_\gamma(x) = \frac{1}{1 - \gamma} \int_x^\infty \left(\frac{W'_+(y)}{W(y)}m_0(y) - m'_0(y) \right) \left(\frac{W(x)}{W(y)} \right)^{\frac{1}{1-\gamma}} dy.$$

This formula generalizes Theorem 2.1 of Wang et al. [24] and may be viewed as an extension of the tax identity, see also Cheung and Landriault [12, (2.16)] and Kyprianou and Zhou [20].

3 The Sparre Andersen renewal risk model with phase-type interarrival times

In this section we generalize Equation (2.2) to the Sparre Andersen renewal risk model with PH interarrival times, see [22] for the analysis of this model with no tax. This generalization turns out to be quite straightforward and can be extended to a wider class of processes, including a Markov-modulated Cramér-Lundberg process as in [25]. A Sparre Andersen risk process is a process of the form (2.1), where $N(t)$ is an independent renewal process with interarrival distribution $A(x)$ concentrated on $(0, \infty)$ (see e.g. Asmussen and Albrecher [7]).

We assume that $A(x)$ is of phase type with parameters $(n, \boldsymbol{\alpha}, A)$, which means that $A(x)$ is the distribution of the life time of a transient Markov chain (MC) on n states with transition rate matrix A and initial distribution given by a row vector $\boldsymbol{\alpha}$. Strictly speaking, these n transient states should be complemented by an absorbing state. Letting $\boldsymbol{a} = -A\mathbf{1}$ be the column vector of transition rates into the absorbing state, one can identify the PH density by $f(x) = \boldsymbol{\alpha}e^{Ax}\boldsymbol{a}$ for $x \geq 0$.

Let $\phi(x, i)$ be the survival probability assuming that at time 0 the process is at the level x and in phase i , i.e. the MC representing the interarrival times is in state i . Then the survival probability for the original model is given by $\boldsymbol{\alpha}\boldsymbol{\phi}(x)$, where $\boldsymbol{\phi}(x)$ is a column vector of $\phi(x, i)$. Let $\boldsymbol{p}(x)$ be a row vector composed of $p(x, j)$, where $p(x, j)$ is the probability that an excursion started at x regains the original level in the phase j without causing ruin. Then Equation (2.2) reads

$$c(1 - \gamma)\boldsymbol{\phi}'_{\gamma}(x) + (A + \boldsymbol{a}\boldsymbol{p}(x))\boldsymbol{\phi}_{\gamma}(x) = \mathbf{0}.$$

To see this we write up to $o(h)$ terms

$$\phi_{\gamma}(x, i) = (1 + a_{ii}\frac{h}{\tilde{c}})\phi_{\gamma}(x+h, i) + a_{ij}\frac{h}{\tilde{c}}\phi_{\gamma}(x+h, j) + a_i\frac{h}{\tilde{c}}\sum_j p(x, j)\phi_{\gamma}(x+h, j)$$

with $\tilde{c} = c(1 - \gamma)$, which immediately leads to the above differential equation. In other words, we obtain

$$\boldsymbol{\phi}'_{\gamma}(x) + \frac{1}{1 - \gamma}\Lambda(x)\boldsymbol{\phi}_{\gamma}(x) = \mathbf{0}, \quad (3.1)$$

where $\Lambda(x) = (A + \boldsymbol{a}\boldsymbol{p}(x))/c$. Interestingly, the matrix $\Lambda(x)$ has a very nice probabilistic interpretation, which we later exploit in the analysis of

a spectrally-negative MAP. This matrix can be identified using scale matrices, the analogues of scale functions of a spectrally-negative Lévy process, see Section 5.

The solution to Equation (3.1) in one dimension has an exponential form, which then leads to the power relation between taxed and original survival probabilities. This is not the case in higher dimensions, because matrices $\Lambda(x), x \geq 0$ do not commute in general. The solution can be expressed using so-called product integrals, see [14, 16]. Our numerical experiments with a simple two-dimensional model indicate that various power relations indeed fail to hold, see Section 7.

Finally, we only indicate the form of a differential equation for the taxed Gerber-Shiu function $\alpha \mathbf{m}_\gamma(x)$, where $\mathbf{m}_\gamma(x)$ is a column vector composed from $m_\gamma(x, i)$. Similarly to (3.1) we get

$$(1 - \gamma) \mathbf{m}'_\gamma(x) + \Lambda(x) \mathbf{m}_\gamma(x) + \mathbf{r}_w(x) = \mathbf{0},$$

where $r_w(x, i)$ is a penalty term resulting from immediate ruin. Since the vector $\mathbf{r}_w(x)$ again does not depend on taxation, one arrives at

$$(1 - \gamma) \mathbf{m}'_\gamma(x) + \Lambda(x) \mathbf{m}_\gamma(x) - \mathbf{m}'_0(x) - \Lambda(x) \mathbf{m}_0(x) = \mathbf{0}.$$

4 Spectrally-negative Markov additive processes as an umbrella of various risk processes

In this section we present a definition of a MAP and show that a number of important risk processes can be represented using spectrally-negative MAPs, that is, MAPs which have negative jumps only. Strictly speaking, a MAP is a bivariate process $(X(t), J(t))$, where $J(t)$ is a Markov chain (MC) representing an exogenous background process, and $X(t)$ is a so-called additive component modulated by $J(t)$. Nonetheless we often say MAP meaning $X(t)$, so for example a spectrally-negative MAP is a MAP whose additive component can have only negative jumps. A MAP is a generalization of a Lévy process in the sense that $X(t)$ has stationary and independent increments conditioned on the state of the modulating process $J(t)$, see [6, Ch. XI] for an introduction to MAPs.

Definition 4.1. *A bivariate process $(X(t), J(t))$ is called MAP if, given $\{J(T) = i\}$, the shifted process $(X(T + t) - X(T), J(T + t))$ is independent*

from $(X(t), J(t)), 0 \leq t \leq T$ and has the same law as $(X(t) - X(0), J(t))$ given $\{J(0) = i\}$ for all i and $T > 0$.

Furthermore, one can replace the deterministic T in the above definition by a stopping time. The resulting property is called the strong Markov property for MAPs.

It is common to assume that $J(t)$ is an irreducible MC with a finite state space $E = \{1, \dots, n\}$, which we do throughout this work. It can be shown that $X(t)$ evolves as some Lévy process $X_i(t)$ while $J(t) = i$, and in addition a transition of $J(t)$ from i to j triggers a jump of $X(t)$ distributed as U_{ij} , where $J(t)$ and all the components in the construction are assumed to be independent. This construction presents an alternative often-used definition of a MAP.

The law of a spectrally-negative MAP is characterized by a certain matrix-valued function $F(\theta)$, which is the analogue of the Laplace exponent of a spectrally-negative Lévy process. In particular, assuming $X(0) = 0$, the following holds true

$$\mathbb{E}[e^{\theta X(t)}; J(t)] = e^{F(\theta)t},$$

where the (i, j) -th element of the matrix on the left is given by

$$\mathbb{E}(e^{\theta X(t)}; J(t) = j | J(0) = i).$$

The $n \times n$ matrix $F(\theta)$ is given by

$$F_{ij}(\theta) = \begin{cases} \psi_i(\theta) + q_{ii}, & \text{if } i = j, \\ q_{ij}G_{ij}(\theta), & \text{if } i \neq j, \end{cases}$$

where q_{ij} are the elements of the transition rate matrix Q of $J(t)$, $\psi_i(\theta) = \log \mathbb{E}e^{\theta X_i(1)}$ is the Laplace exponent of the Lévy process $X_i(t)$, and $G_{ij}(\theta) = \mathbb{E}e^{\theta U_{ij}}$. Let us now consider some special cases of spectrally-negative MAPs relevant in risk theory.

Example 4.1. *A Markov-modulated Cramér-Lundberg process with tax was considered in [25]. This model is retrieved from a spectrally-negative MAP $(X(t), J(t))$ by putting $U_{ij} = 0$ and $X_i(t) = c_i t - \sum_{k=1}^{N_i(t)} C_k$, where $N_i(t)$ is a Poisson process of intensity λ_i and C_k are iid positive random variables. Hence $G_{ij}(\theta) = 1$ and $\psi_i(\theta) = c_i \theta - \lambda_i(1 - \mathbb{E}e^{-\theta C_1})$.*

Example 4.2. *A Sparre Andersen renewal risk process with PH interarrival times, as considered in Section 3, is another special case of a spectrally-negative MAP $(X(t), J(t))$. Here the MC $J(t)$ lives on n states. It makes a*

jump from i to j without causing a jump of $X(t)$ with rate a_{ij} ; it makes a jump from i to j and triggers a jump $-C_k$ with rate $a_i\alpha_j$. Hence for $i \neq j$ it holds that $F_{ij}(\theta) = a_{ij} + a_i\alpha_j\mathbb{E}e^{-\theta C_1}$, that is $q_{ij} = a_{ij} + a_i\alpha_j$ and U_{ij} is an appropriate mixture of 0 and $-C_1$. Then $F_{ii}(\theta) = a_{ii} + a_i\alpha_i\mathbb{E}e^{-\theta C_1} + c\theta$, because $q_{ii} = -\sum_{j \neq i} q_{ij} = a_{ii} + a_i\alpha_i$ and $\psi_i(\theta) = c\theta - a_i\alpha_i(1 - \mathbb{E}e^{-\theta C_1})$ which corresponds to a compound Poisson process with intensity $a_i\alpha_i$, jumps distributed as $-C_1$ and drift c . In matrix notation we have

$$F(\theta) = A + \mathbf{a}\boldsymbol{\alpha}\mathbb{E}e^{-\theta C_1} + c\theta\mathbb{I}_n,$$

where \mathbb{I}_n is an $n \times n$ identity matrix. Finally, we note that $J(0)$ should be distributed according to $\boldsymbol{\alpha}$.

Finally, we illustrate that it is also possible to incorporate positive jumps of phase type into the model. To do so we enlarge the state space of the background process and replace PH jumps by linear stretches of the unit slope. This procedure is commonly known as fluid embedding.

Example 4.3. One can extend the model of Section 3 to allow more general jumps. We assume that C_k is a mixture of C_k^- with probability p and $-C_k^+$ with probability $1 - p$, where C_k^- are any iid positive random variables and C_k^+ are iid positive random variables of phase type. Let $\boldsymbol{\beta}, B, \mathbf{b}$ be the initial distribution, the transition rate matrix and the vector of rates into the absorbing state corresponding to this PH distribution having, say, n_b phases. The corresponding spectrally-negative MAP $(X(t), J(t))$ is constructed as follows. We let $n = n_a + n_b$, where the first n_a phases represent the PH interarrival times. The matrix $F(\theta)$ can be written in a block form as

$$F(\theta) = \begin{pmatrix} A + p\mathbf{a}\boldsymbol{\alpha}\mathbb{E}e^{-\theta C_1^-} + c\theta\mathbb{I}_{n_a} & (1-p)\mathbf{a}\boldsymbol{\beta} \\ \mathbf{b}\boldsymbol{\alpha} & B + \theta\mathbb{I}_{n_b} \end{pmatrix}$$

and $J(0)$ is to be distributed according to $(\boldsymbol{\alpha}, \mathbf{0})$. It is noted that the original process is retrieved by the time change, so that the time when $J(t) > n_a$ is not counted. This can be achieved by the state-dependent killing as discussed in Remark 6.4.

5 Review of some exit problems

Consider a spectrally-negative MAP $(X(t), J(t))$ and assume that none of the underlying Lévy processes $X_i(t)$ is a.s. non-increasing. This assumption is satisfied by all the examples given in Section 4, and allows to simplify the notation. Define the first passage times as follows

$$\tau_a^+ = \inf\{t \geq 0 : X(t) > a\}, \quad \tau_b^- = \inf\{t \geq 0 : X(t) < -b\},$$

where $a, b \geq 0$. The two-sided exit problem concerns identification of the matrix $\mathbb{P}[\tau_a^+ < \tau_b^-, J(\tau_a^+)]$, whose ij th element is given by $\mathbb{P}[\tau_a^+ < \tau_b^-, J(\tau_a^+) = j | J(0) = i]$. The following result solving this problem was proven in [18]:

Theorem 5.1. *There exists a unique continuous function $W : [0, \infty) \rightarrow \mathbb{R}^{n \times n}$ such that $W(x)$ is invertible for all $x > 0$,*

$$\mathbb{P}[\tau_a^+ < \tau_b^-, J(\tau_a^+)] = W(b)W(a+b)^{-1} \text{ for all } a, b \geq 0 \text{ with } a+b > 0,$$

and

$$\int_0^\infty e^{-\theta x} W(x) dx = F(\theta)^{-1}$$

for all sufficiently large $\theta > 0$.

The matrix-valued function $W(x)$ is a generalization of a scale function of a spectrally-negative Lévy process, and hence we call it a scale matrix.

In the framework of spectrally-negative MAPs it is often advantageous to identify certain associated MCs. For example, using the Definition 4.1 of a MAP with $T = \tau_x^+$ we see that the process $J(\tau_x^+)$ is itself a MC as a function of $x \geq 0$. We denote its transition rate matrix by Λ , so that $\mathbb{P}[J(\tau_x^+)] = e^{\Lambda x}$. This MC can be transient, that is, strictly speaking, we may need to add an absorbing state. In this case the life time of $J(\tau_x^+)$, which coincides with the overall supremum of $X(t)$, has a PH distribution characterized by the matrix Λ .

Next consider the reflected process $Y(t) = X(t) - \bar{X}(t)$, where $\bar{X}(t)$ is the running supremum. Roughly speaking, this process depicts excursions of $X(t)$ from its maximum. Suppose we kill $X(t)$ (send it to some absorbing state) at the stopping time

$$T_a = \inf\{t \geq 0 : Y(t) < -a\},$$

i.e. at the first time when the height of an excursion from the maximum exceeds $a > 0$. Again, using the strong Markov property for MAPs, we see that $J(\tau_x^+), x \geq 0$ is a MC also in the presence of killing. We denote its transition rate matrix by $\Lambda(a)$, so that $\mathbb{P}[\tau_x^+ < T_a; J(\tau_x^+)] = e^{\Lambda(a)x}$. It was shown in [18] that for $a > 0$ the right and left derivatives $W'_+(a)$ and $W'_-(a)$ exist and

$$\Lambda(a) = -W'_+(a)W(a)^{-1}, \quad \Lambda(a-) = -W'_-(a)W(a)^{-1}, \quad (5.1)$$

which generalizes the well-known identities in the theory of Lévy processes.

Remark 5.1. *We remark that in the case of a Lévy process the life time of the transient MC characterized by $\Lambda(a)$ reduces to an exponential random variable. Hence $\bar{X}(T_a)$ is exponentially distributed with rate $W'_+(a)/W(a)$. This quantity can be seen as the total amount of dividends paid until ruin in a Lévy model with the barrier dividend strategy, where the initial capital and the barrier are both placed at the level a .*

Let us briefly discuss the proof of the left identity in (5.1). These ideas will be important in the analysis of the taxed process. One starts by establishing a bound for any $0 < h < \delta$:

$$\mathbb{P}[\tau_h^+ < T_a; J(\tau_h^+)] \leq \mathbb{P}[\tau_h^+ < \tau_a^-; J(\tau_h^+)] \leq \mathbb{P}[\tau_h^+ < T_{a+\delta}; J(\tau_h^+)], \quad (5.2)$$

which follows directly from a sample path analysis. Then

$$\Lambda(a) \leq \lim_{h \downarrow 0} (\mathbb{P}[\tau_h^+ < \tau_a^-; J(\tau_h^+)] - \mathbb{I})/h \leq \Lambda(a + \delta),$$

because $\mathbb{P}[\tau_h^+ < T_a; J(\tau_h^+)] = \mathbb{I} + \Lambda(a)h$ up to $o(h)$ terms. Finally, from the probabilistic interpretation it follows that $\Lambda(a + \delta) \rightarrow \Lambda(a)$ as $\delta \rightarrow 0$, and so we obtain

$$W(a)W(a + h)^{-1} = \mathbb{P}[\tau_h^+ < \tau_a^-; J(\tau_h^+)] = \mathbb{I} + \Lambda(a)h \quad (5.3)$$

as $h \downarrow 0$ up to $o(h)$ terms.

6 The taxed process

In the setting of MAP-driven risk processes it is important that the taxation rate is allowed to depend on the state of the environment process $J(t)$. Hence we consider a vector $\boldsymbol{\gamma}(x) = (\gamma_1(x), \dots, \gamma_n(x))$ of taxation rates and define a diagonal matrix $\Gamma(x)$ with $1/(1 - \gamma_i(x))$ on the diagonal, where all $\gamma_i(x) \in [0, 1)$. So the taxed version of $X(t)$ is defined by

$$X_\gamma(t) = X(t) - \int_0^t \gamma_{J(s)}(\bar{X}(s)) d\bar{X}(s) = X(t) - \int_0^{\bar{X}(t)} \gamma_{J(\tau_x^+)}(x) dx.$$

Similarly, to the one-dimensional case we choose to illustrate the ideas for $\gamma_i(x) \equiv \gamma_i$ for all i . Denoting the law of the taxed process by \mathbb{P}^γ we present following important result.

Lemma 6.1. *It holds that*

$$\mathbb{P}^\gamma[\tau_h^+ < \tau_a^-; J(\tau_h^+)] = \mathbb{I} + \Gamma\Lambda(a)h$$

up to $o(h)$ terms as $h \downarrow 0$.

Proof. Consider $\mathbb{P}^\gamma[\tau_h^+ < T_a; J(\tau_h^+)]$, where T_a does not depend on taxation. Firstly, observe that

$$\mathbb{P}^\gamma[\tau_h^+ < T_a; J(\tau_h^+) = i | J(0) = i] = 1 + \lambda(a)_{ii} \frac{h}{1 - \gamma_i} + o(h),$$

because h under taxation corresponds to $h/(1-\gamma_i)$ without taxation assuming that $J(\tau_x^+)$ does not jump. Secondly, for $i \neq j$ we get

$$\mathbb{P}^\gamma[\tau_h^+ < T_a; J(\tau_h^+) = j | J(0) = i] = \lambda(a)_{ij} \frac{h}{1 - \gamma_i} + o(h),$$

because the probability of two or more jumps of $J(\tau_x^+)$ in the interval $\max_i h/(1-\gamma_i)$ is of order $o(h)$. This can be summarized in matrix notation up to $o(h)$ terms as follows

$$\mathbb{P}^\gamma[\tau_h^+ < T_a; J(\tau_h^+)] = \mathbb{I} + \Gamma\Lambda(a)h.$$

Finally, the same bound as in (5.2) applies to taxed processes, which then yields the result. \square

Let us now consider the main quantity of interest, the matrix $\Phi_\gamma(x, y)$ defined by

$$\Phi_\gamma(x, y) = \mathbb{P}_x^\gamma[\tau_y^+ < \tau_0^-; J(\tau_y^+)],$$

where $0 \leq x \leq y$. This quantity is more convenient to work with than the survival vector. Moreover, the latter can be obtained from $\Phi_\gamma(x, y)$ by taking the limit $y \rightarrow \infty$. Using the strong Markov property for MAPs, it is not difficult to see that $\Phi_\gamma(x, y)$ are transition probability matrices of a certain (non-stationary) Markov process with a finite number of states. In the following we derive Kolmogorov's equations characterizing these transition probabilities.

It can be seen that $\Phi_\gamma(x, y)$ is continuous in both x and y , because a MAP as considered in the present paper can not touch a fixed level without immediately passing it a.s., see [18]. Furthermore, Lemma 6.1 shows that $\Phi_\gamma(x, x+h) = \mathbb{I} + \Gamma\Lambda(x)h$ up to $o(h)$ terms. Using the strong Markov property for MAPs we write

$$\Phi_\gamma(x, y+h) = \Phi_\gamma(x, y)\Phi_\gamma(y, y+h),$$

which then results in the following theorem.

Theorem 6.1. *The transition probability matrices $\Phi_\gamma(x, y)$ solve the Kolmogorov's forward equation*

$$\frac{\partial_+}{\partial y} \Phi_\gamma(x, y) = \Phi_\gamma(x, y)\Gamma\Lambda(y) = -\Phi_\gamma(x, y)\Gamma W'_+(y)W(y)^{-1} \text{ for } y \geq x \quad (6.1)$$

with the initial condition $\Phi_\gamma(x, y) = \mathbb{I}$. For $\partial_-/\partial y$ one needs to replace $\Lambda(y)$ and $W'_+(y)$ with $\Lambda(y-)$ and $W'_-(y)$.

Similarly, one can obtain Kolmogorov's backward equation

$$\frac{\partial}{\partial x} \Phi_\gamma(x, y) = -\Gamma\Lambda(x)\Phi_\gamma(x, y) = \Gamma W'(x)W(x)^{-1}\Phi_\gamma(x, y) \text{ for } 0 \leq x \leq y \quad (6.2)$$

with the initial condition $\Phi_\gamma(y, y) = \mathbb{I}$.

Remark 6.1. *The above non-stationary MC can be seen as an extension of ideas from Section 2, where the killing event was identified with the first point of a nonhomogenous Poisson process. Note also that our proof identifies the transition rate matrix $\Gamma\Lambda(y)$ defining this non-stationary Markov process with the transition rate matrix of the stationary taxed MC $J(\tau_x^+)$, $\tau_x < T_y$.*

Remark 6.2. *In many cases the matrix $W(x)$ is continuously differentiable, and so in particular $W'_-(x) = W'_+(x)$ for all $x > 0$. This is for instance the case when all the underlying Lévy processes $X_i(t)$ satisfy the following assumption: Either $X_i(t)$ is of unbounded variation on compacts, or the jump measure of $X_i(t)$ as well as the distributions of U_{ij} for all j have no atoms.*

One can check that in the case of no taxation the Equations (6.1) and (6.2) indeed admit the solution $\Phi_0(x, y) = W(x)W(y)^{-1}$. In one dimension, i.e. in the case of a spectrally-negative Lévy, process we get an exponential function leading to a power relation between $\Phi_\gamma(x, y)$ and $\Phi_0(x, y)$:

$$\Phi_\gamma(x, y) = e^{\int_x^y \frac{1}{1-\gamma} \Lambda(z) dz} = \Phi_0(x, y)^{\frac{1}{1-\gamma}}.$$

In higher dimensions $\Phi_\gamma(x, y)$ can be expressed using so-called product integrals, see [14, 16]. The problem arises from the fact that $\Gamma\Lambda(a)$ and $\Gamma\Lambda(b)$ do not commute in general. Our numerical experiments with a simple two-dimensional model and constant taxation rate γ indicate that various power relations (inspired by the one-dimensional case) between $\Phi_\gamma(x, y)$ and $\Phi_0(x, y)$ fail to hold.

Remark 6.3. *It is possible to show that $\Phi_\gamma(x, y) = W_\gamma(x)W_\gamma(y)^{-1}$ for some matrix-valued function $W_\gamma(\cdot)$; in one dimension it is given by $W(\cdot)^{\frac{1}{1-\gamma}}$. Then Equation (6.2) shows that $W'_\gamma(x) = -\Gamma\Lambda(x)W_\gamma(x)$, which is the same differential equation as (6.2), but with a different initial condition. It is unlikely that $W_\gamma(x)$ can be identified via its transform similarly to Theorem 5.1.*

Remark 6.4. *The above analysis can be applied in exactly the same form to an exponentially killed MAP. This then yields Kolmogorov's equations for $\mathbb{E}_x^\gamma[e^{-\delta\tau_y^+}; \tau_y^+ < \tau_0^-, J(\tau_y^+)]$, where $\delta > 0$ is the rate of killing. In certain cases, as e.g. in Example 4.3, one may need to use state-dependent killing. This allows to 'ignore' the time spent in auxiliary states representing positive jumps of phase type. In terms of the matrix $F(\theta)$ characterizing the MAP, killing amounts to replacing $F(\theta)$ with $F(\theta) - \delta\mathbb{I}$, or rather $F(\theta) - \delta D$, where D is a diagonal matrix with 0 and 1 on the diagonal depending if the corresponding state is auxiliary or not. Some further details about state-dependent killing of MAPs can be found in [11, 17].*

Remark 6.5. *Suppose now that the taxation rates $\gamma_i(x)$ are function of x . Assume that $\gamma_i(x) \in [0, 1)$ are right-continuous functions bounded away from 1, so that $1/(1 - \gamma_i(x))$ are right-continuous, too. One can repeat the arguments of Lemma 6.1 to obtain*

$$\mathbb{P}^\gamma[\tau_h^+ < \tau_a^-; J(\tau_h^+)] = \mathbb{I} + \Gamma(a)\Lambda(a)h$$

up to $o(h)$ terms as $h \downarrow 0$. This shows that the Kolmogorov's Equation (6.1) holds with Γ replaced by $\Gamma(y)$.

7 The Sparre Andersen model with PH interarrivals and PH jumps

Let us consider a special case of a spectrally-negative MAP for which the scale matrix $W(x)$ and then also $\Lambda(x)$ can be identified (semi-)explicitly. We reconsider the model from Section 3 and assume that the jumps C_k have PH distribution with n_b phases, transition rate matrix B , the initial distribution β , and $\mathbf{b} = B\mathbf{1}$ (see also [9]). Similarly to Example 4.3, we introduce n_b auxiliary states to represent the jumps, and let $n = n_a + n_b$ and $X_i(t) = -t$ for $i > n_a$. Hence we get a continuous piecewise linear MAP characterized by

$$F(\theta) = \begin{pmatrix} A + (c\theta - \delta)\mathbb{I}_{n_a} & \mathbf{a}\beta \\ \mathbf{b}\alpha & B - \theta\mathbb{I}_{n_b} \end{pmatrix},$$

where $\delta \geq 0$ is the rate of exponential killing. Such a fluid model allows for an explicit identification of $W(x)$.

Let Λ_a and Λ_b be $n_a \times n_a$ and $n_b \times n_b$ transition rate matrices of the MCs $J(\tau_x^+)$ and $J(\tau_x^-)$. Let also $\Pi_{a,b}$ be an $n_a \times n_b$ matrix, such that its i th row contains the distribution $\mathbb{P}_i[J(\tau_0^-)]$. Similarly, we define the $n_b \times n_a$ matrix $\Pi_{b,a}$ as the matrix of initial distributions of $J(\tau_0^+)$. There are various

ways to obtain matrices $\Lambda_a, \Lambda_b, \Pi_{a,b}$ and $\Pi_{b,a}$ from $F(\theta)$ including iterative and spectral methods [5, 13, 23], see also the fundamental works [10, 21] on Wiener-Hopf factorisation for matrices.

In the following we work exclusively with the original n_a dimensional MAP with jumps using matrices Λ, Π, \cdot defined above. The scale matrix $W(x)$ of the original process can be written as $[e^{-\Lambda_a x} - \Pi_{a,b} e^{\Lambda_b x} \Pi_{b,a}]C$, where C is a constant matrix irrelevant in the following, see [8, 17]. Hence we get

$$\Lambda(x) = -W'(x)W(x)^{-1} = [e^{-\Lambda_a x} \Lambda_a + \Pi_{a,b} e^{\Lambda_b x} \Lambda_b \Pi_{b,a}] [e^{-\Lambda_a x} - \Pi_{a,b} e^{\Lambda_b x} \Pi_{b,a}]^{-1}. \quad (7.1)$$

Next, one can use Kolmogorov's Equation (6.1) to numerically solve for $\Phi_\gamma(x, y)$.

7.1 A numerical example

In order to illustrate the effort required to obtain $\Phi_\gamma(x, y)$ we consider a simple example. We assume that the distribution of interarrivals of claims is Erlang(2) with rates λ , and that the claims are exponential with rate μ . So we get $n_a = 2, n_b = 1$ and

$$F(\theta) = \begin{pmatrix} -\lambda - \delta + c\theta & \lambda & 0 \\ 0 & -\lambda - \delta + c\theta & \lambda \\ \mu & 0 & -\mu - \theta \end{pmatrix}.$$

We put $\delta = 0, \lambda = c = 1, \mu = 2$ and use the spectral method to compute matrices Λ and Π, \cdot , which amounts to finding the zeros of $\det(F(\theta))$ and the corresponding null spaces, see e.g. [13]. We find

$$\Lambda_a = \begin{pmatrix} -1 & 1 \\ 0.732051 & -0.732051 \end{pmatrix}, \quad \Pi_{a,b} = \begin{pmatrix} 0.133975 \\ 0.366025 \end{pmatrix}$$

$$\Lambda_b = -1.73205, \quad \Pi_{b,a} = (0.732051, 0.267949).$$

As a check we note that $\Lambda_a \mathbf{1} = \mathbf{0}$, because the process $X(t)$ drifts to $+\infty$ and hence Λ_a is a transition rate matrix of a recurrent MC. Next we use Equation (7.1) to find the 2×2 matrix $\Lambda(x)$.

Choosing taxation rate $\gamma = 0.2$ and initial capital $x = 1$ we numerically solve the system of differential equations in (6.1). In fact, we are only interested in the first row of $\Phi_\gamma(x, y)$, because the initial distribution of the phase process corresponding to the Erlang(2) distribution is $(1, 0)$. Hence we are concerned with a system of 2 differential equations. Summing up

these two solutions we obtain $\mathbb{P}_x^\gamma(\tau_y^+ < \tau_0^-)$, which is depicted in Figure ?? in bold. Its non-taxed version is plotted using the thinner solid line (on top). The dashed line represents $\mathbb{P}_x(\tau_y^+ < \tau_0^-)^{1/(1-\gamma)}$, and the dotted line gives the sum of the first row of $\Phi_0(x, y)^{1/(1-\gamma)}$. The plot illustrates that both the identity $\mathbb{P}_x^\gamma(\tau_y^+ < \tau_0^-) = \mathbb{P}_x(\tau_y^+ < \tau_0^-)^{1/(1-\gamma)}$ and the matrix version $\Phi_\gamma(x, y) = \Phi_0(x, y)^{1/(1-\gamma)}$ of the tax identity indeed do not hold for general MAP risk processes.

Acknowledgement. We thank Dominik Kortschak for useful advice.

References

- [1] H. Albrecher, S. Borst, O.J. Boxma and J. Resing (2009) The tax identity in risk theory – a simple proof and an extension. *Insurance: Mathematics and Economics* 44(2), 304–306.
- [2] H. Albrecher, S. Borst, O.J. Boxma and J. Resing (2011) Ruin excursions, the G/G/ ∞ queue and tax payments in renewal risk models. *Journal of Applied Probability* 48A, 3–14.
- [3] H. Albrecher and C. Hipp (2007) Lundberg’s risk process with tax. *Blätter der DGVM* 28, 13–28.
- [4] H. Albrecher, J. Renaud and X. Zhou (2008) A Lévy insurance risk process with tax. *Journal of Applied Probability* 45(2), 363–375.
- [5] S. Asmussen (1995) Stationary distributions for fluid flow models with or without Brownian noise. *Stochastic Models* 11, 21–49.
- [6] S. Asmussen (2003) *Applied Probability and Queues*. Springer, Berlin.
- [7] S. Asmussen and H. Albrecher (2010) *Ruin Probabilities*. Second Edition. World Scientific, New Jersey.
- [8] F. Avram, M.R. Pistorius and M. Usabel (2003) The two barriers ruin problem via a Wiener Hopf decomposition approach. *Annals of the University of Craiova-Mathematics and Computer Science Series* 30, 38–44.
- [9] F. Avram and M. Usabel (2004) Ruin probabilities for risk processes with phase-type waiting times. *Astin Bulletin* 34 (2), 315–332.

- [10] M.T. Barlow, L.C.G. Rogers and D. Williams (1980) Wiener-Hopf factorization for matrices. *Séminaire de Probabilités XIV*, Lecture Notes in Math. 784, Springer, Berlin, 324–331.
- [11] L. Breuer (2012) Occupation times for Markov-modulated Brownian motion. *Journal of Applied Probability* (in press).
- [12] E.C.K. Cheung and D. Landriault (2012) On a risk model with surplus-dependent premium and tax rates. *Methodology and Computing in Applied Probability* 14(2), 233–251.
- [13] B. D’Auria, J. Ivanovs, O. Kella, and M. Mandjes (2012) First passage of a Markov additive process and generalized Jordan chains. *Journal of Applied Probability* 47(4), 1048–1057.
- [14] R. L. Dobrušin (1953) Generalization of Kolmogorov’s equations for Markov processes with a finite number of possible states. *Mat. Sbornik N.S.* 33(75), 567–596.
- [15] H.U. Gerber (1979) *An Introduction to Mathematical Risk Theory*, Heubner Foundation, Homewood, Illinois.
- [16] G. S. Goodman and S. Johansen (1973) Kolmogorov’s differential equations for non-stationary, countable state Markov processes with uniformly continuous transition probabilities. *Proc. Cambridge Philos. Soc.* 73, 119–138.
- [17] J. Ivanovs (2011) *One-sided Markov additive Processes and Related Exit Problems*, Ph.D. Thesis, Universiteit van Amsterdam.
- [18] J. Ivanovs and Z. Palmowski (2012) Occupation densities in solving exit problems for Markov additive processes and their reflections. *Stochastic Processes and their Applications* (in press).
- [19] L. Ji, C. Zhang and E. Hashorva. On risk processes with two classes of claims and tax. *Preprint, University of Lausanne*.
- [20] A.E. Kyprianou and X. Zhou (2009) General tax structures and the Lévy insurance risk model. *Journal of Applied Probability* 46(4), 1146–1156.
- [21] R.R. London, H.P. McKean, L.C.G. Rogers and D. Williams (1982) A martingale approach to some Wiener-Hopf problems II. *Séminaire de Probabilités XVI*, Lecture Notes in Math. 920, Springer, Berlin, 68–90.

- [22] J. Ren (2007) The discounted joint distribution of the surplus prior to ruin and the deficit at ruin in a Sparre Andersen model. *North American Actuarial Journal* 11(3), 128–136.
- [23] L.C.G. Rogers (1994) Fluid models in queueing theory. *Advances in Applied Probability* 4, 390–413.
- [24] W. Wang, R. Ming and Y. Hu (2010) On the expected discounted penalty function for risk process with tax. *Statistics & Probability Letters* 81(4), 489–501.
- [25] J. Wei, H. Yang and R. Wang (2010) On the Markov-modulated insurance risk model with tax. *Blätter der DGVM* 31 (1), 65–78.