

# On queues with service and interarrival times depending on waiting times

O.J. Boxma · M. Vlasiou

Received: 26 July 2006 / Revised: 26 November 2006 / Published online: 20 March 2007  
© Springer Science+Business Media, LLC 2007

**Abstract** We consider an extension of the standard G/G/1 queue, described by the equation  $W \stackrel{D}{=} \max\{0, B - A + YW\}$ , where  $\mathbb{P}[Y = 1] = p$  and  $\mathbb{P}[Y = -1] = 1 - p$ . For  $p = 1$  this model reduces to the classical Lindley equation for the waiting time in the G/G/1 queue, whereas for  $p = 0$  it describes the waiting time of the server in an alternating service model. For all other values of  $p$ , this model describes a FCFS queue in which the service times and interarrival times depend linearly and randomly on the waiting times. We derive the distribution of  $W$  when  $A$  is generally distributed and  $B$  follows a phase-type distribution, and when  $A$  is exponentially distributed and  $B$  deterministic.

**Keywords** Lindley recursion · G/G/1 queue · Alternating service model

**Mathematics Subject Classification (2000)** 60K25 · 90B22

---

This research has been carried out when M. Vlasiou was affiliated with EURANDOM, The Netherlands.

---

O.J. Boxma  
EURANDOM and Department of Mathematics & Computer Science, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands  
e-mail: boxma@win.tue.nl

M. Vlasiou (✉)  
H. Milton Stewart School of Industrial & Systems Engineering, Georgia Institute of Technology, 765 Ferst Drive, Atlanta, GA 30332-0205, USA  
e-mail: vlasiou@gatech.edu

## 1 Introduction

One of the most fundamental relations in queuing and random walk theory is Lindley's recursion [11]:

$$W_{n+1} = \max\{0, B_n - A_n + W_n\}. \quad (1.1)$$

In queuing theory, it represents a relation between the waiting times of the  $n$ th and  $(n + 1)$ st customer in a single server queue,  $A_n$  indicating the interarrival time between the  $n$ th and  $(n + 1)$ st customer and  $B_n$  denoting the service time of the  $n$ th customer. In the applied probability literature there has been a considerable amount of interest in generalisations of Lindley's recursion, namely the class of Markov chains described by the recursion  $W_{n+1} = g(W_n, X_n)$ . For earlier work on such stochastic recursions see, for example, Brandt et al. [5] and Borovkov and Foss [3]. Many structural properties of this recursion have been derived. For example, Asmussen and Sigman [2] develop a duality theory, relating the steady-state distribution to a ruin probability associated with a risk process. More references in this domain can be found in Asmussen and Schock Petersen [1] and Seal [16]. An important assumption which is often made in these studies is that the function  $g(w, x)$  is non-decreasing in its main argument  $w$ . For example, in [2] this assumption is crucial for their duality theory to hold.

In this paper we consider a generalisation of Lindley's recursion, for which the monotonicity assumption does not hold. In particular, we study the Lindley-type recursion

$$W_{n+1} = \max\{0, B_n - A_n + Y_n W_n\}, \quad (1.2)$$

where for every  $n$ , the random variable  $Y_n$  is equal to plus or minus one according to the probabilities  $\mathbb{P}[Y_n = 1] = p$  and  $\mathbb{P}[Y_n = -1] = 1 - p$ ,  $0 \leq p \leq 1$ . The sequences  $\{A_n\}$

and  $\{B_n\}$  are assumed to be independent sequences of i.i.d. non-negative random variables. Our main goal is to derive the steady-state distribution of  $\{W_n, n = 1, 2, \dots\}$ , when it exists.

Equation (1.2) reduces to the classical Lindley recursion [11] when  $\mathbb{P}[Y_n = 1] = 1$  for every  $n$ . Furthermore, if  $\mathbb{P}[Y_n = -1] = 1$ , then (1.2) describes the waiting time of the server in an alternating service model with two service points. For a description of the model for this special case and related results see Park et al. [13] and Vlastou et al. [20–23].

Studying a recursion that contains both Lindley's classical recursion and the recursion in [13, 20–23] as special cases seems of interest in its own right. Additional motivation for studying the recursion is supplied by the fact that, for  $0 < p < 1$ , the resulting model can be interpreted as a special case of a queueing model in which service and interarrival times depend on waiting times. We shall now discuss the latter model.

Consider an extension of the standard G/G/1 queue in which the service times and the interarrival times depend linearly and randomly on the waiting times. Namely, the model is specified by a stationary and ergodic sequence of four-tuples of non-negative random variables  $\{(A_n, B_n, \widehat{A}_n, \widehat{B}_n)\}$ ,  $n \geq 0$ . The sequence  $\{W_n\}$  is defined recursively by

$$W_{n+1} = \max\{0, \overline{B}_n - \overline{A}_n + W_n\},$$

where

$$\overline{A}_n = A_n + \widehat{A}_n W_n,$$

$$\overline{B}_n = B_n + \widehat{B}_n W_n.$$

We interpret  $W_n$  as the waiting time and  $\overline{B}_n$  as the service time of customer  $n$ . Furthermore, we take  $\overline{A}_n$  to be the interarrival time between customers  $n$  and  $n + 1$ . We call  $B_n$  the *nominal service time* of customer  $n$  and  $A_n$  the *nominal interarrival time* between customers  $n$  and  $n + 1$ , because these would be the actual times if the additional shift were omitted, that is, if  $\mathbb{P}[\widehat{A}_n = \widehat{B}_n = 0] = 1$ .

Evidently, the waiting times satisfy the generalised Lindley recursion (1.2), where we have written  $Y_n = 1 + \widehat{B}_n - \widehat{A}_n$ . This model—for generally distributed random variables  $Y_n$ —has been introduced in Whitt [24], where the focus is on conditions for the process to converge to a proper steady-state limit, and on approximations for this limit. There are very few exact results known for queueing models in which interarrival and/or service times depend on waiting times; we refer to Whitt [24] for some references.

Whitt [24] builds upon previous results by Vervaat [19] and Brandt [4] for the unrestricted recursion  $W_{n+1} = Y_n W_n + X_n$ , where  $X_n = B_n - A_n$ . There has been considerable previous work on this unrestricted recursion, due to its close connection to the problem of the ruin of an insurer

who is exposed to a stochastic economic environment. Such an environment has two kinds of risk, which were called by Norberg [12] insurance risk and financial risk. Indicatively, we mention the work by Tang and Tsitsiashvili [17], and by Kalashnikov and Norberg [10]. In the more general framework,  $W_n$  may represent an inventory in time period  $n$  (e.g. cash),  $Y_n$  may represent a multiplicative, possibly random, decay or growth factor between times  $n$  and  $n + 1$  (e.g. interest rate) and  $B_n - A_n$  may represent a quantity that is added or subtracted between times  $n$  and  $n + 1$  (e.g. deposit minus withdrawal). Obviously, the positive-part operator is appropriate for many applications [24].

This paper presents an exact analysis of the steady-state distribution of  $\{W_n, n = 1, 2, \dots\}$  as given by (1.2) with  $\mathbb{P}[Y_n = 1] = p$  and  $\mathbb{P}[Y_n = -1] = 1 - p$ . For  $0 < p < 1$ , this amounts to analysing the above-described G/G/1 extension where  $\widehat{A}_n = \widehat{B}_n$  with probability  $p$ , and  $\widehat{A}_n = 2 + \widehat{B}_n$  with probability  $1 - p$ . This problem, and state-dependent queueing processes in general, is connected to LaPalice queueing models, introduced by Jacquet [9], where customers are scheduled in such a way that the period between two consecutively scheduled customers is greater than or equal to the service time of the first customer.

This paper is organised in the following way. In Sect. 2 we comment on the stability of the process  $\{W_n\}$ , as it is defined by recursion (1.2). In the remainder of the paper it is assumed that the steady-state distribution of  $\{W_n\}$  exists. Section 3 is devoted to the determination of the distribution of  $W$  when  $A$  is generally distributed and  $B$  has a phase-type distribution. In Sect. 4 we determine the distribution of  $W$  when  $A$  is exponentially distributed and  $B$  is deterministic. At the end of each section we compare the results that we derive to the already known results for Lindley's recursion (i.e. for  $p = 1$ ) and to the equivalent results for the Lindley-type recursion arising for  $p = 0$ .

At the end of this introduction we mention a few notational conventions. For a random variable  $X$  we denote its distribution by  $F_X$  and its density by  $f_X$ . Furthermore, we shall denote by  $f^{(i)}$  the  $i$ th derivative of the function  $f$ . The Laplace-Stieltjes transforms (LST) of  $A$  and  $W$  are respectively denoted by  $\alpha$  and  $\omega$ . To keep expressions simple, we also use the function  $\phi$  defined as  $\phi(s) = \omega(s)\alpha(s)$ .

## 2 Stability

The following result on the convergence of the process  $\{W_n\}$  to a proper limit  $W$  is shown in Whitt [24]. It is included here only for completeness.

From recursion (1.2), it is obvious that if we replace  $Y_n$  by  $\max\{0, Y_n\}$  and  $B_n - A_n$  by  $\max\{0, B_n - A_n\}$ , then the resulting waiting times will be at least as large as the ones given by (1.2). Moreover, when we make this change, the positive-part operator is not necessary anymore.

**Lemma 1** (Whitt [24, Lemma 1]) *If  $W_n$  satisfies (1.2), then with probability 1,  $W_n \leq Z_n$  for all  $n$ , where*

$$Z_{n+1} = \max\{0, Y_n\}Z_n + \max\{0, B_n - A_n\}, \quad n \geq 0, \quad (2.1)$$

and  $Z_0 = W_0 \geq 0$ .

So if  $W_n$  satisfies (1.2),  $Z_n$  satisfies (2.1), and  $Z_n$  converges to the proper limit  $Z$ , then  $\{W_n\}$  is tight and  $\mathbb{P}[W > x] \leq \mathbb{P}[Z > x]$  for all  $x$ , where  $W$  is the limit in distribution of any convergent subsequence of  $\{W_n\}$ . This observation, combined with Theorem 1 of Brandt [4], which implies that  $Z_n$  satisfying (2.1) converges to a proper limit if  $\mathbb{P}[\max\{0, Y_n\} = 0] = \mathbb{P}[Y_n \leq 0] > 0$ , leads to the following theorem.

**Theorem 1** (Whitt [24, Theorem 1]) *The series  $\{W_n\}$  is tight for all  $\rho = \mathbb{E}[B_0]/\mathbb{E}[A_0]$  and  $W_0$ . If, in addition,  $0 \leq p < 1$  and  $\{Y_n, B_n - A_n\}$  is a sequence of independent vectors with*

$$\mathbb{P}[Y_0 \leq 0, B_0 - A_0 \leq 0] > 0,$$

*then the events  $\{W_n = 0\}$  are regeneration points with finite mean time and  $\{W_n\}$  converges in distribution to a proper limit  $W$  as  $n \rightarrow \infty$  for all  $\rho$  and  $W_0$ .*

Naturally, for  $p = 1$ , i.e. for the classical Lindley recursion, we need the additional condition that  $\rho < 1$ .

Therefore, assume that the sequences  $\widehat{B}_n - \widehat{A}_n$  and  $B_n - A_n$  are independent stationary sequences, that are also independent of one another, and that for all  $n$ ,  $A_n$  and  $B_n$  are non-negative. Then the conditions of Theorem 1 hold, so there exists a proper limit  $W$ , and for the system in steady-state we write

$$W \stackrel{\mathcal{D}}{=} \max\{0, B - A + YW\}, \quad (2.2)$$

where “ $\stackrel{\mathcal{D}}{=}$ ” denotes equality in distribution, where  $A, B$  are generic random variables distributed like  $A_n, B_n$ , and where  $\mathbb{P}[Y = 1] = p$  and  $\mathbb{P}[Y = -1] = 1 - p$ .

*Remark 1* For  $x \geq 0$  (2.2) yields that

$$\begin{aligned} F_W(x) &= \mathbb{P}[W \leq x] \\ &= p\mathbb{P}[X + W \leq x] + (1 - p)\mathbb{P}[X - W \leq x], \end{aligned}$$

where  $X = B - A$  (note that  $\mathbb{P}[X < 0] > 0$ ). Assuming that the distribution  $F_X$  of the random variable  $X$  is continuous, the last term is equal to  $1 - \mathbb{P}[X - W \geq x]$ , which gives us that

$$\begin{aligned} F_W(x) &= p \int_{-\infty}^x F_W(x - y) dF_X(y) \\ &\quad + (1 - p) \left( 1 - \int_x^{\infty} F_W(y - x) dF_X(y) \right). \end{aligned}$$

This means that the limiting distribution of  $W$ , provided that  $F_X$  is continuous, satisfies the functional equation

$$\begin{aligned} F(x) &= p \int_{-\infty}^x F(x - y) dF_X(y) \\ &\quad + (1 - p) \left( 1 - \int_x^{\infty} F(y - x) dF_X(y) \right). \end{aligned} \quad (2.3)$$

Therefore, there exists at least one function that is a solution to (2.3). It can be shown that in fact there exists a unique measurable bounded function  $F : [0, \infty) \rightarrow \mathbb{R}$  that satisfies this functional equation.

To show this, consider the space  $\mathcal{L}^\infty([0, \infty))$ , i.e. the space of measurable and bounded functions on the real line with the norm

$$\|F\| = \sup_{t \geq 0} |F(t)|.$$

In this space define the mapping

$$\begin{aligned} (\mathcal{T}F)(x) &= p \int_{-\infty}^x F(x - y) dF_X(y) \\ &\quad + (1 - p) \left( 1 - \int_x^{\infty} F(y - x) dF_X(y) \right). \end{aligned}$$

Note that  $\mathcal{T}F : \mathcal{L}^\infty([0, \infty)) \rightarrow \mathcal{L}^\infty([0, \infty))$ , i.e.,  $\mathcal{T}F$  is measurable and bounded. For two arbitrary functions  $F_1$  and  $F_2$  in this space we have

$$\begin{aligned} \|(\mathcal{T}F_1) - (\mathcal{T}F_2)\| &= \sup_{x \geq 0} |(\mathcal{T}F_1)(x) - (\mathcal{T}F_2)(x)| \\ &= \sup_{x \geq 0} \left| p \int_{-\infty}^x [F_1(x - y) - F_2(x - y)] dF_X(y) \right. \\ &\quad \left. + (1 - p) \int_x^{\infty} [F_2(y - x) - F_1(y - x)] dF_X(y) \right| \\ &\leq \sup_{x \geq 0} \left( p \int_{-\infty}^x |F_1(x - y) - F_2(x - y)| dF_X(y) \right. \\ &\quad \left. + (1 - p) \int_x^{\infty} |F_2(y - x) - F_1(y - x)| dF_X(y) \right) \\ &\leq \sup_{x \geq 0} \left( p \int_{-\infty}^x \sup_{t \geq 0} |F_1(t) - F_2(t)| dF_X(y) \right. \\ &\quad \left. + (1 - p) \int_x^{\infty} \sup_{t \geq 0} |F_2(t) - F_1(t)| dF_X(y) \right) \\ &= \|F_1 - F_2\| \sup_{x \geq 0} [p F_X(x) + (1 - p)(1 - F_X(x))]. \end{aligned}$$

Note that the supremum appearing above is less than or equal to  $\max\{p, 1 - p\}$  for  $p \in (0, 1)$  and equal to  $1 - F_X(0)$  for  $p = 0$ . Therefore, since for  $p \neq 1$  it holds

that  $\sup_{x \geq 0} [p F_X(x) + (1 - p)(1 - F_X(x))] < 1$ , for these values of the parameter  $p$  we have a contraction mapping. Furthermore, we know that  $\mathcal{L}^\infty([0, \infty))$  is a Banach space, therefore by the Fixed Point Theorem we have that (2.3) has a unique solution; for a related discussion, see also [20].

In the next two sections we determine the distribution of  $W$  for two cases in which  $\{A_n\}$  and  $\{B_n\}$  are independent and i.i.d. sequences of non-negative random variables.

### 3 The GI/PH case

In this section we assume that  $A$  is generally distributed, while  $B$  follows a particular phase-type distribution. Specifically, we assume that with probability  $\kappa_n$  the nominal service time  $B$  follows an Erlang distribution with parameter  $\mu$  and  $n$  phases, i.e.,

$$F_B(x) = \sum_{n=1}^N \kappa_n \left( 1 - e^{-\mu x} \sum_{j=0}^{n-1} \frac{(\mu x)^j}{j!} \right) = \sum_{n=1}^N \kappa_n \sum_{j=n}^{\infty} e^{-\mu x} \frac{(\mu x)^j}{j!}, \quad x \geq 0, \tag{3.1}$$

with LST  $\sum_{n=1}^N \kappa_n (\mu / (\mu + s))^n$ . These distributions, i.e. mixtures of Erlang distributions, are special cases of Coxian or phase-type distributions. It is sufficient to consider only this class, since it may be used to approximate any given continuous distribution on  $[0, \infty)$  arbitrarily close; see Schassberger [15]. Following the proof in [22], one can show that for such an approximation of  $F_B$ , the error in the resulting waiting time approximation can be bounded.

We are interested in the distribution of  $W$ . In order to derive the distribution of  $W$ , we shall first derive the LST of  $F_W$ . We follow a method based on Wiener-Hopf decomposition. A straightforward calculation yields for values of  $s$  such that  $\text{Re}(s) = 0$ :

$$\begin{aligned} \omega(s) &= \mathbb{E}[e^{-sW}] \\ &= p \mathbb{E}[e^{-s \max\{0, B-A+W\}}] \\ &\quad + (1-p) \mathbb{E}[e^{-s \max\{0, B-A-W\}}] \\ &= p \mathbb{P}[W + B \leq A] + p \mathbb{E}[e^{-s(B-A+W)}] \\ &\quad - p \mathbb{E}[e^{-s(B-A+W)}; W + B \leq A] \\ &\quad + (1-p) \mathbb{P}[B \leq W + A] \\ &\quad + (1-p) \mathbb{E}[e^{-s(B-A-W)}; B \geq W + A]; \end{aligned} \tag{3.2}$$

here  $A$ ,  $B$  and  $W$  are independent random variables. The Lindley-type equation  $W \stackrel{D}{=} \max\{0, B - A - W\}$  for  $A$  generally distributed and  $B$  phase-type has already been

analysed in Vlasiou and Adan [21], and the LST of the corresponding  $W$  is given there. From (3.8) of [21] we can readily copy an expression for the last two terms appearing in (3.2), so  $\omega$  can now be written as

$$\begin{aligned} \omega(s) &= p \mathbb{P}[W + B \leq A] + p \alpha(-s) \omega(s) \sum_{n=1}^N \kappa_n \left( \frac{\mu}{\mu + s} \right)^n \\ &\quad - p \mathbb{E}[e^{-s(B-A+W)}; W + B \leq A] \\ &\quad + (1-p) \left[ 1 - \sum_{n=1}^N \sum_{i=0}^{n-1} \kappa_n \frac{(-\mu)^i}{i!} \phi^{(i)}(\mu) \right. \\ &\quad \left. \times \left( 1 - \left( \frac{\mu}{\mu + s} \right)^{n-i} \right) \right]. \end{aligned}$$

So for  $\text{Re}(s) = 0$  we have that

$$\begin{aligned} \omega(s) &\left[ 1 - p \alpha(-s) \sum_{n=1}^N \kappa_n \left( \frac{\mu}{\mu + s} \right)^n \right] \\ &= p \mathbb{P}[W + B \leq A] - p \mathbb{E}[e^{-s(B-A+W)}; W + B \leq A] \\ &\quad + (1-p) \left[ 1 - \sum_{n=1}^N \sum_{i=0}^{n-1} \kappa_n \frac{(-\mu)^i}{i!} \phi^{(i)}(\mu) \right. \\ &\quad \left. \times \left( 1 - \left( \frac{\mu}{\mu + s} \right)^{n-i} \right) \right]. \end{aligned} \tag{3.3}$$

Cohen [6, pp. 322–323] shows by applying Rouché’s theorem that the function

$$\begin{aligned} &1 - p \alpha(-s) \sum_{n=1}^N \kappa_n \left( \frac{\mu}{\mu + s} \right)^n \\ &\equiv \frac{1}{(\mu + s)^N} \left[ (\mu + s)^N - p \alpha(-s) \sum_{n=1}^N \kappa_n \mu^n (\mu + s)^{N-n} \right] \end{aligned}$$

has exactly  $N$  zeros  $\xi_i(p)$  in the left-half plane if  $0 < p < 1$  (it is assumed that  $\alpha(\mu) \neq 0$ , which is not an essential restriction) or if  $p = 1$  and  $\mathbb{E}[B] < \mathbb{E}[A]$ . Naturally, this statement is not valid if  $p = 0$ ; therefore, this case needs to be excluded from this point on. So we rewrite (3.3) as follows

$$\begin{aligned} \omega(s) &\prod_{i=1}^N (s - \xi_i(p)) \\ &= \frac{\prod_{i=1}^N (s - \xi_i(p))}{(\mu + s)^N - p \alpha(-s) \sum_{n=1}^N \kappa_n \mu^n (\mu + s)^{N-n}} \\ &\quad \times \left[ p (\mu + s)^N \mathbb{P}[W + B \leq A] \right. \\ &\quad \left. - p (\mu + s)^N \mathbb{E}[e^{-s(B-A+W)}; W + B \leq A] \right] \end{aligned}$$

$$\begin{aligned}
 &+ (1 - p) \left[ (\mu + s)^N - \sum_{n=1}^N \sum_{i=0}^{n-1} \kappa_n \frac{(-\mu)^i}{i!} \phi^{(i)}(\mu) \right. \\
 &\left. \times ((\mu + s)^N - \mu^{n-i} (\mu + s)^{N-n+i}) \right]. \tag{3.4}
 \end{aligned}$$

The left-hand side of (3.4) is analytic for  $\text{Re}(s) > 0$  and continuous for  $\text{Re}(s) \geq 0$ , and the right-hand side of (3.4) is analytic for  $\text{Re}(s) < 0$  and continuous for  $\text{Re}(s) \leq 0$ . So from Liouville’s theorem [18] we have that both sides of (3.4) are the same  $N$ th degree polynomial, say,  $\sum_{i=0}^N q_i s^i$ . Hence,

$$\omega(s) = \frac{\sum_{i=0}^N q_i s^i}{\prod_{i=1}^N (s - \xi_i(p))}. \tag{3.5}$$

In the expression above, the constants  $q_i$  are not determined so far, while the roots  $\xi_i(p)$  are known. In order to obtain the transform, observe that  $\omega$  is a fraction of two polynomials of degree  $N$ . So, ignoring the special case of multiple zeros  $\xi_i(p)$ , partial fraction decomposition yields that (3.5) can be rewritten as

$$\omega(s) = c_0 + \sum_{i=1}^N \frac{c_i}{(s - \xi_i(p))}, \tag{3.6}$$

which implies that the waiting time distribution has a mass at the origin that is given by

$$\mathbb{P}[W = 0] = \lim_{s \rightarrow \infty} \mathbb{E}[e^{-sW}] = c_0$$

and has a density that is given by

$$f_W(x) = \sum_{i=1}^N c_i e^{\xi_i(p)x}.$$

All that remains is to determine the  $N + 1$  constants  $c_i$ . To do so, we work as follows.

We shall substitute (3.6) in the left-hand side of (3.4), and express the terms  $\mathbb{P}[W + B \leq A]$  and  $\mathbb{E}[e^{-s(B-A+W)}; W + B \leq A]$  that appear at the right-hand side of (3.4) in terms of the constants  $c_i$ . Note that the terms  $\phi^{(i)}(\mu)$  that appear at the right-hand side of (3.4) can also be expressed in terms of the constants  $c_i$ . Thus we obtain a new equation that we shall differentiate a total of  $N$  times. We shall evaluate each of these derivatives for  $s = 0$  and thus we obtain a linear system of  $N$  equations for the constants  $c_i, i = 0, \dots, N$ . The last equation that is necessary to uniquely determine the constants  $c_i$  is the normalisation equation

$$c_0 + \int_0^\infty f_W(x) dx = 1. \tag{3.7}$$

To begin with, note that

$$\begin{aligned}
 \mathbb{P}[W + B \leq A] &= \mathbb{P}[W = 0] \mathbb{P}[B \leq A] \\
 &+ \int_0^\infty \mathbb{P}[B \leq A - x] \sum_{i=1}^N c_i e^{\xi_i(p)x} dx, \tag{3.8}
 \end{aligned}$$

with

$$\begin{aligned}
 \mathbb{P}[B \leq A] &= \int_0^\infty \sum_{n=1}^N \kappa_n \left( e^{-\mu x} \sum_{j=n}^\infty \frac{(\mu x)^j}{j!} \right) dF_A(x) \\
 &= \sum_{n=1}^N \sum_{i=n}^\infty \kappa_n \frac{(-\mu)^i}{i!} \alpha^{(i)}(\mu), \tag{3.9}
 \end{aligned}$$

and

$$\begin{aligned}
 &\int_0^\infty \mathbb{P}[B \leq A - x] \sum_{i=1}^N c_i e^{\xi_i(p)x} dx \\
 &= \int_0^\infty \int_0^\infty \mathbb{P}[B \leq y - x] \sum_{i=1}^N c_i e^{\xi_i(p)x} dx dF_A(y) \\
 &= \int_0^\infty \int_0^y e^{-\mu(y-x)} \sum_{n=1}^N \\
 &\quad \times \sum_{j=n}^\infty \kappa_n \frac{(\mu(y-x))^j}{j!} \sum_{i=1}^N c_i e^{\xi_i(p)x} dx dF_A(y) \\
 &= \sum_{n=1}^N \sum_{j=n}^\infty \sum_{i=1}^N \sum_{k=j+1}^\infty \kappa_n c_i \frac{\mu^j (\mu + \xi_i(p))^{k-j-1}}{k! (-1)^k} \alpha^{(k)}(\mu). \tag{3.10}
 \end{aligned}$$

Likewise, we have that

$$\begin{aligned}
 &\mathbb{E}[e^{-s(B-A+W)}; W + B \leq A] \\
 &= \mathbb{P}[W = 0] \mathbb{E}[e^{-s(B-A)}; B \leq A] \\
 &\quad + \int_0^\infty \mathbb{E}[e^{-s(B-A+x)}; x + B \leq A] \sum_{i=1}^N c_i e^{\xi_i(p)x} dx, \tag{3.11}
 \end{aligned}$$

with

$$\begin{aligned}
 &\mathbb{E}[e^{-s(B-A)}; B \leq A] \\
 &= \int_0^\infty \int_0^x e^{-s(y-x)} \sum_{n=1}^N \kappa_n \mu e^{-\mu y} \frac{(\mu y)^{n-1}}{(n-1)!} dy dF_A(x) \\
 &= \int_0^\infty e^{xs} \sum_{n=1}^N \kappa_n \left( \frac{\mu}{\mu + s} \right)^n
 \end{aligned}$$

$$\begin{aligned} & \times \sum_{i=n}^{\infty} e^{-x(\mu+s)} \frac{x^i(\mu+s)^i}{i!} dF_A(x) \\ & = \sum_{n=1}^N \sum_{i=n}^{\infty} \kappa_n \mu^n \frac{(-1)^i}{i!} (\mu+s)^{i-n} \alpha^{(i)}(\mu), \end{aligned} \tag{3.12}$$

and

$$\begin{aligned} & \int_0^{\infty} \mathbb{E}[e^{-s(B-A+x)}; x+B \leq A] \sum_{i=1}^N c_i e^{\xi_i(p)x} dx \\ & = \int_0^{\infty} \int_0^y \int_0^{y-x} e^{-s(z-y+x)} \sum_{n=1}^N \kappa_n \mu e^{-\mu z} \frac{(\mu z)^{n-1}}{(n-1)!} \\ & \quad \times \sum_{i=1}^N c_i e^{\xi_i(p)x} dz dx dF_A(y) \\ & = \int_0^{\infty} \int_0^y \sum_{n=1}^N \kappa_n \left(\frac{\mu}{\mu+s}\right)^n e^{-s(x-y)} \\ & \quad \times \sum_{j=n}^{\infty} e^{-(\mu+s)(y-x)} \frac{(\mu+s)^j (y-x)^j}{j!} \\ & \quad \times \sum_{i=1}^N c_i e^{\xi_i(p)x} dx dF_A(y) \\ & = \sum_{n=1}^N \sum_{j=n}^{\infty} \sum_{i=1}^N \sum_{k=j+1}^{\infty} \kappa_n c_i \left(\frac{\mu}{\mu+s}\right)^n \\ & \quad \times \frac{(\mu+s)^j (\mu+\xi_i(p))^{k-j-1}}{k!(-1)^k} \alpha^{(k)}(\mu). \end{aligned} \tag{3.13}$$

So, using (3.9) and (3.10), substitute (3.8) in the right-hand side of (3.4), and similarly for (3.11). Furthermore, as mentioned before, substitute (3.6) into the left-hand side of (3.4) to obtain an expression, where both sides can be reduced to an  $N$ th degree polynomial in  $s$ . By evaluating this polynomial and all its derivatives for  $s = 0$  we obtain  $N$  equations binding the constants  $c_i$ . These equations, and the normalisation equation (3.7), form a linear system for the constants  $c_i$ ,  $i = 0, \dots, N$ , that uniquely determines them (see also Remark 2 below). For example, the first equation, evaluated at  $s = 0$ , yields that

$$c_0 - \sum_{i=1}^N \frac{c_i}{\xi_i(p)} = \frac{1-p}{1-p\alpha(0)} = 1,$$

since  $\alpha(0) = 1$ . We summarise the above in the following theorem.

**Theorem 2** Consider the recursion given by (1.2), and assume that  $0 < p < 1$ . Let (3.1) be the distribution of the random variable  $B$ . Then the limiting distribution of the waiting

time has mass  $c_0$  at the origin and a density on  $[0, \infty)$  that is given by

$$f_W(x) = \sum_{i=1}^N c_i e^{\xi_i(p)x}.$$

In the above equation, the constants  $\xi_i(p)$ , with  $\text{Re}(\xi_i(p)) < 0$ , are the  $N$  roots of

$$(\mu+s)^N - p\alpha(-s) \sum_{n=1}^N \kappa_n \mu^n (\mu+s)^{N-n} = 0,$$

and the  $N + 1$  constants  $c_i$  are the unique solution to the linear system described above.

*Remark 2* Although the roots  $\xi_i(p)$  and coefficients  $c_i$  may be complex-valued, the density and the mass  $c_0$  at zero will be positive. This follows from the fact that there is a unique equilibrium distribution and thus a unique solution to the linear system for the coefficients  $c_i$ . Of course, it is also clear that each root  $\xi_i(p)$  and coefficient  $c_i$  have a companion conjugate root and conjugate coefficient, which implies that the imaginary parts appearing in the density cancel.

*Remark 3* In case that  $\xi_i(p)$  has multiplicity greater than one for one or more values of  $i$ , the analysis proceeds in essentially the same way. For example, if  $\xi_1(p) = \xi_2(p)$ , then the partial fraction decomposition of  $\omega$  becomes

$$\omega(s) = c_0 + \frac{c_1}{(s - \xi_1(p))^2} + \sum_{i=2}^N \frac{c_i}{s - \xi_i(p)},$$

the inverse of which is given by

$$f_W(x) = c_1 x e^{\xi_1(p)x} + \sum_{i=2}^N c_i e^{\xi_i(p)x}.$$

*Remark 4* For the nominal service time  $B$  we have considered only mixtures of Erlang distributions, mainly because this class approximates well any continuous distribution on  $[0, \infty)$  and because we can illustrate the techniques we use without complicating the analysis. However, we can extend this class by considering distributions with a rational Laplace transform. The analysis in [21] can be extended to such distributions, and the analysis in Cohen [6, Sect. II.5.10] is already given for such distributions, so the results given there can be implemented directly.

*Remark 5* The analysis we have presented so far can be directly extended to the case where  $Y$  takes any finite number of negative values. In other words, let the distribution of  $Y$  be given by  $\mathbb{P}[Y = 1] = p$ , and for  $i = 1, \dots, n$ ,

$\mathbb{P}[Y = -u_i] = p_i$ , where  $u_i > 0$  and  $\sum_i p_i = 1 - p$ . Then, for example, (3.3) becomes

$$\begin{aligned} \omega(s) & \left[ 1 - p \alpha(-s) \sum_{n=1}^N \kappa_n \left( \frac{\mu}{\mu + s} \right)^n \right] \\ & = p \mathbb{P}[W + B \leq A] - p \mathbb{E}[e^{-s(B-A+W)}; W + B \leq A] \\ & \quad + \sum_{i=1}^n p_i \mathbb{P}[B \leq u_i W + A] \\ & \quad + \sum_{i=1}^n p_i \left[ \sum_{n=1}^N \kappa_n \left( \frac{\mu}{\mu + s} \right)^n \alpha(-s) \omega(-u_i s) \right. \\ & \quad \left. - \mathbb{E}[e^{-s(B-A-u_i W)}; B \leq u_i W + A] \right]. \end{aligned}$$

Following the same steps as below (3.3), we can conclude that the waiting time density is again given by a mixture of exponentials of the form

$$f_W(x) = \sum_{i=1}^N \hat{c}_i e^{\hat{\xi}_i(p)x},$$

where the new constants  $\hat{c}_i$  (and the mass of the distribution at zero, given by  $\hat{c}_0$ ) are to be determined as the unique solution to a linear system of equations. The only additional remark necessary when forming this linear system is to observe that both the probability  $\mathbb{P}[B \leq u_i W + A]$  and the expectation  $\mathbb{E}[e^{-s(B-A-u_i W)}; B \leq u_i W + A]$  can be expressed linearly in terms of the constants  $\hat{c}_i$ .

*The case  $p = 0$*

We have seen that the case where  $Y_n = -1$  for all  $n$ , or in other words the case  $p = 0$ , had to be excluded from the analysis. Equation (3.4) is still valid if we take the constants  $\xi_i(0)$  to be defined as in Theorem 2. However, one cannot apply Liouville’s theorem to the resulting equation. The transform can be inverted directly. As it is shown in [21], the terms  $\phi^{(i)}(\mu)$  that remain to be determined follow by differentiating (3.4)  $N - 1$  times and evaluating  $\omega^j(s)$  at  $s = \mu$  for  $i = 0, \dots, N - 1$ . The density in this case is a mixture of Erlang distributions with the same scale parameter  $\mu$  for all exponential phases. As we can see, for  $p = 0$  the resulting density is intrinsically different from the one described in Theorem 2.

*The case  $p = 1$*

If  $p = 1$  and  $\mathbb{E}[B] < \mathbb{E}[A]$ , then we are analysing the steady-state waiting time distribution of a G/PH/1 queue. Equa-

tion (3.4) now reduces to

$$\begin{aligned} \omega(s) & \prod_{i=1}^N (s - \xi_i(1)) \\ & = \frac{\prod_{i=1}^N (s - \xi_i(1))}{(\mu + s)^N - \alpha(-s) \sum_{n=1}^N \kappa_n \mu^n (\mu + s)^{N-n}} \\ & \quad \times [(\mu + s)^N \mathbb{P}[W + B \leq A] \\ & \quad - (\mu + s)^N \mathbb{E}[e^{-s(B-A+W)}; W + B \leq A]]. \end{aligned} \tag{3.14}$$

Earlier we have already observed that the right-hand side of (3.14) is equal to an  $N$ th degree polynomial  $\sum_{i=0}^N q_i s^i$ . Inspection of the right-hand side of (3.14) reveals that it has an  $N$ -fold zero in  $s = -\mu$ . Indeed, all zeros of the numerator of the quotient in the right-hand side cancel against zeros of the denominator, and the term

$$\mathbb{P}[W + B \leq A] - \mathbb{E}[e^{-s(B-A+W)}; W + B \leq A]$$

is finite for  $s = -\mu$ . Hence,

$$\sum_{i=0}^N q_i s^i = q_N (\mu + s)^N. \tag{3.15}$$

Combining (3.14) and (3.15), we conclude that

$$\omega(s) \prod_{i=1}^N (s - \xi_i(1)) = q_N (\mu + s)^N,$$

and since  $\omega(0) = 1$ , the last equation gives us that

$$q_N = \frac{\prod_{i=1}^N (-\xi_i(1))}{\mu^N}.$$

Thus, we have that

$$\omega(s) = \left( \frac{\mu + s}{\mu} \right)^N \prod_{i=1}^N \frac{\xi_i(1)}{\xi_i(1) - s},$$

which is in agreement with II.5.190 in [6, p. 324].

**4 The M/D case**

We have examined so far the case where the nominal interarrival time  $A$  is generally distributed and the nominal service time  $B$  follows a phase-type distribution. In other words, we have studied the case which is in a sense analogous to the ordinary G/PH/1 queue. We now would like to study the reversed situation; namely, the case analogous to the M/G/1 queue.

The M/G/1 queue has been studied in much detail. However, the analogous alternating service model—i.e., take

$\mathbb{P}(Y = -1) = 1$  in (1.2), so  $p = 0$ —seems to be more complicated to analyse. As shown in [20], if  $p = 0$ , the density of  $W$  satisfies a generalised Wiener-Hopf equation, for which no solution is known in general. The presently available results for the distribution of  $W$  with  $p = 0$  are developed in [20], where  $B$  is assumed to belong to a class strictly bigger than the class of functions with rational Laplace transforms, but not completely general. Moreover, the method developed in [20] breaks down when applied to (2.2) with  $Y$  not identically equal to  $-1$ .

We shall refrain from trying to develop an alternative approach for the M/G case with a more general distribution for  $B$  than the one treated in Sect. 3. Instead, we give a detailed analysis of the M/D case:  $A$  is exponentially distributed and  $B$  is deterministic. This case is neither contained in the G/PH case of the previous section nor has it been treated (for the special choice of  $p = 0$ ) in [20]. Its analysis is of interest for various reasons. To start with, the model generalises the classical M/D/1 queue; additionally, the analysis illustrates the difficulties that arise when studying (2.2) in case  $A$  is exponentially distributed and  $B$  is generally distributed; finally, the different effects of Lindley’s classical recursion and of the Lindley-type recursion discussed in [20] are clearly exposed. As we shall see in the following, the analysis can be practically split into two parts, where each part follows the analysis of the corresponding model with  $Y \equiv 1$ , or  $Y \equiv -1$ .

### 4.1 Deterministic nominal service times

As before, consider (2.2), and assume that  $Y = 1$  with probability  $p$  and  $Y = -1$  with probability  $1 - p$ . Let  $A$  be exponentially distributed with rate  $\lambda$  and  $B$  be equal to  $b$ , where  $b > 0$ . Furthermore, we shall denote by  $\pi_0$  the mass of the distribution of  $W$  at zero; that is,  $\pi_0 = \mathbb{P}[W = 0]$ .

For this setting, we have from (2.2) that for  $x \geq 0$ ,

$$\begin{aligned}
 F_W(x) &= \mathbb{P}[\max\{0, b - A + YW\} \leq x] \\
 &= \mathbb{P}[b - A + YW \leq x] \\
 &= p \mathbb{P}[b - A + W \leq x] + (1 - p) \mathbb{P}[b - A - W \leq x] \\
 &= p \pi_0 \mathbb{P}[b - A \leq x] \\
 &\quad + p \int_0^\infty \mathbb{P}[b - A \leq x - y] f_W(y) dy \\
 &\quad + (1 - p) \pi_0 \mathbb{P}[b - A \leq x] \\
 &\quad + (1 - p) \int_0^\infty \mathbb{P}[b - A \leq x + y] f_W(y) dy \\
 &= \pi_0 \mathbb{P}[A \geq b - x] + p \int_0^\infty \mathbb{P}[A \geq b - x + y] f_W(y) dy \\
 &\quad + (1 - p) \int_0^\infty \mathbb{P}[A \geq b - x - y] f_W(y) dy. \tag{4.1}
 \end{aligned}$$

So, for  $0 \leq x < b$  the above equation reduces to

$$\begin{aligned}
 F_W(x) &= \pi_0 e^{-\lambda(b-x)} + p \int_0^\infty e^{-\lambda(b-x+y)} f_W(y) dy \\
 &\quad + (1 - p) \int_0^{b-x} e^{-\lambda(b-x-y)} f_W(y) dy \\
 &\quad + (1 - p) \int_{b-x}^\infty f_W(y) dy, \tag{4.2}
 \end{aligned}$$

and for  $x \geq b$ , (4.1) reduces to

$$\begin{aligned}
 F_W(x) &= \pi_0 + p \int_0^{x-b} f_W(y) dy \\
 &\quad + p \int_{x-b}^\infty e^{-\lambda(b-x+y)} f_W(y) dy + (1 - p)(1 - \pi_0), \tag{4.3}
 \end{aligned}$$

where we have utilised the normalisation equation

$$\pi_0 + \int_0^\infty f_W(y) dy = 1. \tag{4.4}$$

In the following, we shall derive the distribution on the interval  $[0, b)$  and on the interval  $[b, \infty)$  separately. At this point though, one should note that from (2.2) it is apparent that for  $A$  exponentially distributed and  $B = b$ , the distribution of  $W$  is continuous on  $(0, \infty)$ . Also, one can verify that (4.2) for  $x = b$  reduces to (4.3) for  $x = b$ . The fact that  $F_W$  is continuous on  $(0, \infty)$  will be used extensively in the sequel. Notice also that from (4.2) and (4.3) we can immediately see that we can differentiate  $F_W(x)$  for  $x \in (0, b)$  and  $x \in (b, \infty)$ ; see, for example, Titchmarsh [18, p. 59].

#### The distribution on $[0, b)$

In all subsequent equations it is assumed that  $x \in (0, b)$ . In order to derive the distribution of  $W$  on  $[0, b]$ , we differentiate (4.2) once to obtain

$$\begin{aligned}
 f_W(x) &= \lambda \pi_0 e^{-\lambda(b-x)} + \lambda p \int_0^\infty e^{-\lambda(b-x+y)} f_W(y) dy \\
 &\quad + \lambda(1 - p) \int_0^{b-x} e^{-\lambda(b-x-y)} f_W(y) dy \\
 &\quad - (1 - p) e^{-\lambda(b-x)} e^{\lambda(b-x)} f_W(b - x) \\
 &\quad + (1 - p) f_W(b - x).
 \end{aligned}$$

We rewrite this equation after noticing that the terms in the last two lines cancel, while the sum of the integrals in the first two lines can be rewritten by using (4.2). Thus, we have that



$$\begin{aligned}
 f_W(x) &= \lambda\pi_0 e^{-\lambda(b-x)} + \lambda \left( F_W(x) - \pi_0 e^{-\lambda(b-x)} \right. \\
 &\quad \left. - (1-p) \int_{b-x}^{\infty} f_W(y) dy \right) \\
 &= \lambda F_W(x) - \lambda(1-p) \int_{b-x}^{\infty} f_W(y) dy. \tag{4.5}
 \end{aligned}$$

In order to obtain a linear differential equation, differentiate (4.5) once more, which leads to

$$f'_W(x) = \lambda f_W(x) - \lambda(1-p)f_W(b-x). \tag{4.6}$$

Equation (4.6) is a homogeneous linear differential equation, not of a standard form because of the argument  $b-x$  that appears at the right-hand side. To solve it, we substitute  $x$  for  $b-x$  in (4.6) to obtain

$$f'_W(b-x) = \lambda f_W(b-x) - \lambda(1-p)f_W(x). \tag{4.7}$$

Then, we differentiate (4.6) once more to obtain

$$f''_W(x) = \lambda f'_W(x) + \lambda(1-p)f'_W(b-x),$$

and we eliminate the term  $f'_W(b-x)$  by using (4.7). Thus, we conclude that

$$f''_W(x) = \lambda^2 p(2-p)f_W(x). \tag{4.8}$$

For  $p \neq 0$ , the solution to this differential equation is given by

$$f_W(x) = d_1 e^{r_1 x} + d_2 e^{r_2 x}, \tag{4.9}$$

where  $r_1$  and  $r_2$  are given by

$$r_{1,2} = \pm \lambda \sqrt{p(2-p)}, \tag{4.10}$$

and the constants  $d_1$  and  $d_2$  will be determined by the initial conditions. Namely, the solution needs to satisfy (4.6) and the condition  $F_W(0) = \pi_0$ . Thus, for the first equation, substitute the general solution we have derived into (4.6). For the second equation, first rewrite (4.5) as follows:

$$f_W(x) = \lambda F_W(x) - \lambda(1-p) \left( 1 - \pi_0 - \int_0^{b-x} f_W(y) dy \right),$$

then substitute  $f_W(x)$  from (4.9), and finally evaluate the resulting equation for  $x = 0$ . This system uniquely determines  $d_1$  and  $d_2$ . Specifically, we have that

$$\begin{aligned}
 d_1 &= \frac{\lambda^2(1-p)(1-p(1-\pi_0)-2\pi_0)r_1}{(e^{br_1}-1)\lambda^2(2-p)(1-p)+e^{br_1}r_1(r_1-\lambda(2-p))}, \\
 d_2 &= \frac{e^{br_1}\lambda(1-p(1-\pi_0)-2\pi_0)r_1(\lambda-r_1)}{(e^{br_1}-1)\lambda^2(2-p)(1-p)+e^{br_1}r_1(r_1-\lambda(2-p))},
 \end{aligned}$$

where in the process we have assumed that  $p \neq 1$ . Up to this point we have that the waiting-time distribution on  $[0, b]$  is given by

$$F_W(x) = \frac{d_1}{r_1}(e^{r_1 x} - 1) + \frac{d_2}{r_2}(e^{r_2 x} - 1) + \pi_0, \tag{4.11}$$

where  $d_1$  and  $d_2$  are known up to the probability  $\pi_0$ . The cases for  $p = 0$  and  $p = 1$  follow directly from (4.8) and will be handled separately in the sequel.

*The distribution on  $[b, \infty)$*

As before, we obtain a differential equation by differentiating (4.3) once, and substituting the resulting integrals by using (4.3) once more. Thus, we obtain the equation

$$\begin{aligned}
 f_W(x) &= \lambda \left( F_W(x) - \pi_0 - (1-p)(1-\pi_0) \right. \\
 &\quad \left. - p \int_0^{x-b} f_W(y) dy \right),
 \end{aligned}$$

which can be reduced to

$$f_W(x) = \lambda(F_W(x) - 1 + p - p F_W(x-b)). \tag{4.12}$$

Equation (4.12) is a delay-differential equation that can be solved recursively. Observe that for  $x \in (b, 2b)$ , the term  $F_W(x-b)$  has been derived in the previous step, so for  $x \in (b, 2b)$ , (4.12) reduces to an ordinary linear differential equation from which we can easily derive the distribution of  $W$  in the interval  $(b, 2b)$ .

For simplicity, denote by  $F_i(x)$  the distribution of  $W$  when  $x \in [ib, (i+1)b]$ , and analogously denote by  $f_i(x)$  the density of  $W$ , when  $x \in (ib, (i+1)b)$ . Then (4.12) states that

$$f_i(x) = \lambda(F_i(x) - 1 + p - p F_{i-1}(x-b)),$$

which leads to an expression for  $F_i$  that is given in terms of an indefinite integral that is a function of  $x$ , that is,

$$\begin{aligned}
 F_i(x) &= e^{\lambda x} \left[ \int \lambda(-1 + p - p F_{i-1}(x-b))e^{-\lambda x} dx + \gamma_i \right], \\
 i &\geq 1.
 \end{aligned} \tag{4.13}$$

The constants  $\gamma_i$  can be derived by exploiting the fact that the waiting-time distribution is continuous. In particular, every  $\gamma_i$  is determined by the equation

$$F_i(ib) = F_{i-1}(ib). \tag{4.14}$$

Solving (4.13) recursively, we obtain that

$$\begin{aligned}
 F_i(x) &= 1 - p^i(1 - \pi_0) - p^i \left( \frac{d_1}{r_1} + \frac{d_2}{r_2} \right) \\
 &+ \sum_{j=1}^2 \left( \frac{\lambda p}{\lambda - r_j} \right)^i \frac{d_j}{r_j} e^{r_j(x-ib)} \\
 &+ x \sum_{j=0}^{i-1} (-\lambda p)^j \gamma_{i-j} \frac{(x-jb)^{j-1}}{j!} e^{\lambda(x-jb)}. \quad (4.15)
 \end{aligned}$$

Observe that for  $i = 0$ , if we define the empty sum at the right-hand side to be equal to zero, then the above expression reduces to (4.11). Notice that, since we have made use of the distribution on  $[0, b)$  as it is given by (4.11), (4.15) is not valid for  $p = 0$  or  $p = 1$ . From (4.14) we now have that for every  $i \geq 1$ ,

$$\begin{aligned}
 \gamma_i &= e^{-\lambda ib} (1 - p) p^{i-1} \left( \pi_0 - 1 - \frac{d_1 - d_2}{r_1} \right) \\
 &- \sum_{j=1}^2 \frac{e^{-\lambda ib} d_j}{r_j} \left( \frac{\lambda p}{\lambda - r_j} \right)^i \left( 1 - \frac{e^{br_j}(\lambda - r_j)}{\lambda p} \right) \\
 &+ i \sum_{j=1}^{i-1} \frac{e^{-\lambda jb} (i-j)^{j-1} (-\lambda p b)^j (\gamma_{i-1-j} - \gamma_{i-j})}{j!} \\
 &+ \gamma_{i-1}, \quad (4.16)
 \end{aligned}$$

where we have assumed that  $\gamma_0 = 0$ , and that for  $i = 1$ , the second sum is equal to zero. Dividing the expressions for  $d_i$  by  $r_i$  leads to the simplification

$$1 - p^i(1 - \pi_0) - p^i \left( \frac{d_1}{r_1} + \frac{d_2}{r_2} \right) = 1 - \frac{p^i}{2 - p}.$$

Recall that  $d_1$  and  $d_2$ , and thus also all constants  $\gamma_i$ , are known in terms of  $\pi_0$ . The probability  $\pi_0$  that still remains to be determined will be given by the normalisation (4.4). Notice though, that since the waiting-time distribution is determined recursively for every interval  $[ib, (i + 1)b]$ , (4.4) yields an infinite sum. The sum is well defined, since a unique density exists. The above findings are summarised in the following theorem.

**Theorem 3** Consider the recursion given by (1.2), and assume that  $0 < p < 1$ . Let  $A$  be exponentially distributed with rate  $\lambda$  and  $B$  be equal to  $b$ , where  $b > 0$ . Then for  $x \in [ib, (i + 1)b]$ ,  $i = 0, 1, \dots$ , the limiting distribution of the waiting time is given by

$$\begin{aligned}
 F_W(x) &= 1 - \frac{p^i}{2 - p} + \sum_{j=1}^2 \left( \frac{\lambda p}{\lambda - r_j} \right)^i \frac{d_j}{r_j} e^{r_j(x-ib)} \\
 &+ x \sum_{j=0}^{i-1} (-\lambda p)^j \gamma_{i-j} \frac{(x-jb)^{j-1}}{j!} e^{\lambda(x-jb)},
 \end{aligned}$$

where the constants  $\gamma_i$  are given by (4.16) and the probability  $\pi_0$  is given by the normalisation equation (4.4).

One might expect though that (4.4) may not be suitable for numerically determining  $\pi_0$ . However, if the probability  $p$  is not too close to one, or in other words, if the system does not almost behave like an M/D/1 queue, then one can numerically approximate  $\pi_0$  from the normalisation equation. As an example, in Fig. 1 we display a typical plot of the waiting-time distribution. We have chosen  $b = 1$ ,  $\lambda = 2$ , and  $p = 1/3$ .

For  $p$  close to one, we can see from the expressions for  $d_1$  and  $d_2$  that both the numerators and the denominators of these two constants approach zero. Furthermore, the denominators  $\lambda - r_j$ ,  $j = 1, 2$  that appear in the waiting-time distribution also approach zero, which makes Theorem 3 unsuitable for numerical computations for values of  $p$  close to one. Moreover, we also see that very large values of the parameter  $\lambda$  may also lead to numerical problems, since  $\lambda$  is involved in the exponent of almost all exponential terms that appear in the waiting-time distribution.

As one can observe from Fig. 1, and show from Theorem 3,  $F_W$  is not differentiable for  $x = b$ . This is not surprising, as the waiting-time distribution is defined by two different equations; namely (4.2) for  $x < b$  and (4.3) for  $x \geq b$ . Furthermore, from (4.5) we have that

$$f_W(b^-) = \lambda F_W(b) - \lambda(1 - p)(1 - \pi_0),$$

and from (4.12) we have that

$$f_W(b^+) = \lambda(F_W(b) - 1 + p - p\pi_0).$$

That is,  $f_W(b^-) - f_W(b^+) = \lambda\pi_0$ .

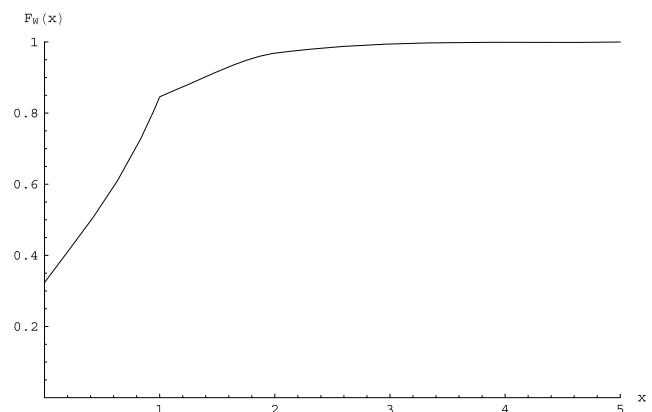


Fig. 1 The waiting time distribution for  $b = 1$ ,  $\lambda = 2$ , and  $p = 1/3$

The case  $p = 0$

Observe that if  $p = 0$  then the support of  $W$  is the interval  $[0, b]$ . To determine the density of the waiting time, we insert  $p = 0$  into (4.8). Thus, we obtain that

$$f''_W(x) = 0,$$

from which we immediately have that

$$f_W(x) = v_1x + v_2,$$

for some constants  $v_1$  and  $v_2$  such that (4.6) is satisfied. The latter condition implies that for every  $x \in (0, b)$  the following equation must hold:

$$v_1 = \lambda(v_1x + v_2) - \lambda(v_1(b - x) + v_2).$$

From this we conclude that  $v_1$  is equal to zero, i.e. the waiting time has a mass at zero and is uniformly distributed on  $(0, b)$ . To determine the mass  $\pi_0$  and the constant  $v_2$ , we evaluate (4.5) at  $x = 0$  and we use the normalisation (4.4), keeping in mind that  $f_W(x) = 0$  for  $x \in [b, \infty)$ . These two equations yield that if  $p = 0$ , then

$$f_W(x) = \frac{\lambda}{1 + \lambda b}, \quad 0 < x < b, \quad \text{and} \quad \pi_0 = \frac{1}{1 + \lambda b}. \tag{4.17}$$

Evidently, the density in this case is quite different from the density for  $p \neq 0$ , which is on  $(0, b)$  a mixture of two exponentials; see (4.9).

Another way to see that  $f_W(x) = \lambda\pi_0$ ,  $0 < x < b$ , is as follows. Recall that for  $p = 0$  and  $x \geq b$  we have that  $f_W(x) = 0$ . Equation (4.5) can now be written as

$$f_W(x) = \lambda\pi_0 + \lambda\mathbb{P}[W \in (0, x)] - \lambda\mathbb{P}[W \in (b - x, b)].$$

Replacing  $x$  by  $b - x$  shows that  $f_W(x) = f_W(b - x)$ , which implies that  $\mathbb{P}[W \in (0, x)] = \mathbb{P}[W \in (b - x, b)]$  and finally that  $f_W(x) = \lambda\pi_0$ ,  $0 < x < b$ . It seems less straightforward to explain probabilistically that  $W$ , given that  $W > 0$ , is uniformly distributed. With a view towards the recursion  $W \stackrel{D}{=} \max\{0, b - A - W\}$ , we believe that this property is related to the fact that, if  $n$  Poisson arrivals occur in some interval, then they are distributed like the  $n$  order statistics of the uniform distribution on that interval; see Ross [14, Sect. 2.3].

The case  $p = 1$

For the M/D/1 queue, Erlang [7] derived the following expression for the waiting-time distribution:

$$\mathbb{P}[W \leq x] = (1 - \rho) \sum_{j=0}^i \frac{(-\lambda(x - jb))^j}{j!} e^{\lambda(x - jb)},$$

$$ib \leq x < (i + 1)b,$$

where  $\rho$  is the traffic intensity. Recall that for the M/D/1 queue we have that  $F_W(0) = 1 - \rho$ . We see that for  $p = 1$  (4.5) indeed leads to the waiting-time distribution  $(1 - \rho)e^{\lambda x}$ , as it is given by Erlang’s expression for the first interval  $[0, b)$ . For  $x \geq b$ , one needs to recursively solve (4.13) in order to obtain Erlang’s expression. However, since the recursive solution we have obtained for our model makes use of  $F_W(x)$  as it is given by (4.11), which is not valid for  $p = 1$ , the waiting-time distribution we have obtained in Theorem 3 cannot be extended to the case for  $p = 1$ .

The terms both in Erlang’s expression for the waiting-time distribution of an M/D/1 queue and in Theorem 3 alternate in sign and in general are much larger than their sum. Thus, the numerical evaluation of the sum may be hampered by roundoff errors due to the loss of significant digits, in particular under heavy traffic. For the M/D/1 queue, however, a satisfactory solution has been given by Franx [8] in a way that only a finite sum of positive terms is involved; thus, this expression presents no numerical complications, not even for high traffic intensities. For our model, extending Franx’s approach is a challenging problem as the representation of various quantities appearing in [8] which are related to the queue length at service initiations is not straightforward.

As we see, the waiting-time distribution in Theorem 3 is quite similar to Erlang’s expression, so we expect that eventually the solution will suffer from roundoff errors. However, a significant difference in the numerical computation between the M/D/1 queue and the model described by recursion (1.2) arises when computing  $\pi_0$ . For any single server queue we know a priori that  $\mathbb{P}[W = 0] = 1 - \rho$ . In our model,  $\pi_0$  has to be computed from the normalisation equation, where the numerical complications when calculating the waiting-time distribution become apparent. In particular, as  $p$  tends to 1, i.e. as the system behaves almost like an M/D/1 queue, the computation of  $\pi_0$  becomes more problematic.

As a final observation, we note that the effects of Lindley’s classical recursion and of the Lindley-type recursion discussed in [20] are quite apparent. The analysis for our model is in a sense separated into two parts: the derivation of the waiting-time distribution in  $[0, b)$  and in  $[b, \infty)$ . In the first part, we see that (4.6) is quite similar to the differential equation appearing in [22] for the derivation of the waiting-time distribution in case  $p = 0$  and  $B$  follows a polynomial distribution. Moreover, one could use the same technique to derive a solution, but (4.6) is too simple to call for such means. In the second part, we see the effects of the M/D/1 queue, as we eventually derive  $F_W$  in a recursive manner. Furthermore, this model inherits all the numerical difficulties appearing in the classical solution for the M/D/1 queue, plus the additional difficulties of computing  $\pi_0$ . For Lindley’s recursion,  $\pi_0$  is known beforehand, while for the

Lindley-type recursion described in [13, 20–23]  $\pi_0$  is derived by the normalisation equation.

**Acknowledgements** The authors would like to thank Ivo Adan for his helpful remarks and his assistance in simulating this model.

## References

1. Asmussen S, Schock Petersen S. Ruin probabilities expressed in terms of storage processes. *Adv Appl Probab* 1989;**20**(4):913–916.
2. Asmussen S, Sigman K. Monotone stochastic recursions and their duals. *Probab Eng Inf Sci* 1996;**10**(1):1–20.
3. Borovkov AA, Foss S. Stochastically recursive sequences. *Sib Adv Math* 1992;**2**:16–81.
4. Brandt A. The stochastic equation  $Y_{n+1} = A_n Y_n + B_n$  with stationary coefficients. *Adv Appl Probab* 1986;**18**(1):211–220.
5. Brandt A, Franken P, Lisek B. Stationary stochastic models. vol. 78, *Mathematische Lehrbücher und Monographien, II. Abteilung: Mathematische Monographien*. Berlin: Akademie Verlag; 1990.
6. Cohen JW. The single server queue. Amsterdam: North-Holland; 1982.
7. Erlang AK. The theory of probabilities and telephone conversations. In: Brockmeyer E, Halstrøm HL, Jensen A, editors. *The life and works of A.K. Erlang. Applied mathematics and computing machinery series*. vol 6. Acta polytechnica Scandinavica. 2nd ed. 1960, p. 131–137. English translation. First published in “Nyt tidsskrift for matematik” B. vol 20. 1909, p. 33.
8. Franx GJ. A simple solution for the M/D/c waiting time distribution. *Oper Res Lett* 2001;**29**(5):221–229.
9. Jacquet P. Subexponential tail distribution in LaPalice queues. *Perform Eval Rev* 1992;**20**(1):60–69.
10. Kalashnikov V, Norberg R. Power tailed ruin probabilities in the presence of risky investments. *Stoch Process Appl* 2002;**98**:211–228.
11. Lindley DV. The theory of queues with a single server. *Proc Camb Philos Soc* 1952;**48**:277–289.
12. Norberg R. Ruin problems with assets and liabilities of diffusion type. *Stoch Process Appl* 1999;**81**:255–269.
13. Park BC, Park JY, Foley RD. Carousel system performance. *J Appl Probab* 2003;**40**(3):602–612.
14. Ross SM. *Stochastic processes*. 2nd ed. New York: Wiley; 1996.
15. Schassberger R. *Warteschlangen*. Wien: Springer; 1973.
16. Seal HL. Risk theory and the single server queue. *Mitt Ver Schweiz Versicher* 1972;**72**:171–178.
17. Tang Q, Tsitsiashvili G. Precise estimates for the ruin probability in finite horizon in a discrete-time model with heavy-tailed insurance and financial risks. *Stoch Process Appl* 2003;**108**:299–325.
18. Titchmarsh EC. *Theory of functions*. 2nd ed. London: Oxford University Press; 1968.
19. Vervaat W. On a stochastic difference equation and a representation of non-negative infinitely divisible random variables. *Adv Appl Probab* 1979;**11**(4):750–783.
20. Vlasiou M. A non-increasing Lindley-type equation. Technical Report 2005-015, Eurandom, Eindhoven, The Netherlands, 2005. Available at <http://www.eurandom.nl>.
21. Vlasiou M, Adan IJBF. An alternating service problem. *Probab Eng Inf Sci* 2005;**19**(4):409–426.
22. Vlasiou M, Adan IJBF. Exact solution to a Lindley-type equation on a bounded support. *Oper Res Lett* 2007;**35**(1):105–113.
23. Vlasiou M, Adan IJBF, Wessels J. A Lindley-type equation arising from a carousel problem. *J Appl Probab* 2004;**41**(4):1171–1181.
24. Whitt W. Queues with service times and interarrival times depending linearly and randomly upon waiting times. *Queueing Syst* 1990;**6**(4):335–351.