

## APPROXIMATIONS FOR THE GERBER–SHIU EXPECTED DISCOUNTED PENALTY FUNCTION IN THE COMPOUND POISSON RISK MODEL

SUSAN M. PITTS,\* *University of Cambridge*

KONSTADINOS POLITIS,\*\* *University of Piraeus*

### Abstract

In the classical risk model with initial capital  $u$ , let  $\tau(u)$  be the time of ruin,  $X_+(u)$  be the risk reserve just before ruin, and  $Y_+(u)$  be the deficit at ruin. Gerber and Shiu (1998) defined the function  $m_\delta(u) = E[e^{-\delta\tau(u)}w(X_+(u), Y_+(u))\mathbf{1}(\tau(u) < \infty)]$ , where  $\delta \geq 0$  can be interpreted as a force of interest and  $w(r, s)$  as a penalty function, meaning that  $m_\delta(u)$  is the expected discounted penalty payable at ruin. This function is known to satisfy a defective renewal equation, but easy explicit formulae for  $m_\delta(u)$  are only available for certain special cases for the claim size distribution. Approximations thus arise by approximating the desired  $m_\delta(u)$  by that associated with one of these special cases. In this paper a functional approach is taken, giving rise to first-order correction terms for the above approximations.

*Keywords:* Time of ruin; deficit at ruin; surplus prior to ruin; defective renewal equation; functional approach

2000 Mathematics Subject Classification: Primary 60K10; 91B30  
Secondary 60K05

### 1. Introduction

Ruin theory is a traditional focus of interest and research in insurance mathematics, and recent interest has included the study of a particular function that arises in ruin theory, namely Gerber and Shiu's expected discounted penalty function (see, for example, Gerber and Shiu (1998), Lin and Willmot (1999), and Lin and Willmot (2000)). In this paper we consider this function for the compound Poisson risk model. We define the function below, after we have introduced various notation and definitions for this risk model.

Suppose that claims arrive at a Poisson process rate  $\lambda$  and that the claims  $Y_1, Y_2, \dots$  are independent, identically distributed, nonnegative random variables, independent of the claims arrival process, with distribution function  $H$ , mean  $\mu$  (assumed finite), and with the Laplace–Stieltjes transform of  $H$  denoted by  $\tilde{h}(s) = E[e^{-sY_1}] = \int_{[0, \infty)} e^{-st}H(dt)$ . Premium income is assumed to accrue linearly in time at rate  $c > 0$ , and we assume positive safety loading, that is,  $\lambda\mu < c$ . The risk reserve or surplus at time  $t$  is then

$$R(t) = u + ct - \sum_{i=1}^{N(t)} Y_i,$$

Received 9 February 2005; revision received 7 February 2007.

\* Postal address: Statistical Laboratory, Centre for Mathematical Sciences, University of Cambridge, Wilberforce Road, Cambridge CB3 0WB, UK. Email address: s.pitts@statslab.cam.ac.uk

\*\* Postal address: Department of Statistics and Insurance Science, University of Piraeus, 80 Karaoli & Demetriou Street, Piraeus, 185 34, Greece.

where  $u$  is the initial capital and  $N(t)$  is the number of claims arriving up to time  $t$ . The sum is defined to be zero if  $N(t) = 0$ .

The expected discounted penalty function involves several quantities arising in ruin theory. The first of these is  $\tau(u)$ , the *time of ruin* starting from initial capital  $u$ , given by

$$\tau(u) = \inf\{t \geq 0: R(t) < 0\},$$

where we set  $\tau(u) = \infty$  if ruin never occurs, i.e. if  $R(t) \geq 0$  for all  $u \geq 0$ . We observe that the positive safety loading condition implies that  $\psi(u)$ , the probability of ruin with initial capital  $u$ , given by  $\psi(u) = P(\tau(u) < \infty)$ , satisfies  $\psi(u) < 1$  (Asmussen (2000, Corollary III.1.4)). Two further important quantities are the *surplus prior to ruin*,  $X_+(u)$ , given by

$$X_+(u) = R(\tau(u)-),$$

and the *deficit at ruin*,  $Y_+(u)$ , given by

$$Y_+(u) = |R(\tau(u))|.$$

For  $\delta \geq 0$  and a nonnegative function  $w(r, s)$ , the *expected discounted penalty function* is defined to be

$$m_\delta(u) = E[e^{-\delta\tau(u)}w(X_+(u), Y_+(u)) \mathbf{1}(\tau(u) < \infty)],$$

where  $\mathbf{1}(A)$  is the indicator function of the event  $A$ .

The name ‘expected discounted penalty function’ is motivated by thinking of  $\delta$  as a force of interest and of  $w(X_+(u), Y_+(u))$  as a ‘penalty’ incurred at ruin. An alternative interpretation is to think of  $\delta$  as the argument of a Laplace transform. Different choices for the function  $w$  give rise to different quantities of interest in ruin theory. For example, if  $w(r, s) = 1$  for all  $r, s \geq 0$ , then

$$m_\delta(u) = E[e^{-\delta\tau(u)} \mathbf{1}(\tau(u) < \infty)],$$

which is the Laplace transform of  $\tau(u)$  on the event  $\{\tau(u) < \infty\}$ . Still with  $w(r, s) = 1$ , if we take  $\delta = 0$  then we find that  $m_{\delta=0}(u)$  is  $\psi(u)$ , the probability of ruin. Other choices for  $w$  include  $w(r, s) = r^k s^l$  for some  $k, l \geq 0$ , in which case, with  $\delta = 0$ , we find that  $m_{\delta=0}(u) = E[X_+(u)^k Y_+(u)^l \mathbf{1}(\tau(u) < \infty)]$ , giving rise to moments of the surplus just before ruin, and the deficit at ruin. Another set of examples is given by considering functions such as  $w(r, s) = \mathbf{1}(s > y)$  for some fixed  $y \geq 0$ . The corresponding expected discount penalty function when  $\delta = 0$  is then  $m_{\delta=0}(u) = P(Y_+(u) > y, \tau(u) < \infty)$ , thus leading to a consideration of functions related to the distribution function of  $Y_+(u)$ . Similar approaches lead to the distribution function of  $X_+(u)$  and the joint distribution of  $(X_+(u), Y_+(u))$ . See Lin and Willmot (1999), (2000) and Gerber and Shiu (1998), who also discussed an example where an appropriate choice of  $w$  leads to an application in option pricing.

All of the above examples indicate that the expected discounted penalty function includes a wide range of quantities of importance and interest in ruin theory, making it a natural and useful object of study. One fruitful approach is to observe that the function  $m_\delta(u)$  satisfies the defective renewal equation given below, and to use results from renewal theory to provide access to  $m_\delta(u)$ . This approach has been developed in Gerber and Shiu (1998) and Lin and Willmot (1999), (2000), and we give an outline here; see also Willmot and Lin (2001, Section 9.2). The defective renewal equation involves the root  $\rho \equiv \rho(\delta)$  of the equation

$$\lambda \tilde{h}(\rho) - \lambda + c\rho = \delta. \tag{1}$$

There is a unique nonnegative solution  $\rho$  to (1) with  $\rho(0) = 0$  (see, for example, Willmot and Lin (2001, p. 160)). With this  $\rho$ , we have

$$m_\delta(u) = \frac{\lambda}{c} \int_0^u m_\delta(u - t) \int_{(t,\infty)} e^{-\rho(y-t)} H(dy) dt + v(u), \tag{2}$$

where

$$v(u) = \frac{\lambda}{c} \int_u^\infty e^{-\rho(t-u)} \int_{(t,\infty)} w(t, y - t) H(dy) dt. \tag{3}$$

We assume for the moment that  $w \geq 0$  is such that the integrals in  $v$  are finite, although later we will give explicit conditions on  $w$ . Note that  $w$  only appears in (2) via  $v$ . Equation (2) can be written in the form

$$m_\delta(u) = \int_0^u m_\delta(u - t) f(t) dt + v(u), \tag{4}$$

where, for  $t \geq 0$ , we have

$$f(t) = \frac{\lambda}{c} \int_{(t,\infty)} e^{-\rho(y-t)} H(dy). \tag{5}$$

It is easy to see that  $f(t)$  is a nonnegative function with total mass at most  $\lambda \int_0^\infty (1 - H(y)) dy/c$ , and this is less than 1 by the positive safety loading assumption that  $\lambda\mu/c < 1$ . This means that  $f(t)$  is a defective probability density function, and that (4) is a defective renewal equation. Let  $F(t)$  be the distribution function associated with  $f(t)$ . General results for defective renewal equations (for a discussion of these, see Willmot and Lin (2001, Section 9.1)) show that the solution to the defective renewal equation (4) involves  $U(F)$ , the renewal function associated with  $F$ , where

$$U(F) = \sum_{n=0}^\infty F^{*n}.$$

In this definition,  $F^{*n}$  is the  $n$ th Lebesgue–Stieltjes convolution power of  $F$ , and  $F^{*0}(t) = \mathbf{1}(t \geq 0)$ . We note here that  $(U(F))(t)$  is finite for all  $t \geq 0$  and that  $\lim_{t \rightarrow \infty} (U(F))(t) < \infty$ , because  $F$  is defective. With this definition for  $U(F)$ , the solution to the defective renewal equation (4) is given by

$$m_\delta(u) = \int_{[0,u]} v(u - t)(U(F))(dt). \tag{6}$$

Although (6) is indeed a simple representation of  $m_\delta(u)$  in terms of the convolution of  $v$  and  $U(F)$ , it is not always easy to obtain explicit formulae for it. However, for certain choices of  $w$  and  $h$  this can be done. For example, when  $w(r, s) = 1$  and the claim sizes are exponentially distributed, an explicit formula for  $m_\delta(u)$  can be found from the formula given in Willmot and Lin (2001, Example 9.2.1, p. 163), although we note that even this seemingly easy case is not a simple one-line calculation. Other special cases leading to explicit formulae include those where the claim size distribution is a mixture or combination of exponentials, or a mixture of Erlang distributions (all with the same rate parameter); see, for example, Lin and Willmot (1999), (2000).

Now suppose that we wish to find the expected discounted penalty function for another, more ‘difficult’, claim size distribution  $H$ . One approach that is adopted in similar circumstances in

many applications in applied probability is to approximate  $H$  by another distribution, say  $H_0$ , an exponential distribution with the same mean as  $H$ , for which the expected discounted penalty function is explicitly known. The expected discounted penalty function for  $H_0$  is then used as an approximation to the desired expected discounted penalty function, for  $H$ . Since the approximating expected discounted penalty function has an explicit formula, we note that this approximation is an approximation formula, not just an approximation. We write  $\Phi$  for the functional or map that takes  $H$  onto  $m_\delta(u)$ , so that this approximation may be represented in rough terms as

$$\Phi(H) \approx \Phi(H_0).$$

If the functional  $\Phi$  is continuous in an appropriate sense, then the approximating expected discounted penalty function will be ‘close’ to the desired one when  $H$  is ‘close’ to  $H_0$ , that is,

$$\text{dist}_2(\Phi(H), \Phi(H_0)) \text{ is small when } \text{dist}_1(H, H_0) \text{ is small,}$$

where  $\text{dist}_1$  denotes an appropriate distance between the two input distributions and  $\text{dist}_2$  denotes a distance between the two output functions.

Suppose now that, in addition to being continuous, the functional  $\Phi$  is differentiable in an appropriate sense, with derivative  $\Phi'_{H_0}$  at  $H_0$ . Then the above approximation, which we call the *zeroth-order approximation*, may be refined or corrected to obtain a new approximation, the *first-order approximation*, given, again in rough terms, by

$$\Phi(H) \approx \Phi(H_0) + \Phi'_{H_0}(H - H_0). \quad (7)$$

If the derivative  $\Phi'_{H_0}$  can be explicitly obtained (which is quite likely, since  $\Phi'_{H_0}$  will be a linear map), then the first-order approximation will again be an approximation formula.

From the above ‘rough terms’ description, we see that this method of approximation, sometimes called the *functional approach*, may in principle be applied to many different stochastic models, and it has been successfully applied in several different areas of applied probability (Grübel (1989), Grübel and Pitts (1992), Politis and Pitts (1998), Pitts (1998), (1999), (2004)).

In this paper we define the functional  $\Phi$  that takes as input the claim size distribution and produces as output the corresponding expected discounted penalty function. In order to make this definition precise, we need to specify the spaces that are to form the domain and codomain for  $\Phi$ , including appropriate topologies, and this is done in Subsection 2.1. Our main result is to show that  $\Phi$  is appropriately differentiable, and to find an expression for its derivative at  $H_0$ . This result is stated in Subsection 2.2, where its use in obtaining new first-order approximations to  $m_\delta(u)$  is outlined. Section 3 contains a particular example of how the approximation works in practice. The proof of the main theorem is given via a series of intermediate results, in Section 4, and Section 5 contains conclusions and remarks.

## 2. Results

### 2.1. The functional $\Phi$

In this subsection we give a formal definition of  $\Phi$  by expressing it in terms of several simpler maps. As part of this definition, we need to specify relevant spaces for the domains and codomains of  $\Phi$  and of the intermediate maps. Our aim is to use spaces for which the functionals are all differentiable, but which are still probabilistically appealing.

The domain for the functional  $\Phi$  has to contain the claim size distribution. However, from (7) we see that when we apply the functional approach we will evaluate the derivative of  $\Phi$  at

the difference between two probability measures, a quantity which is not itself a probability measure. Hence, a more natural space for the domain of  $\Phi$  is the space,  $M$ , of finite, signed measures on the Borel sets  $\mathcal{B} \equiv \mathcal{B}([0, \infty))$  of  $[0, \infty)$ . If  $a$  is an element of  $M$  then let  $|a|$  be the total variation measure of  $a$  given by  $|a|(E) = \sup \sum_{i=1}^{\infty} |a(E_i)|$ , where  $E$  is in  $\mathcal{B}$  and the supremum is taken over all partitions  $\{E_i\}$  of  $E$  in  $\mathcal{B}$ . The total variation norm of  $a$  is  $\|a\|_{TV} = |a|([0, \infty))$ , and it can be shown that  $(M, \|\cdot\|_{TV})$  is a Banach space. See Rudin (1986, Chapter 6) for details of the space of finite, complex measures on  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ , of which  $M$  is a subspace, and also Dunford and Schwartz (1958, Section III.7.4) for Banach space properties.

Recall that we assume throughout that the claim sizes have finite first moments. In some of the intermediate stages of our decomposition of  $\Phi$  we need various ‘moments’ of the measures to be finite. This motivates the introduction of the spaces

$$M_k = \left\{ a \in M : \int_{[0, \infty)} t^k |a|(dt) < \infty \right\},$$

for  $k \geq 0$ . For  $a \in M_k$ , let  $\|a\|_k = \int_{[0, \infty)} (1+t)^k |a|(dt)$ . It can be checked that  $(M_k, \|\cdot\|_k)$  is a Banach space.

From (3) and (5), both  $v(u)$  and  $f(t)$  involve the quantity  $\rho$  satisfying (1), so we first consider the functional that takes a finite, signed measure to a quantity corresponding to  $\rho$ . Let

$$A = \left\{ a \in M_1 : \int_{[0, \infty)} t |a|(dt) < \frac{c}{\lambda} \right\}.$$

For a proper probability measure  $h$ , the condition  $h \in A$  is the same as the condition of positive safety loading, so  $A$  is a natural set to consider. For  $a \in M$ , let  $\tilde{a}(t) = \int_{[0, \infty)} e^{-ty} a(dy)$  and define the function  $g_a : t \mapsto \lambda \tilde{a}(t) - \lambda \tilde{a}(0) + ct$  for  $t \geq 0$ . The function  $g_a$  is continuous (right continuous at 0), maps 0 onto 0, and has derivative  $\lambda \tilde{a}'(t) + c$ . For  $a \in A$ ,  $|\lambda \int_{[0, \infty)} ye^{-ty} a(dy)| \leq \lambda \int_{[0, \infty)} y |a|(dy) < c$  and, so, for all  $t \geq 0$ ,

$$g'_a(t) = \lambda \tilde{a}'(t) + c = c - \lambda \int_{[0, \infty)} ye^{-ty} a(dy) \geq c - \lambda \int_{[0, \infty)} y |a|(dy) > 0. \tag{8}$$

Hence,  $g_a$  is invertible, and for  $a \in A$  we define the functional  $\Psi$ , with  $\Psi(a) = \rho_a$ , such that, for  $\delta \geq 0$ ,

$$(\Psi(a))(\delta) = \rho_a(\delta) = (\lambda \tilde{a} - \lambda \tilde{a}(0) + c\iota)^{-1}(\delta) = g_a^{-1}(\delta), \tag{9}$$

where  $\iota : t \mapsto t$ . For a proper probability measure  $h$ , (1) shows that  $\Psi(h)$  is the required function  $\rho(\cdot)$ . The function  $\rho_a$  has  $\rho_a(0) = 0$  and is increasing and continuous, with  $\rho_a(\delta) \rightarrow \infty$  as  $\delta \rightarrow \infty$ , meaning that  $\rho_a$  is unbounded on  $[0, \infty)$ . Hence, we restrict to  $\delta \in [0, K]$  for some fixed, positive  $K < \infty$ , in which case  $\rho_a : \delta \mapsto \rho_a(\delta)$ ,  $\delta \in [0, K]$ , is in the space,  $C_{[0, K]}$ , of bounded, continuous functions on  $[0, K]$ . We give this space the supremum norm. In summary, we have

$$\Psi : \begin{cases} A \rightarrow C_{[0, K]}, \\ a \mapsto \rho_a. \end{cases}$$

For later component functionals, we use special cases of the space

$$\mathcal{E} = \{\varphi : [0, K] \rightarrow (B, \|\cdot\|), \varphi \text{ continuous}\},$$

where  $(B, \|\cdot\|)$  is a Banach space. The continuity of  $\varphi$  means that  $\|\varphi(\delta) - \varphi(\delta_1)\|$  tends to 0 as  $\delta$  tends to  $\delta_1$ . For  $\varphi \in \mathcal{E}$ , define  $\|\varphi\|_{\mathcal{E}} = \sup_{\delta \in [0, K]} \|\varphi(\delta)\|$ . It can be shown that  $(\mathcal{E}, \|\cdot\|_{\mathcal{E}})$  is a Banach space. Let  $(\mathcal{D}, \|\cdot\|_{\mathcal{D}})$  be the resulting  $\mathcal{E}$ -space if  $B$  is  $D_{[0, \infty]}$ , the space of right-continuous functions on  $[0, \infty]$  with left-hand limits, with supremum norm. Let  $(\mathcal{C}, \|\cdot\|_{\mathcal{C}})$  be the resulting  $\mathcal{E}$ -space when  $B$  is  $C_{[0, \infty]}$ , the space of continuous functions on  $[0, \infty]$  with supremum norm, so that  $\mathcal{C}$  is contained in  $\mathcal{D}$ . Let  $(\mathcal{M}, \|\cdot\|_{\mathcal{M}})$  be the  $\mathcal{E}$ -space that results if  $(B, \|\cdot\|)$  is  $(M, \|\cdot\|_{TV})$ .

For future reference, for  $a$  and  $b$  in  $\mathcal{M}$  we define  $a * b$  to be the element of  $\mathcal{M}$  given by

$$(a * b)(\delta) = a(\delta) * b(\delta),$$

meaning that  $(\mathcal{M}, \|\cdot\|_{\mathcal{M}}, *)$  is a commutative (real) Banach algebra with identity  $e$  given by  $e(\delta) = e_0$  for all  $\delta$ , where  $e_0$  is the measure with unit mass at 0. Note that we use the same notation for convolution in  $M$  and in  $\mathcal{M}$ , but the meaning should always be clear from the context. Also for future reference, we note that the element  $a$  in  $\mathcal{M}$  is invertible if there is an element  $a^{*(-1)}$  in  $\mathcal{M}$  such that  $a * a^{*(-1)} = e$ ; then  $a^{*(-1)}$  is called the inverse of  $a$ .

Care is needed in interpreting expressions which have  $\delta$  as an argument. If  $f$  is in  $\mathcal{C}$  or  $\mathcal{D}$  then  $f(\delta)$  is a function on  $[0, \infty)$ , and it makes sense to write  $(f(\delta))(u)$  for  $u \geq 0$ . If  $b$  is in  $\mathcal{M}$  then  $b(\delta)$  is a finite, signed measure, and it makes sense to write  $(b(\delta))(E)$  for a Borel set  $E$ . This is in contrast to  $\rho_a(\delta)$  defined in (9), where  $\rho_a(\delta)$  is in  $\mathbb{R}$  for each  $\delta$  in  $[0, K]$ . Thus,  $\rho_a(\delta)(y - t)$  means the product of two real numbers,  $\rho_a(\delta)$  and  $(y - t)$ .

We now turn to the component map that takes a finite, signed measure onto an appropriate counterpart of  $v$ , which is given for a probability measure in (3). This equation involves the function  $w(r, s)$ , so we need now to be precise about conditions on  $w$ . We assume from now on that  $w$  is measurable and that there exist nonnegative  $k$  and  $l$ , and a finite constant  $C_w$ , such that

$$0 \leq w(r, s) \leq C_w r^k s^l \quad \text{for all } r, s \geq 0. \tag{10}$$

All the examples for  $w(r, s)$  in Section 1 satisfy (10), and any nonnegative, bounded  $w$  satisfies this condition with  $k = l = 0$ .

For the counterpart of the  $v$ -map, for  $a \in A \cap M_{k+l+1}$  and each  $\delta \in [0, K]$  we define the function  $v_a(\delta)$  by

$$(v_a(\delta))(u) = \frac{\lambda}{c} \int_u^\infty e^{-\rho_a(\delta)(t-u)} \int_{(t, \infty)} w(t, y - t) a(dy) dt, \tag{11}$$

for  $u \geq 0$ . If  $w$  satisfies (10) and  $a \in M_{k+l+1}$ , then we have

$$\begin{aligned} \int_0^\infty \int_{(t, \infty)} w(t, y - t) |a|(dy) dt &\leq C_w \int_0^\infty \int_{(t, \infty)} t^k y^l |a|(dy) dt \\ &\leq \frac{C_w}{k + 1} \int_{(0, \infty)} y^{k+l+1} |a|(dy) \\ &< \infty. \end{aligned}$$

Thus, if  $a$  is in  $M_{k+l+1}$  then  $v_a(\delta)$  is bounded and, hence, in  $C_{[0, \infty]}$ . It can be checked that  $v_a(\delta)$  is continuous in  $\delta$  (as a map from  $[0, K]$  to  $C_{[0, \infty]}$ ), which means that  $v_a$  is in  $\mathcal{C}$ . Note that we need  $a \in A$  because the expression on the right-hand side of (11) involves  $\rho_a$ . Summarising the above, we write  $\Phi_1$  for the  $v$ -functional, and have

$$\Phi_1: \begin{cases} A \cap M_{k+l+1} \rightarrow \mathcal{C}, \\ a \mapsto v_a. \end{cases}$$

From (6), we need to consider counterparts of the defective probability measure whose density is given in (5) and its associated renewal measure. For  $a$  in  $A \subseteq M_1$  and  $\delta \in [0, K]$ , let  $f_a(\delta)$  be the measure on  $([0, \infty), \mathcal{B})$  given by

$$(f_a(\delta))(E) = \int_E \frac{\lambda}{c} \int_{(t, \infty)} e^{-\rho_a(\delta)(y-t)} a(dy) dt \tag{12}$$

for all  $E$  in  $\mathcal{B}$ . Then, since  $a$  is in  $M_1$ , we know that  $f_a(\delta)$  is a finite, signed measure, and we can check that  $f_a$  is in  $\mathcal{M}$ . For the proper probability measure  $h$  corresponding to the claim size distribution,  $f_h(\delta)$  is a defective probability measure whose density is the same as the density in (5). Let  $\Psi_1$  be the functional defined by

$$\Psi_1: \begin{cases} A \rightarrow \mathcal{M}, \\ a \mapsto f_a. \end{cases}$$

For general  $a$  in  $A$ ,

$$\sup_{\delta \in [0, K]} \|f_a(\delta)\|_{TV} \leq \frac{\lambda}{c} \int_{[0, \infty)} y|a|(dy) < 1,$$

meaning that  $\|f_a\|_{\mathcal{M}} < 1$ . Using Rudin (1986, Theorem 18.3), we see that  $e - f_a$  is invertible in  $\mathcal{M}$  and that  $(e - f_a)^{*(-1)} = \sum_{n=0}^{\infty} f_a^{*n}$ , where  $f_a^{*0}$  is the identity element  $e$ . This means that  $((e - f_a)^{*(-1)})(\delta) = \sum_{n=0}^{\infty} f_a(\delta)^{*n} = (e_0 - f_a(\delta))^{*(-1)}$ , where the superscript ‘ $*(-1)$ ’ in the final expression refers to convolution inverse in  $\mathcal{M}$ . For  $a$  in  $A$ , let  $u_a$  in  $\mathcal{M}$  be defined by

$$u_a = (e - f_a)^{*(-1)} = \sum_{n=0}^{\infty} f_a^{*n}. \tag{13}$$

To avoid confusion with the initial surplus  $u$ , notice that the quantity  $u_a$  is in  $\mathcal{M}$ , and always appears with a subscript denoting the input element of  $A$  giving rise to  $f_a$  and, hence, to  $u_a$ . When  $h$  is the claim size distribution,  $u_h(\delta)$  is the (finite) renewal measure associated with the defective probability measure with density given in (5). Let  $\Phi_2$  be defined by

$$\Phi_2: \begin{cases} A \rightarrow \mathcal{M}, \\ a \mapsto u_a. \end{cases}$$

We have now given precise definitions of the functionals that take a finite, signed measure  $a$  in the subset  $A$  of  $M_1$  and map it onto the various quantities needed in order to find the counterpart of the expected discounted penalty function  $m_\delta(u)$  as given in (6). The final step is the convolution of an element of  $\mathcal{C}$  and an element of  $\mathcal{M}$ . Let  $f$  be in  $\mathcal{C}$  and  $b$  in  $\mathcal{M}$ . Then, for  $\delta$  in  $[0, K]$  and  $t \geq 0$ , we define  $f * b$  to be given by  $(f * b)(\delta) = f(\delta) * b(\delta)$ , where

$$(f(\delta) * b(\delta))(t) = \int_{[0, t]} (f(\delta))(t - y)(b(\delta))(dy).$$

We note that this is yet another use of the symbol ‘ $*$ ’, and again the meaning should be clear from the context. For general  $f$  in  $\mathcal{C}$  and  $b$  in  $\mathcal{M}$ , it can be shown that  $f * b$  is in  $\mathcal{D}$ . However, we are interested in applying this with  $f = v_a$  and  $b = u_a$  for  $a$  in  $A \cap M_{k+l+1}$ . The measure  $u_a(\delta)$  has a unit mass at 0 and is absolutely continuous on  $(0, \infty)$ , so  $v_a(\delta) * u_a(\delta)$  is an element of  $C_{[0, \infty]}$ . Thus, in this case,  $v_a * u_a$  is in  $\mathcal{C}$ .

We write  $\Phi$  for the functional that takes  $a$  in  $A \cap M_{k+l+1}$  as input and maps it onto the output  $\Phi(a) = v_a * u_a$  in  $\mathcal{C}$ , that is,

$$\Phi: \begin{cases} A \cap M_{k+l+1} \rightarrow \mathcal{C}, \\ a \mapsto v_a * u_a. \end{cases}$$

For the claim size distribution  $h$ , we see that  $\Phi(h) = v_h * u_h$  is the element  $m$  in  $\mathcal{C}$  given by

$$(m(\delta))(u) = \int_{[0,u]} (v_h(\delta))(u - y)(u_h(\delta))(dy), \tag{14}$$

and (6) shows that this is the expected discounted penalty function associated with  $h$ . Thus, the functional  $\Phi$  is our expected-discounted-penalty-function functional. It may be written in terms of component maps as

$$\Phi(a) = \Phi_1(a) * \Phi_2(a). \tag{15}$$

**2.2. The derivative of  $\Phi$**

In this subsection we give the formal framework for the derivatives of the functionals, and we state our main result, which gives the derivative of the functional  $\Phi$ . This gives rise to a first-order approximation for the expected discounted penalty function, as explained in Section 1.

It turns out that our functionals are Fréchet differentiable. Suppose that  $(B_i, \|\cdot\|_i)$ ,  $i = 1, 2$ , are Banach spaces, and that  $A$  is a nonempty, open set in  $B_1$ . The map  $\Phi: A \rightarrow B_2$  is *Fréchet differentiable* at  $a \in A$  if there is a linear, bounded map  $\Phi'_a: B_1 \rightarrow B_2$  such that, given an  $\varepsilon > 0$ , there exists an  $\eta > 0$  such that

$$\|b\|_1 < \eta \implies \|\Phi(a + b) - \Phi(a) - \Phi'_a(b)\|_2 \leq \varepsilon \|b\|_1.$$

The map  $\Phi'_a$  is called the *Fréchet derivative* of  $\Phi$  at  $a$ . Fréchet differentiability is a strong form of differentiability of maps between Banach spaces, and Fréchet derivatives obey the chain rule, which is important for our approach to the functional  $\Phi$  in Subsection 2.1 as the convolution of compositions of intermediate functionals.

We first state some of the intermediate derivatives that appear in the formula for the derivative of  $\Phi$  in our main result below. The proofs that these are indeed the derivatives of the maps in question appear later, in Section 4.

First, the derivative at  $a$  of the functional  $\Psi$  (see (9)) is given by

$$(\Psi'_a(b))(\delta) = -\frac{\lambda(\tilde{b}(\rho_a(\delta)) - \tilde{b}(0))}{\lambda\tilde{a}'(\rho_a(\delta)) + c}$$

for a finite, signed measure  $b$  and for  $\delta$  in  $[0, K]$ . Next we consider the functional  $\Phi_1: a \mapsto v_a$ , where  $v_a$  is as defined in (11). For a finite, signed measure  $b$  and for  $\delta$  in  $[0, K]$ ,  $\Phi'_{1,a}(b)$  is the element of  $\mathcal{C}$  such that  $(\Phi'_{1,a}(b))(\delta)$  is the function defined by

$$\begin{aligned} ((\Phi'_{1,a}(b))(\delta))(u) &= -\frac{\lambda}{c} \int_u^\infty e^{-\rho_a(\delta)(t-u)} (\Psi'_a(b))(\delta)(t - u) \int_{(t,\infty)} w(t, y - t) a(dy) dt \\ &\quad + \frac{\lambda}{c} \int_u^\infty e^{-\rho_a(\delta)(t-u)} \int_{(t,\infty)} w(t, y - t) b(dy) dt. \end{aligned}$$

Finally we turn to  $\Psi_1 : a \mapsto f_a$ , where  $f_a(\delta)$  is the element of  $M$  given in (12). For a finite, signed measure  $b$  and for  $\delta$  in  $[0, K]$ ,  $\Psi'_{1,a}(b)$  is the element of  $\mathcal{M}$  defined by

$$((\Psi'_{1,a}(b))(\delta))(E) = \int_E \left( -\frac{\lambda}{c} \int_{(t,\infty)} (\Psi'_a(b))(\delta)(y-t)e^{-\rho_a(\delta)(y-t)} a(dy) + \frac{\lambda}{c} \int_{(t,\infty)} e^{-\rho_a(\delta)(y-t)} b(dy) \right) dt$$

for all  $E$  in  $\mathcal{B}$ .

In the statement of the theorem below, observe that, although the functional  $\Phi$  is defined for  $a$  in  $A \cap M_{k+l+1}$ , we need  $a$  to be in  $A \cap M_{k+l+3}$  for our proof of the differentiability of the intermediate map  $\Phi_1 : a \mapsto v_a$  to hold (see Section 4). For this reason, we consider the restriction of  $\Phi$  to this smaller set at the outset.

**Theorem 1.** *The functional  $\Phi : A \cap M_{k+l+3} \rightarrow \mathcal{C}$  with  $\Phi(a) = m = v_a * u_a$  as in (14) is Fréchet differentiable at  $a$  in  $A \cap M_{k+l+3}$ , with Fréchet derivative at  $a$ ,  $\Phi'_a$ , given by*

$$\Phi'_a(b) = v_a * (\Psi'_{1,a}(b) * (e - f_a)^{*(-2)}) + \Phi'_{1,a}(b) * u_a$$

for  $b$  in  $M_{k+l+3}$ .

The derivative involves  $(e - f_a)^{*(-2)}$ , where  $b^{*(-2)}$  is  $(b^{*(-1)})^{*2}$  for any element  $b$  of  $\mathcal{M}$ . Notice that  $\Phi'_a(b)$ , while complicated, is indeed a linear function of  $b$ . Notice also that  $\Phi'_a(b)$  is in  $\mathcal{C}$ , and so  $(\Phi'_a(b))(\delta)$  is in  $C_{[0,\infty)}$  for  $\delta$  in  $[0, K]$ . As seen in Subsection 2.1, the constant  $K$ ,  $0 < K < \infty$ , can be chosen to be arbitrarily large, meaning that the (tacit) restriction in Theorem 1 that  $\delta$  be in  $[0, K]$  is not a real limitation on the applicability of the result in practice.

The proof of this theorem is given in Section 4, and proceeds by finding the derivatives of the intermediate maps in the definition of  $\Phi$  in (15) and then putting these together. The reader familiar with these methods will readily identify the derivatives of the component maps in the definitions of the quantities before the theorem and in the form of the derivative of  $\Phi$  in the statement of the theorem itself.

Now suppose that we wish to use the derivative of  $\Phi$  to obtain a first-order approximation to  $\Phi(h_1) = m^{(1)}$  for some given claim size distribution  $h_1$  for which we do not have an easily available explicit expression. As suggested in Section 1, we may approximate  $h_1$  by another distribution, say  $h_0$ , for which we do have an easily available expression for  $\Phi(h_0) = m^{(0)}$ , and this gives rise to the zeroth-order approximation,  $\Phi(h_0)$ , for  $\Phi(h_1)$ . Using the ideas outlined in Section 1, a first-order approximation is given by

$$\Phi(h_1) \approx \Phi(h_0) + \Phi'_{h_0}(h_1 - h_0). \tag{16}$$

We now translate this into an expression involving more familiar functions and measures, rather than elements of the  $\mathcal{E}$ -type spaces  $\mathcal{C}$  and  $\mathcal{M}$ . Let  $m_\delta^{(1)}(u)$  and  $m_\delta^{(0)}(u)$  be the expected discounted penalty functions associated with  $h_1$  and  $h_0$ , respectively. Then (16) becomes

$$\begin{aligned} m_\delta^{(1)}(u) &\approx m_\delta^{(0)}(u) + ((\Phi'_{h_0}(h_1 - h_0))(\delta))(u) \\ &= m_\delta^{(0)}(u) - \int_0^u (v_{h_0}(\delta))(u-t)((\Psi'_{1,h_0}(h_1 - h_0))(\delta) * (u_{h_0}(\delta)^{*2}))(dt) \\ &\quad + \int_{[0,u]} ((\Phi'_{1,h_0}(h_1 - h_0))(\delta))(u-t)(u_{h_0}(\delta))(dt). \end{aligned}$$

### 3. Example

We now illustrate our method, and assess the quality of the approximation, by the use of a particular example. We consider the case where  $w(r, s) = 1$ , meaning that

$$m_\delta(u) = E[e^{-\delta\tau(u)} \mathbf{1}(\tau(u) < \infty)].$$

We choose the distribution whose output we wish to approximate,  $h_1$ , to be a gamma distribution with parameters  $\frac{3}{2}$  and  $\frac{3}{2}$ , meaning that it has a density

$$\frac{3\sqrt{6x}e^{-3x/2}}{2\sqrt{\pi}},$$

and we note that this has a nonscale parameter which is not an integer. For this case, the Laplace transform of the time to ruin does not have a closed-form expression and we approximate it by that associated with our null input,  $h_0$ .

The simplest choice of the null input is clearly the exponential distribution. However, this has a density with a mode at 0, which is not the case for the gamma density above, and it is a single-parameter family, so our input approximation (although very easy to implement in this case) may be regarded as rather crude. The family of combinations of exponential distributions, on the other hand, has several attractive features and we choose our input claim size distribution,  $h_0$ , to be a member of that family. More specifically, a combination of exponential distributions studied by, *inter alios*, and in the context of ruin theory, Gerber *et al.* (1987) has a density with the general form

$$\sum_{i=1}^k A_i \beta_i e^{-\beta_i x}, \quad x \geq 0,$$

subject to the obvious condition  $\sum_{i=1}^k A_i = 1$ . When all the coefficients  $A_i$  are positive, one speaks of a mixture of  $k$  exponentials; in the more general case we consider here, some of the  $A_i$  can be negative and we have a combination of exponentials.

We note in particular that the Laplace transform of the time to ruin in the case where claim sizes follow a combination of exponentials was discussed in Willmot and Lin (2001, Section 9.2). For general claim size distributions  $H$ , Willmot and Lin (2001) showed that when  $w = 1$  the expected discounted penalty function associated with a distribution  $h$  can be written as the tail of a compound geometric distribution like so:

$$m_\delta(u) = \sum_{n=1}^{\infty} (1 - \phi_0) \phi_0^n \overline{F_0^{*n}}(u).$$

Here  $\phi_0 = \lambda(1 - \tilde{h}(\rho))/(\rho c)$ ,  $\lambda$  and  $c$  are respectively the arrival and premium rates in the compound Poisson model,  $\rho$  is the positive solution to (1),  $F_0$  has density

$$F'_0(y) = \frac{\rho e^{\rho y} \int_{(y, \infty)} e^{-\rho t} H(dt)}{1 - \tilde{h}(\rho)},$$

and  $\overline{F_0^{*n}} = 1 - F_0^{*n}$  denotes the tail of the  $n$ -fold convolution power of  $F_0$ . It is clear that when  $H$  is a combination of exponential distributions, so too is  $F_0$ . Moreover, an interesting result

that can be seen easily from the last equation is that, if  $h$  has a density  $\alpha_1 e^{-\beta_1 y} + \alpha_2 e^{-\beta_2 y} + \dots + \alpha_n e^{-\beta_n y}$ , then  $F'_0(y)$  has the form  $F'_0(y) = \alpha'_1 e^{-\beta_1 y} + \alpha'_2 e^{-\beta_2 y} + \dots + \alpha'_n e^{-\beta_n y}$ , meaning that the two functions have the same exponent terms. This is true for *any* choice of  $\delta$ , that is, only the coefficients of these terms depend on  $\delta$ . It then follows from standard results in ruin theory (see, for instance, Gerber *et al.* (1987)) that  $m_\delta(u)$  has the form

$$m_\delta(u) = C_1 e^{-\rho_1 u} + C_2 e^{-\rho_2 u} + \dots + C_n e^{-\rho_n u} \tag{17}$$

where the constants  $C_i$  and  $\rho_i$ ,  $i = 1, 2, \dots, n$ , can be found easily.

For the illustration of our method, it is sufficient to consider a combination of two exponential distributions, that is, we assume that the claim size distribution  $h_0$  has a density of the form

$$A_1 \beta e^{-\beta x} + (1 - A_1) \gamma e^{-\gamma x}. \tag{18}$$

We assume without loss of generality that  $\beta < \gamma$ ; then, a necessary condition for this to be a probability density is that  $0 \leq A_1 \leq \gamma / (\gamma - \beta)$ . In order to obtain the values of  $A_1$ ,  $\beta$ , and  $\gamma$ , we need three equations. One way of obtaining these is to equate the first three moments of the two input distributions,  $h_0$  and  $h_1$ . This yields the system of equations

$$\frac{A_1}{\beta} + \frac{1 - A_1}{\gamma} = 1, \quad \frac{2A_1}{\beta^2} + \frac{2(1 - A_1)}{\gamma^2} = \frac{5}{3}, \quad \frac{6A_1}{\beta^3} + \frac{6(1 - A_1)}{\gamma^3} = \frac{35}{9}, \tag{19}$$

where the three quantities on the respective right-hand sides are the first three moments of the gamma distribution. It is straightforward to check that the solution to these equations is given by

$$A_1 = \frac{5 + 4\sqrt{10}}{10}, \quad \beta = \frac{10 - \sqrt{10}}{5}, \quad \gamma = \frac{10 + \sqrt{10}}{5}.$$

Note that, to four decimal places,  $A_1 = 1.7649$ , so the density in (18) is a combination (not a mixture) of exponentials. The values of the various parameters that have been used in this particular illustration are  $\delta = 1$ ,  $\lambda = 1$ , and  $c = 2$ .

Following the arguments above, and using a simple program in MAPLE<sup>®</sup>, it is readily verified that the expected discounted penalty function associated with the null input density  $h_0$  indeed has the form in (17) with  $n = 2$  and

$$C_1 = -0.039337, \quad \rho_1 = 2.848353, \quad C_2 = 0.347519, \quad \rho_2 = 0.874381.$$

Note that  $-\rho_1$  and  $-\rho_2$  are the two negative solutions to the generalised Lundberg equation (1).

As noted above, the output,  $m_\delta^{(1)}$ , associated with  $h_1$  cannot be found analytically, so we have used simulation to obtain this. In particular, to find the simulated values of  $m_\delta^{(1)}$  we used the well-known method of exponential change of measure (see, for instance, Asmussen (2000)) and we ran 5000 simulations for each choice of the initial surplus  $u$ . On the other hand, the derivative of the map  $\Phi$ , taken at  $h_0$ , needed for the approximation in Theorem 1 can be found explicitly. In fact, this requires rather lengthy and tedious calculations, so we should mention that it is difficult to implement the approximation in Theorem 1 without the aid of a computer algebra package, such as MATHEMATICA<sup>®</sup> or MAPLE. However, using one of these packages, all the quantities involved are readily available. Figure 1 shows the zeroth-order approximation in (16) for the above choice of inputs. The solid line in the figure represents the ‘true’ output,  $m_\delta^{(1)}(u)$  (obtained by simulation), corresponding to  $h_1$ , while the dashed line depicts  $m_\delta^{(0)}(u)$ . It can be seen that the zeroth-order approximation in this case matches the shape of the unknown

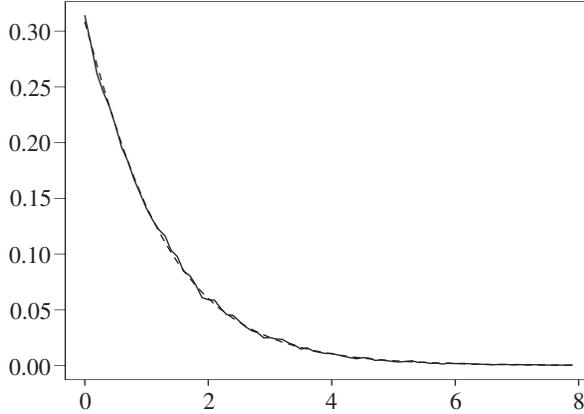


FIGURE 1:  $m_{\delta}^{(0)}(u)$  (dashed line) and  $m_{\delta}^{(1)}(u)$  (solid line) when  $h_1$  is gamma  $(\frac{3}{2}, \frac{3}{2})$  and  $h_0$  is a combination of exponentials.

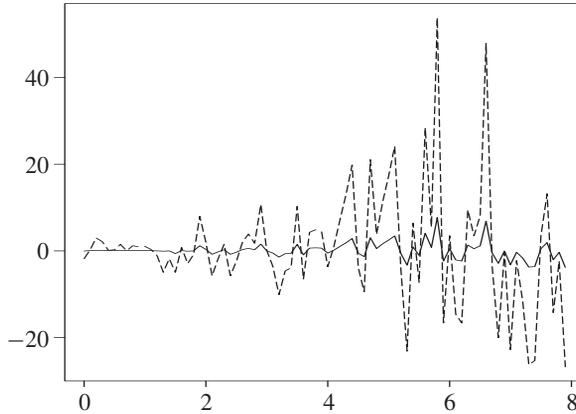


FIGURE 2: Relative percentage errors of the zeroth-order approximation (dashed line) and the first-order approximation (solid line) when  $h_1$  is gamma  $(\frac{3}{2}, \frac{3}{2})$  and  $h_0$  is a combination of exponentials.

output very well. Since  $m_{\delta}^{(1)}(u)$  is obtained by simulation for each value of  $u$  separately, the solid line in the graph is not smooth, although of course the underlying exact solution is indeed smooth.

In order to gain some insight into the advantage of using the approximation suggested by Theorem 1, in Figure 2 we show the relative percentage errors of the zeroth- and first-order approximations. More explicitly, the quantities shown in the graph are

$$\frac{m_{\delta,app}^{(i)}(u) - m_{\delta}^{(1)}(u)}{m_{\delta}^{(1)}(u)} \times 100$$

for  $i = 0, 1$ , where  $m_{\delta,app}^{(0)}(u) = m_{\delta}^{(0)}(u)$  and  $m_{\delta,app}^{(1)}(u)$  is the first-order approximation on the right-hand side of (16), using the derivative of the map  $\Phi$  from Theorem 1 and for the above choice of inputs. It can be seen from the graph that the first-order approximation improves drastically upon the zeroth-order one, as it reduces the relative percentage error by roughly six

to ten times. In particular, for the range of values shown in the graph, the maximum absolute values of the relative errors for the zeroth- and first-order approximations are 53.9% and 7.7%, respectively.

### 4. Proofs

Recall that in Subsection 2.1 we showed that  $\Phi(a) = \Phi_1(a) * \Phi_2(a)$ , where  $\Phi_1(a)$  and  $\Phi_2(a)$  both involve  $\Psi(a) = \rho_a$ . Our first step is to obtain the Fréchet differentiability of the map  $\Psi$  defined in (9).

**Proposition 1.** *The map  $\Psi: A \subseteq M_1 \rightarrow C_{[0,K]}$  with  $(\Psi(a))(\delta) = (\lambda\tilde{a} - \lambda\tilde{a}(0) + ct)^{-1}(\delta)$  is Fréchet differentiable at  $a$  in  $A$ , with Fréchet derivative at  $a$ ,  $\Psi'_a$ , given by*

$$(\Psi'_a(b))(\delta) = -\frac{\lambda(\tilde{b}(\rho_a(\delta)) - \tilde{b}(0))}{\lambda\tilde{a}'(\rho_a(\delta)) + c}$$

for  $b$  in  $A$  and  $\delta$  in  $[0, K]$ .

*Proof.* Let  $g_a(t) = \lambda\tilde{a}(t) - \lambda\tilde{a}(0) + ct$ ,  $t \geq 0$ , as in Subsection 2.1, meaning that  $g_a$  is differentiable on  $[0, \infty)$  (right differentiable at 0) and, for  $a \in A$ , we have  $\rho_a(\delta) = g_a^{-1}(\delta)$ . Hence, for  $b$  such that  $a + b$  is in  $A$ , using the mean value theorem we obtain

$$\begin{aligned} 0 &= g_{a+b}(\rho_{a+b}(\delta)) - g_a(\rho_a(\delta)) \\ &= g_{a+b}(\rho_{a+b}(\delta)) - g_a(\rho_{a+b}(\delta)) + g_a(\rho_{a+b}(\delta)) - g_a(\rho_a(\delta)) \\ &= (g_{a+b} - g_a)(\rho_{a+b}(\delta)) + g'_a(\xi)(\rho_{a+b}(\delta) - \rho_a(\delta)) \end{aligned}$$

for some  $\xi \equiv \xi(\delta)$  between  $\rho_a(\delta)$  and  $\rho_{a+b}(\delta)$ . For  $a$  in  $A$ ,  $g'_a(t) > 0$  for all  $t \geq 0$ , by (8), and we find that

$$\rho_{a+b}(\delta) - \rho_a(\delta) = -\frac{(g_{a+b} - g_a)(\rho_{a+b}(\delta))}{g'_a(\xi)} = -\frac{\lambda(\tilde{b}(\rho_{a+b}(\delta)) - \tilde{b}(0))}{g'_a(\xi)}. \tag{20}$$

This means that

$$\begin{aligned} \left| \rho_{a+b}(\delta) - \rho_a(\delta) + \frac{\lambda(\tilde{b}(\rho_a(\delta)) - \tilde{b}(0))}{g'_a(\rho_a(\delta))} \right| &= \left| -\frac{\lambda(\tilde{b}(\rho_{a+b}(\delta)) - \tilde{b}(0))}{g'_a(\xi)} + \frac{\lambda(\tilde{b}(\rho_a(\delta)) - \tilde{b}(0))}{g'_a(\rho_a(\delta))} \right| \\ &\leq \lambda(|\tilde{b}(\rho_{a+b}(\delta))| + |\tilde{b}(0)|) \left| \frac{1}{g'_a(\xi)} - \frac{1}{g'_a(\rho_a(\delta))} \right| \\ &\quad + \frac{\lambda}{|g'_a(\rho_a(\delta))|} |\tilde{b}(\rho_{a+b}(\delta)) - \tilde{b}(\rho_a(\delta))|. \end{aligned} \tag{21}$$

The first term on the right-hand side of (21) is at most

$$2\lambda\|b\|_1 \left| \frac{1}{g'_a(\xi)} - \frac{1}{g'_a(\rho_a(\delta))} \right|.$$

Let  $[0, t_0]$  be a closed, bounded interval containing  $\rho_a([0, K])$  and  $\rho_{a+b}([0, K])$ . Since  $g'_a$  is uniformly continuous on  $[0, t_0]$ , and since, by (8),  $g'_a(t) \geq c - \lambda \int_{[0,\infty)} y|a|(dy) > 0$  for all  $t \geq 0$ , we see that  $1/g'_a(t)$  is uniformly continuous on  $[0, t_0]$ . Hence, given an  $\varepsilon > 0$ , there exists an  $\eta_1 > 0$  such that

$$|y - t| < \eta_1 \implies \left| \frac{1}{g'_a(y)} - \frac{1}{g'_a(t)} \right| < \frac{\varepsilon}{4\lambda}. \tag{22}$$

We now show that we can make  $|\xi - \rho_a(\delta)|$  smaller than  $\eta_1$  by making  $\|b\|_1$  small enough. It is easy to check that

$$|g_{a+b}(t) - g_a(t)| \leq 2\lambda\|b\|_{TV} \leq 2\lambda\|b\|_1$$

for all  $t$  in  $[0, t_0]$ , meaning that  $g_{a+b}$  lies between  $g_a - 2\lambda\|b\|_1$  and  $g_a + 2\lambda\|b\|_1$ . For  $\delta$  in  $[0, K]$ , let

$$t_u = g_a^{-1}(\delta + 2\lambda\|b\|_1)$$

and

$$t_l = g_a^{-1}(\max\{0, \delta - 2\lambda\|b\|_1\}).$$

Then  $\rho_a(\delta)$  and  $\rho_{a+b}(\delta)$  are in  $(t_l, t_u)$ , and

$$|\xi - \rho_a(\delta)| \leq |\rho_{a+b}(\delta) - \rho_a(\delta)| \leq t_u - t_l = g_a^{-1}(\delta + 2\lambda\|b\|_1) - g_a^{-1}(\max\{0, \delta - 2\lambda\|b\|_1\}).$$

Since

$$|\delta + 2\lambda\|b\|_1 - \max\{0, \delta - 2\lambda\|b\|_1\}| \leq 4\lambda\|b\|_1,$$

and since the inverse function  $g_a^{-1}$  is uniformly continuous on  $[0, K]$ , we can choose  $\|b\|_1$  sufficiently small that  $|\xi - \rho_a(\delta)| < \eta_1$  for all  $\delta$  in  $[0, K]$ . Hence, by (22), for  $\|b\|_1$  sufficiently small, the first term on the right-hand side of (21) is at most  $\varepsilon\|b\|_1/2$  for all  $\delta$  in  $[0, K]$ .

For the second term on the right-hand side of (21), for  $a$  in  $A$ , we have

$$\frac{1}{|g'_a(\rho_a(\delta))|} \leq \frac{1}{c - \lambda \int_{[0,\infty)} y|a|(dy)} < \infty$$

for all  $\delta$  in  $[0, K]$ , by (8). Using the mean value theorem again, we see that, for some  $\zeta$  between  $\rho_a(\delta)$  and  $\rho_{a+b}(\delta)$ ,

$$|\tilde{b}(\rho_{a+b}(\delta)) - \tilde{b}(\rho_a(\delta))| \leq |\tilde{b}'(\zeta)| |\rho_{a+b}(\delta) - \rho_a(\delta)| \leq \|b\|_1 |\rho_{a+b}(\delta) - \rho_a(\delta)|.$$

Thus, the second term in (21) is at most

$$\frac{1}{c - \lambda \int_{[0,\infty)} y|a|(dy)} |\rho_{a+b}(\delta) - \rho_a(\delta)| \|b\|_1.$$

We have already seen above that we can choose  $\|b\|_1$  small enough that  $|\rho_{a+b}(\delta) - \rho_a(\delta)|$  is less than any particular given  $\varepsilon$  for all  $\delta$  in  $[0, K]$ , so we can choose  $\|b\|_1$  small enough that the second term in (21) is at most  $\varepsilon\|b\|_1/2$  for all  $\delta$  in  $[0, K]$ .

Putting both terms together, we can choose  $\|b\|_1$  sufficiently small that

$$\sup_{\delta \in [0, K]} \left| \rho_{a+b}(\delta) - \rho_a(\delta) + \frac{\lambda(\tilde{b}(\rho_a(\delta)) - \tilde{b}(0))}{g'_a(\rho_a(\delta))} \right| \leq \varepsilon\|b\|_1.$$

It is easily seen that the function

$$(\Psi'_a(b))(\delta) = \frac{-\lambda(\tilde{b}(\rho_a(\delta)) - \tilde{b}(0))}{\lambda\tilde{a}'(\rho_a(\delta)) + c}$$

is a continuous function of  $\delta$  on  $[0, K]$ , implying that  $\Psi'_a(b)$  is in  $C_{[0, K]}$ . Furthermore, the map  $b \mapsto \Psi'_a(b)$  is a bounded linear map from  $M_1$  to  $C_{[0, K]}$ . Hence, the map  $\Psi : a \mapsto \rho_a$  is Fréchet differentiable at  $a$  in  $A$ , with derivative as shown in the statement of the proposition.

Before proving the differentiability of the functional  $\Phi_1$ , which takes  $a$  onto  $v_a$ , we need a preliminary lemma. We write  $\|\Psi'_a\|$  for the operator norm of the bounded linear map  $\Psi'_a$ .

**Lemma 1.** *Suppose that  $\varphi: [0, \infty) \rightarrow \mathbb{R}$  is continuous and differentiable on  $[0, \infty)$ , and that  $\varphi'$  is bounded. Let  $a$  and  $a + b$  be in  $A \subseteq M_1$ .*

(i) *Then, for all  $\delta$  in  $[0, K]$ ,*

$$|\varphi(\rho_{a+b}(\delta)) - \varphi(\rho_a(\delta))| \leq C_a \sup_{v \geq 0} |\varphi'(v)| \|b\|_1,$$

where  $C_a = 2\lambda/(c - \lambda \int_{[0, \infty)} y|a|(dy))$  is a positive, finite constant.

(ii) *Suppose, in addition, that  $\varphi'$  is continuous and differentiable and that  $\varphi''$  is bounded. Then, given  $\varepsilon > 0$ , there exists an  $\eta > 0$ , independent of  $\varphi$ , such that*

$$\begin{aligned} \|b\|_1 < \eta \quad \implies \quad & |\varphi(\rho_{a+b}(\delta)) - \varphi(\rho_a(\delta)) - (\Psi'_a(b))(\delta)\varphi'(\rho_a(\delta))| \\ & \leq \left( \sup_{v \geq 0} |\varphi'(v)| + C_a \sup_{v \geq 0} |\varphi''(v)| \|\Psi'_a\| \right) \varepsilon \|b\|_1, \end{aligned}$$

for all  $\delta$  in  $[0, K]$ .

*Proof.* (i) We use the mean value theorem, together with (20) and (8), to find that, for some  $\zeta$  between  $\rho_{a+b}(\delta)$  and  $\rho_a(\delta)$ ,

$$\begin{aligned} |\varphi(\rho_{a+b}(\delta)) - \varphi(\rho_a(\delta))| &= |\varphi'(\zeta)| |\rho_{a+b}(\delta) - \rho_a(\delta)| \\ &\leq \sup_{v \geq 0} |\varphi'(v)| \frac{2\lambda}{c - \lambda \int_{[0, \infty)} y|a|(dy)} \|b\|_1. \end{aligned}$$

Since  $a$  is in  $A$ , we know that  $C_a = 2\lambda/(c - \lambda \int_{[0, \infty)} y|a|(dy))$  is finite, by (8), and part (i) is proved.

(ii) Using the mean value theorem again, for some  $\zeta$  between  $\rho_{a+b}(\delta)$  and  $\rho_a(\delta)$  we obtain

$$\begin{aligned} & |\varphi(\rho_{a+b}(\delta)) - \varphi(\rho_a(\delta)) - (\Psi'_a(b))(\delta)\varphi'(\rho_a(\delta))| \\ &= |\varphi'(\zeta)(\rho_{a+b}(\delta) - \rho_a(\delta)) - (\Psi'_a(b))(\delta)\varphi'(\rho_a(\delta))| \\ &\leq \sup_{v \geq 0} |\varphi'(v)| |\rho_{a+b}(\delta) - \rho_a(\delta) - (\Psi'_a(b))(\delta)| \\ &\quad + |\varphi'(\rho_a(\delta)) - \varphi'(\zeta)| |(\Psi'_a(b))(\delta)| \\ &\leq \sup_{v \geq 0} |\varphi'(v)| |\rho_{a+b}(\delta) - \rho_a(\delta) - (\Psi'_a(b))(\delta)| \\ &\quad + \sup_{v \geq 0} |\varphi''(v)| |\rho_{a+b}(\delta) - \rho_a(\delta)| |(\Psi'_a(b))(\delta)|, \end{aligned}$$

with yet another use of the mean value theorem for the second term, together with

$$|\zeta - \rho_a(\delta)| \leq |\rho_{a+b}(\delta) - \rho_a(\delta)|.$$

Now we apply the Fréchet differentiability of  $\Psi$  (see Proposition 1) to the first term and (20)

to the second term, to find that, for all  $\delta$  in  $[0, K]$ ,

$$\begin{aligned} & |\varphi(\rho_{a+b}(\delta)) - \varphi(\rho_a(\delta)) - (\Psi'_a(b))(\delta)\varphi'(\rho_a(\delta))| \\ & \leq \sup_{v \geq 0} |\varphi'(v)|\varepsilon\|b\|_1 + C_a\|\Psi'_a\| \|b\|_1 \sup_{v \geq 0} |\varphi''(v)| \|b\|_1 \\ & \leq \left( \sup_{v \geq 0} |\varphi'(v)| + C_a \sup_{v \geq 0} |\varphi''(v)| \|\Psi'_a\| \right) \|b\|_1 \varepsilon, \end{aligned}$$

for small enough  $\|b\|_1$ , which proves part (ii).

We prove the differentiability of  $\Phi_1$  in the next proposition. Recall that we assume that  $w$  satisfies (10) for some fixed, nonnegative  $k$  and  $l$ . As mentioned in Subsection 2.2, we need  $a$  to be in  $A \cap M_{k+l+3}$  for this result to hold.

**Proposition 2.** *The functional  $\Phi_1: A \cap M_{k+l+3} \rightarrow \mathbb{C}$  with  $\Phi_1(a) = v_a$  as defined in (11) is Fréchet differentiable at  $a$  in  $A \cap M_{k+l+3}$ , with Fréchet derivative at  $a$ ,  $\Phi'_{1,a}$ , given by*

$$\begin{aligned} ((\Phi'_{1,a}(b))(\delta))(u) &= -\frac{\lambda}{c} \int_u^\infty e^{-\rho_a(\delta)(t-u)} (\Psi'_a(b))(\delta)(t-u) \int_{(t,\infty)} w(t, y-t)a(dy) dt \\ &+ \frac{\lambda}{c} \int_u^\infty e^{-\rho_a(\delta)(t-u)} \int_{(t,\infty)} w(t, y-t)b(dy) dt. \end{aligned}$$

*Proof.* Suppose that  $a$  and  $a + b$  are in  $A \cap M_{k+l+3}$ . Then

$$\begin{aligned} & |((\Phi_1(a+b))(\delta))(u) - ((\Phi_1(a))(\delta))(u) - ((\Phi'_{1,a}(b))(\delta))(u)| \\ &= \frac{\lambda}{c} \left| \int_u^\infty e^{-\rho_{a+b}(\delta)(t-u)} \int_{(t,\infty)} w(t, y-t)(a+b)(dy) dt \right. \\ &\quad - \int_u^\infty e^{-\rho_a(\delta)(t-u)} \int_{(t,\infty)} w(t, y-t)a(dy) dt \\ &\quad + \int_u^\infty e^{-\rho_a(\delta)(t-u)} (\Psi'_a(b))(\delta)(t-u) \int_{(t,\infty)} w(t, y-t)a(dy) dt \\ &\quad \left. - \int_u^\infty e^{-\rho_a(\delta)(t-u)} \int_{(t,\infty)} w(t, y-t)b(dy) dt \right| \\ &\leq \frac{\lambda}{c} \left| \int_u^\infty (e^{-\rho_{a+b}(\delta)(t-u)} - e^{-\rho_a(\delta)(t-u)} + (\Psi'_a(b))(\delta)(t-u)e^{-\rho_a(\delta)(t-u)}) \right. \\ &\quad \left. \times \int_{(t,\infty)} w(t, y-t)a(dy) dt \right| \\ &\quad + \frac{\lambda}{c} \left| \int_u^\infty (e^{-\rho_{a+b}(\delta)(t-u)} - e^{-\rho_a(\delta)(t-u)}) \int_{(t,\infty)} w(t, y-t)b(dy) dt \right|. \tag{23} \end{aligned}$$

We apply parts (i) and (ii) of Lemma 1 to the expressions outside the inner integrals in the second and first terms on the right-hand side of (23), respectively, with  $\varphi(v) = \varphi_{t,u}(v) = e^{-v(t-u)}$  for  $t \geq u$ . Then we have  $\sup_{v \geq 0} |\varphi'(v)| = t - u$  and  $\sup_{v \geq 0} |\varphi''(v)| = (t - u)^2$ ; by Lemma 1 we have

$$\sup_{\delta \in [0, K]} |e^{-\rho_{a+b}(\delta)(t-u)} - e^{-\rho_a(\delta)(t-u)}| \leq C_a(t-u)\|b\|_1,$$

and, given an  $\varepsilon > 0$ , there exists an  $\eta > 0$ , independent of  $t$  and  $u$ , such that

$$\|b\|_1 < \eta \implies \sup_{\delta \in [0, K]} |e^{-\rho_{a+b}(\delta)(t-u)} - e^{-\rho_a(\delta)(t-u)} + (\Psi'_a(b))(\delta)(t-u) e^{-\rho_a(\delta)(t-u)}| \leq (t-u + C_a(t-u)^2 \|\Psi'_a\|) \|b\|_1 \varepsilon.$$

Thus,

$$\begin{aligned} \|b\|_1 < \eta \implies & |((\Phi_1(a+b))(\delta))(u) - ((\Phi_1(a))(\delta))(u) - ((\Phi'_{1,a}(b))(\delta))(u)| \\ & \leq \frac{\lambda}{c} \int_u^\infty (t-u + C_a(t-u)^2 \|\Psi'_a\|) \varepsilon \|b\|_1 \\ & \quad \times \int_{(t,\infty)} w(t, y-t) |a|(\mathrm{d}y) \mathrm{d}t \\ & \quad + \frac{\lambda}{c} \int_u^\infty (t-u) C_a \|b\|_1 \int_{(t,\infty)} w(t, y-t) |b|(\mathrm{d}y) \mathrm{d}t. \end{aligned} \tag{24}$$

We can check that

$$\begin{aligned} & \int_u^\infty (t-u)^2 \int_{(t,\infty)} w(t, y-t) |a|(\mathrm{d}y) \mathrm{d}t \\ & \leq C_w \int_u^\infty (t-u)^2 \int_{(t,\infty)} t^k (y-t)^l |a|(\mathrm{d}y) \mathrm{d}t \\ & \leq C_w \int_0^\infty t^{k+2} \int_{(t,\infty)} y^l |a|(\mathrm{d}y) \mathrm{d}t \\ & = \frac{C_w}{k+3} \int_{(0,\infty)} y^{k+l+3} |a|(\mathrm{d}y) \\ & \leq \frac{C_w}{k+3} \|a\|_{k+l+3}, \end{aligned} \tag{25}$$

where  $C_w$  is the constant appearing in (10). Hence, the first term on the right-hand side of (24) is finite because  $a$  is in  $M_{k+l+3}$ . This term is bounded above by  $c_1 \|b\|_1 \varepsilon$ , where  $c_1$  is a constant depending on  $a$ , for all  $u \geq 0$  and  $\delta$  in  $[0, K]$ . For the second term, similar reasoning gives an upper bound of  $c_2 \|b\|_1 \|b\|_{k+l+2}$  for all  $u \geq 0$  and all  $\delta$  in  $[0, K]$ , where  $c_2$  is a constant depending on  $a$ . Bearing in mind that  $\|b\|_1 \leq \|b\|_{k+l+2} \leq \|b\|_{k+l+3}$ , we can put all this together to see that, given an  $\varepsilon > 0$ , we can choose  $\|b\|_{k+l+3}$  to be small enough (which means that  $\|b\|_1$  is ‘small enough’) to ensure that

$$|((\Phi_1(a+b))(\delta))(u) - ((\Phi_1(a))(\delta))(u) - ((\Phi'_{1,a}(b))(\delta))(u)| \leq \varepsilon \|b\|_{k+l+3}$$

for all  $u \geq 0$  and for all  $\delta$  in  $[0, K]$ . Hence, we have shown that, for  $\|b\|_{k+l+3}$  small enough, we have

$$\|\Phi_1(a+b) - \Phi_1(a) - \Phi'_{1,a}(b)\|_c \leq \varepsilon \|b\|_{k+l+3}.$$

The map  $\Phi'_{1,a}(b)$  is a linear function of  $x$ , and we see that it is bounded using similar arguments as for (25).

We now consider the functional  $\Phi_2$ , which takes  $a$  onto  $u_a$ , where  $u_a$  is as defined in (13), and we write it as the composition of two simpler functionals. Recall that  $\Psi_1$  takes  $a \in A$  onto

$f_a \in \mathcal{M}$ , where  $f_a(\delta)$  is the finite, signed measure in (12). Let  $\Psi_2$  be the functional that takes  $f_a$  onto  $u_a = (e - f_a)^{*(-1)}$ . Just as for  $v_a$ , we need  $a$  to be in a higher order  $M$ -space,  $M_3$ , for the differentiability of  $\Psi_1$ .

**Lemma 2.** *The functional  $\Psi_1: A \cap M_3 \rightarrow \mathcal{M}$  with  $\Psi_1(a) = f_a$  is Fréchet differentiable at  $a$  in  $A \cap M_3$ . The Fréchet derivative of  $\Psi_1$  at  $a$  is  $\Psi'_{1,a}$ , where  $(\Psi'_{1,a}(b))(\delta)$  is the finite, signed measure given by*

$$\begin{aligned}
 & ((\Psi'_{1,a}(b))(\delta))(E) \\
 &= \int_E \left( -\frac{\lambda}{c} \int_{(t,\infty)} (\Psi'_a(b))(\delta)(y-t)e^{-\rho_a(\delta)(y-t)} a(dy) + \frac{\lambda}{c} \int_{(t,\infty)} e^{-\rho_a(\delta)(y-t)} b(dy) \right) dt
 \end{aligned}$$

for  $\delta$  in  $[0, K]$ ,  $b$  in  $M_3$ , and  $E$  in  $\mathcal{B}$ .

*Proof.* For  $a$  and  $a + b$  in  $A \cap M_3$ ,  $\Psi_1(a + b) = f_{a+b}$  and  $\Psi_1(a) = f_a$  are in  $\mathcal{M}$ . Thus,

$$\begin{aligned}
 & \|(\Psi_1(a + b))(\delta) - (\Psi_1(a))(\delta) - (\Psi'_{1,a}(b))(\delta)\|_{TV} \\
 & \leq \int_0^\infty \left| \frac{\lambda}{c} \int_{(t,\infty)} e^{-\rho_{a+b}(\delta)(y-t)} (a + b)(dy) - \frac{\lambda}{c} \int_{(t,\infty)} e^{-\rho_a(\delta)(y-t)} a(dy) \right. \\
 & \quad \left. + \frac{\lambda}{c} \int_{(t,\infty)} (\Psi'_a(b))(\delta)(y-t)e^{-\rho_a(\delta)(y-t)} a(dy) - \frac{\lambda}{c} \int_{(t,\infty)} e^{-\rho_a(\delta)(y-t)} b(dy) \right| dt \\
 & \leq \frac{\lambda}{c} \int_0^\infty \int_{(t,\infty)} |e^{-\rho_{a+b}(\delta)(y-t)} - e^{-\rho_a(\delta)(y-t)} \\
 & \quad + (\Psi'_a(b))(\delta)(t-y)e^{-\rho_a(\delta)(y-t)}| |a|(dy) dt \\
 & \quad + \frac{\lambda}{c} \int_0^\infty \int_{(t,\infty)} |e^{-\rho_{a+b}(\delta)(y-t)} - e^{-\rho_a(\delta)(y-t)}| |b|(dy) dt. \tag{26}
 \end{aligned}$$

The proof now proceeds along similar lines to that of Proposition 2, and uses Lemma 1 applied to  $\varphi(v) = \varphi_{y,t}(v) = e^{-v(y-t)}$ ,  $y \geq t$ , with part (i) of Lemma 1 applied to the integrand in the second term on the right-hand side of (26) and part (ii) applied to the integrand in the first term. We also use the fact that, for any  $a$  in  $M_{j+1}$ ,

$$\int_0^\infty \int_{(t,\infty)} (y-t)^j |a|(dy) dt \leq \|a\|_{j+1} < \infty. \tag{27}$$

Hence, given an  $\varepsilon_1 > 0$ , we find that for  $\|b\|_1$  small enough, the right-hand side of (26) is at most

$$\frac{\lambda}{c} C_a \|b\|_1 \|b\|_2 + \frac{\lambda}{c} \|b\|_1 \varepsilon_1 (\|a\|_2 + C_a \|\Psi'_a\| \|a\|_3),$$

where  $C_a = 2\lambda/(c - \lambda \int_{[0,\infty)} y|a|(dy))$  and we recall that  $\|\Psi'_a\|$  is the norm of  $\Psi'_a$  as a linear operator. Putting this together, we find that, given an  $\varepsilon > 0$ , we can choose  $\|b\|_3$  to be small enough that

$$\|(\Psi_1(a + b))(\delta) - (\Psi_1(a))(\delta) - (\Psi'_{1,a}(b))(\delta)\|_{TV} \leq \varepsilon \|b\|_3$$

for all  $\delta$  in  $[0, K]$ , whence

$$\|\Psi_1(a + b) - \Psi_1(a) - \Psi'_{1,a}(b)\|_{\mathcal{M}} \leq \varepsilon \|b\|_3.$$

Finally, it is immediate that  $\Psi'_{1,a}$  is linear. Using (27) we see that it is also bounded.

For  $\Psi_2: f_a \mapsto u_a = (e - f_a)^{*(-1)}$ , we use general results for the inverse map in Banach algebras to obtain the derivative.

**Lemma 3.** *The functional  $\Psi_2: V = \{f \in \mathcal{M} : \|f\|_{\mathcal{M}} < 1\} \rightarrow \mathcal{M}$  with  $\Psi_2(f) = (e - f)^{*(-1)}$  is Fréchet differentiable at  $f$  in  $V$ , with Fréchet derivative,  $\Psi'_{2,f}$ , given by*

$$\Psi'_{2,f}(g) = g * (e - f)^{*(-2)}$$

for  $g$  in  $\mathcal{M}$ .

*Proof.* Rudin (1986, Theorem 18.3) showed that  $e - f$  is invertible in  $\mathcal{M}$  if  $f$  is in  $V$ . Corollary 2 to Rudin (1986, Theorem 18.4) showed that the derivative of the map  $d \mapsto d^{*(-1)}$  at  $d$ , where  $d$  is invertible in  $\mathcal{M}$ , is  $b \mapsto -b * d^{*(-2)}$ . Hence, using the chain rule, we see that  $\Psi_2$  is Fréchet differentiable at  $f$  in  $V$ , with

$$\Psi'_{2,f}(b) = b * (e - f)^{*(-2)}.$$

We combine Lemmas 2 and 3 using the chain rule to find the derivative of  $\Phi_2$ .

**Proposition 3.** *The functional  $\Phi_2: A \cap M_3 \rightarrow \mathcal{M}$  with  $\Phi_2(a) = u_a$  as defined in (13) is Fréchet differentiable at  $a$  in  $A \cap M_3$ , with Fréchet derivative at  $a$ ,  $\Phi'_{2,a}$ , given by*

$$(\Phi'_{2,a}(b))(\delta) = (\Psi'_{1,a}(b))(\delta) * (e_0 - f_a(\delta))^{*(-2)}$$

for  $b$  in  $M_3$ .

*Proof of Theorem 1.* Recall that

$$\Phi(a) = m = v_a * u_a = \Phi_1(a) * \Phi_2(a),$$

where  $\Phi_1(a)$  is in  $\mathcal{C}$ ,  $\Phi_2(a)$  is in  $\mathcal{M}$ , and  $\Phi(a)$  is in  $\mathcal{C}$ . For  $a$  in  $A \cap M_{k+l+3}$ , both  $\Phi_1$  and  $\Phi_2$  are Fréchet differentiable at  $a$ , and Propositions 2 and 3 respectively give the forms of their derivatives. For  $f$  in  $\mathcal{C}$  and  $b$  in  $\mathcal{M}$ , we have

$$\|f * b\|_{\mathcal{D}} \leq \|f\|_{\mathcal{C}} \|b\|_{\mathcal{M}}.$$

Using this, we see that the map  $\Phi_3$  taking the pair  $(f, b) \in \mathcal{C} \times \mathcal{M}$  onto  $f * b$  in  $\mathcal{D}$  is Fréchet differentiable, with Fréchet derivative at  $(f, b)$  given by

$$\Phi'_{3,(f,b)}(g, c) = f * c + g * b$$

for  $g$  in  $\mathcal{C}$  and  $c$  in  $\mathcal{M}$ . Hence, we have

$$\begin{aligned} \Phi'_a(b) &= \Phi_1(a) * \Phi'_{2,a}(b) + \Phi'_{1,a}(b) * \Phi_2(a) \\ &= v_a * (\Psi'_{1,a}(b) * (e - f_a)^{*(-2)}) + \Phi'_{1,a}(b) * u_a. \end{aligned}$$

This gives the expression in the statement of Theorem 1.

## 5. Conclusions and miscellaneous remarks

1. In working out the details needed in applying the functional method to give new approximations for the discounted penalty function, the first point to note is that, except in the case where the distribution of claim sizes is a combination of exponential distributions, explicit expressions for the expected discounted penalty function are typically complicated, even when  $w(r, s)$  is identically equal to 1, where we might have expected to obtain simple formulae. This strengthens the case for approximations of the sort given here and for their first-order corrections.

2. Following the remark above, it is of interest to consider the ‘optimal’ way one can approximate an arbitrary input distribution  $h_1$  by a combination of exponentials. Here, in the example of Section 3, we used the straightforward method of moment matching to approximate a gamma distribution by a combination of two exponentials, and, as illustrated in Figure 1, this yielded a very workable approximation for the expected discounted penalty function. There are clearly more sophisticated ways to implement this input fit: an obvious alternative is, instead of matching the first three moments, to match the mean, the variance, and the asymptotics. The resulting approximation, however, differs rather insignificantly to the one used here and, in addition, the system of equations replacing (19) in this case may not be analytically solvable. Among the various other procedures that can be used to approximate an arbitrary distribution on  $[0, \infty)$  by a combination of exponentials, we mention in particular the recent paper by Dufresne (2006).

3. Using the mean value theorem for maps in Banach spaces (see, for example, Cartan (1971, Proposition 3.3.1)), we see that the differentiability of  $\Phi$  gives rise to an upper bound on the error in  $\Phi(h_0)$ , the zeroth-order approximation for  $\Phi(h_1)$ ; indeed

$$\|\Phi(h_1) - \Phi(h_0)\|_c \leq \sup_{0 \leq t \leq 1} \|\Phi'_{(1-t)h_0+th_1}\| \|h_1 - h_0\|_{k+l+3},$$

where  $\|\Phi'_a\|$  denotes the operator norm of  $\Phi'_a$ . After a lot of algebra, and taking further upper bounds on  $\|\Phi'_a\|$ , we find that

$$\begin{aligned} \|\Phi'_{(1-t)h_0+th_1}\| \leq & \frac{\lambda C_w}{l_{h_0}^2} \left( \max\{\|h_0\|_{k+l+1}, \|h_1\|_{k+l+1}\} \left(1 + \frac{2\lambda\mu}{l_{h_0}}\right) \right. \\ & \left. + \frac{2\lambda \max\{\|h_0\|_{k+l+2}, \|h_1\|_{k+l+2}\}}{k+2} + \frac{l_{h_0}}{k+1} \right), \end{aligned}$$

where  $l_a = c - \int_{[0, \infty)} t|a|(dt)$ . This can be substituted into the mean value theorem bound. In principle, it should be possible to find a better bound by taking less crude upper bounds on  $\|\Phi'_a\|$ .

4. We have considered the compound Poisson risk model. The literature has also included work on the expected discounted penalty function for the Sparre Andersen model, where the claims arrive in a renewal process (see, for example, Dickson and Hipp (2001), Cheng and Tang (2003), Gerber and Shiu (2005), and Li and Garrido (2004)). The functional here is more complicated than for the compound Poisson model, and is the subject of further research.

5. In this paper the derivative of the functional of interest turns out to be fairly complicated, and perhaps this goes hand in hand with the difficulties mentioned in point 1 above. To avoid further complications, we have assumed throughout that  $c$  is fixed. Writing  $c = (1 + \theta)\lambda\mu$ ,

where  $\theta = (c - \lambda\mu)/(\lambda\mu)$  is the relative safety loading, we see that we can think of this either as meaning that we regard  $\mu$  and  $\theta$  as fixed, that is, we choose our approximating claim size distribution,  $h_0$ , so that it has the same mean as that of  $h_1$ , the claim size distribution of interest, and we fix the relative safety loading, or alternatively as meaning that we allow  $\mu$  to change, with the premium income fixed, so that a change in  $\mu$  results in an implied change in  $\theta$ . Clearly there are other possible important scenarios that are not covered by these two points of view; for example that in which  $\theta$  is fixed and in addition we allow  $\mu$  to change. The functional method can be extended to cover this case also, by, for example, regarding the ratio  $\lambda/c$  not as a constant but as a functional of the claim size distribution. However, this would add further complications to the derivatives without adding any insight, so we have not attempted this.

6. Another extension is to include  $w$  as input to the functional, meaning that we consider the functional

$$(a, w) \mapsto v_{a,w} * u_a,$$

where we write  $v_{a,w}$  for  $v_a$  to denote the dependence on  $w$ . The advantage of this is that it allows for approximation of  $m$  by the expected discounted penalty function belonging to a claim size distribution and/or  $w$  for which  $m$  is more easily obtained.

7. The cases where explicit expressions are available, even if they are complicated, typically belong to the class of phase-type distributions, and this class of distributions is one where explicit expressions are often available for many functionals of interest in applied probability (for an application to  $Y_+(u)$  in the Sparre Andersen model, see Drekić *et al.* (2004)). This leads to the practical approach of approximating an input distribution for which there is no easy explicit expression for  $m_\delta(u)$  by a ‘near-by’ phase-type distribution. The functional approach then lends itself to producing new corrected approximations that include the first-order correction term to the phase-type approximation.

### Acknowledgement

We thank an anonymous referee for helpful comments on the original version of the manuscript.

### References

- ASMUSSEN, S. (2000). *Ruin Probabilities*. World Scientific, Singapore.
- CARTAN, H. (1971). *Differential Calculus*. Hermann, Paris.
- CHENG, Y. AND TANG, Q. (2003). Moments of the surplus before ruin and the deficit at ruin in the Erlang(2) risk process. *N. Amer. Actuarial J.* **7**, 1–12.
- DICKSON, D. C. M AND HIPPI, C. (2001). On the time to ruin for Erlang(2) risk processes. *Insurance Math. Econom.* **29**, 333–344.
- DREKIĆ, S., DICKSON, D. C. M., STANFORD, D. A. AND WILLMOT, G. E. (2004). On the distribution of the deficit at ruin when claims are phase-type. *Scand. Actuarial J.* **2004**, 105–120.
- DUFRESNE, D. (2006). Fitting combinations of exponentials to probability distributions. *Appl. Stoch. Models Business Industry* **23**, 23–48.
- DUNFORD, N. AND SCHWARTZ, J. T. (1958). *Linear Operators. Part I. General Theory*. John Wiley, New York.
- GERBER, H. U. AND SHIU, E. S. W. (1998). On the time value of ruin. *N. Amer. Actuarial J.* **2**, 48–78.
- GERBER, H. U. AND SHIU, E. S. W. (2005). The time value of ruin in a Sparre Andersen model. *N. Amer. Actuarial J.* **2**, 49–84.
- GERBER, H. U., GOOVAERTS, M. J. AND KAAS, R. (1987). On the probability and severity of ruin. *ASTIN Bull.* **17**, 151–163.
- GRÜBEL, R. (1989). Stochastic models as functionals: some remarks on the renewal case. *J. Appl. Prob.* **26**, 296–303.
- GRÜBEL, R. AND PITTS, S. M. (1992). A functional approach to the stationary waiting time and idle period distributions of the GI/G/1 queue. *Ann. Prob.* **20**, 1754–1778.

- LI, S. AND GARRIDO, J. (2004). On ruin for the Erlang( $n$ ) risk process. *Insurance Math. Econom.* **34**, 391–408.
- LIN, X. S. AND WILLMOT, G. E. (1999). Analysis of a defective renewal equation arising in ruin theory. *Insurance Math. Econom.* **25**, 63–84.
- LIN, X. S. AND WILLMOT, G. E. (2000). The moments of the time of ruin, the surplus before ruin, and the deficit at ruin. *Insurance Math. Econom.* **27**, 19–44.
- PITTS, S. M. (1998). A functional approach to approximations for the observed open times in single ion channels. *J. Math. Anal. Appl.* **219**, 260–278.
- PITTS, S. M. (1999). Approximations for semi-Markov single ion channel models. In *Semi-Markov Models and Applications*, eds J. Janssen and N. Limnios, Kluwer, Dordrecht, pp. 103–115.
- PITTS, S. M. (2004). A functional approach to approximations for the individual risk model. *ASTIN Bull.* **34**, 379–397.
- POLITIS, K. AND PITTS, S. M. (1998). Approximations for solutions of renewal-type equations. *Stoch. Process. Appl.* **78**, 195–216.
- RUDIN, W. (1986). *Real and Complex Analysis*. McGraw-Hill, New York.
- WILLMOT, G. E. AND LIN, X. S. (2001). *Lundberg Approximations for Compound Distributions with Insurance Applications*. Springer, New York.