Performance Analysis for Bernoulli Feedback Queues Subject to Disasters: A System with Batch Poisson Arrivals Under a Multiple Vacation Policy

George C. Mytalas
Dept. of Mathematics, CUNY,
Brooklyn, NY 11210, USA
and
Michael A. Zazanis
Dept. of Statistics, Athens University of Economics and Business
Athens 10434, Greece

Abstract

We analyze an $M/G/1$ system with batch Poisson arrivals and instantaneous Bernoulli feedback, operating under a Multiple Vacation Policy. The system is subject to disasters that occur according to an independent Poisson process and are followed by (random) repair periods with general distribution. The analysis is carried out using the supplementary variable method. The Laplace transform of the time between two consecutive disasters is obtained and the existence of the stationary regime for the system is shown. Besides obtaining the stationary distribution for the number of customers in the system, we use the information regarding the rates of occurrence of various events provided by the supplementary variables solution to obtain a great variety of additional results. These include the Laplace transform of the busy period distribution and the probability that a customer completes service. We indicate areas of application of our model to real life systems, and we use the analytic results obtained to optimize such a system under a Quality of Service constraint. Finally, we analyze a variant of the system subject to disasters even when the server is not busy.

KEYWORDS: BATCH-ARRIVALS, MULTIPLE VACATIONS, FEEDBACK, DISASTERS, OPTIMIZATION.

1 Introduction

We study a single server queue with batch Poisson arrivals, general service times, and Bernoulli feedback. After completing service a customer either joins immediately the tail of the waiting line as a feedback customer with probability r, or departs forever from the system with probability $1 - r$, independently of anything else. The system suffers from disasters which occur according to an independent Poisson process with rate δ and which instantly remove all customers (both in the queue and in service) from the system. Immediately after the occurrence of a disaster the server undergoes a repair of random duration. Further, whenever a departing customer leaves the system empty, the server takes vacations of random length according to a Multiple Vacation policy. We analyze in detail this model under the assumption that disasters affect the system only when the server is busy serving customers. We also examine a variation of this model in which disasters affect the system when the server is under repair or on vacation as well. We will refer to this modified model as a system with *unconditional disasters*.

The proposed model can be applied in the analysis of a storage area network (SAN) or an email contact center subject to random failures. Storage Area Networks are highly reliable solutions for mass storage of data and as such they can become primary targets of distributed denial-of-service-attacks (DDoS attacks) Such attacks are assumed to occur according to a Poisson process with rate δ. Requests for obtaining data arrive according to a Poisson arrival process with rate λ and we allow for batch arrivals to increase the flexibility of the model. Each request, when processed, may result in the need for further data retrieval and in this case the request joins the queue as a feedback customer. When the system becomes idle, maintenance and other secondary operations are initiated, corresponding to vacations. When a disaster in the form of a DDoS attack occurs, the system is down for the duration of the attack to which is added the duration of restorative operations. This corresponds to the repair period of our model.

The combination of features we study, namely batch arrivals, multiple vacations, Bernoulli feedback, and the presence of disasters has not been studied in the vast literature of such queueing systems. Of course these features have been studied separately and in other combinations. When these features interact, new difficulties and questions arise. The present paper addresses some of them. In the queueing literature during the last two decades there has been a number of papers analyzing the interplay of feedback or retrials with vacations. Among those we mention [3], [6], [14], [24], [33], [29]. Also, a number of papers have analyzed the behavior of queueing systems with feedback or retrials, subject to disasters or unreliable servers. See in particular [5], [11], [12], [21], [16]. Also [8], [23], are among the papers who have studied queueing systems with vacations and disasters or unreliable servers.

The present paper is the first to study the interplay of vacations, Bernoulli feedback, and disasters with subsequent repairs. We analyze the system using the supplementary variables technique and we obtain the joint stationary distribution of the number of customers in the system and the supplementary variable, which is the age of the service, vacation, or repair time, according to the state of the server. We take full advantage of the additional information regarding the supplementary variables in order to obtain the *rates of occurrence of various events, and from these, the corresponding Palm (event-stationary) distributions.* In particular, we obtain the distribution of the number of customers at the epoch of a typical service completion in §3.2, and at a typical busy period initiation epoch in §4.1. The Laplace transform of busy period is obtained in §4.2 and several performances measures in §5. In §6 we examine the cycle structure of the model and obtain the Laplace transform of the time between two consecutive disasters and the stability condition for the system. The case where system suffers from disasters not only when the server is operational but also during repair and vacation periods is examined in §8. Finally in §7 and §9 we use of the performance measures and results obtained to discuss the problem of determining the optimal service rate while satisfying a Quality of Service constraint and to obtain numerical results.

For background on queues with vacations, a subject that has been studied thoroughly, we refer the reader to the books of Takagi [28] and Tian and Zhang [30]. A brief summary of the research on vacation queueing systems during the decade 2000-2010 is provided by Ke et al. [15]. For an account of the latest results on vacation queueing systems in the past decade see Jain et al. [13] who also use the supplementary variables technique (SVT).

Random events (failures) that occur during the busy period may cause a server breakdowns and service interruptions. Various types of system failures with corresponding effects on the customers present have been studied in the literature including simple failures that do not affect the customers present, and failures that affect the customer in service, acting in essentially as “negative customers”. Such systems are known as G-systems. The type of server failures we consider here are *disasters* (also refereed to as catastrophes) causing the removal of all customers from the system (queue and server). Queues with disasters are natural models for communication or computer systems subject to catastrophic failures resulting in the loss of all entities currently in process, or waiting to be processed. Jain and Sigman [10] analyzed $M/G/1$ queues in the presence of disasters and derived a corresponding generalization of the Pollaczek-Khinchine formula. Yechiali [37] studied $M/M/c$ queues with disasters and customers that become impatient during the repair periods following disasters.

Disasters may for instance represent Distributed Denial of Service (DDoS) attacks to the servers of Storage Area Networks (SAN) used to provide high reliability mass storage of data. When these DDoS occur they cause network resources and data to become unavailable to their intended clients. The client requests affected by the DDoS correspond to customers that are removed by the disaster. This application is discussed in Kim and Lee [19] who analyze a queueing system with disasters and server breakdowns using the supplementary variables technique. Communication systems using intrinsically unreliable channels subject to clearing events may also be modelled as queues with disasters. In manufacturing systems disasters may represent catastrophic failures resulting in machine breakdowns and damaging work in process inventory.

The busy period of $M/G/1$ queues (with processor sharing discipline) have been studied by Yashkov and Yashkova [36]. Semenova [25] studied queues with hysteresis control in the presence of disasters that occur according to a Markov Arrival Process (MAP). The transient behavior of markovian queues subject to disasters have been studied by Kumar et al. [20] and recently by Jain and Singh [12] who consider a markovian queue with feedback. Recently, Li and Wang [22] analyzed the equilibrium balking strategies in markovian queues with retrials subject to disasters, while Sun et al. [26] compared systems with normal failures and site-clearing disasters. Mytalas and Zazanis [23] analyzed a batch queueing model, operating under a MAV (Multiple Adapted Vacations) policy and subject to disasters.

Queueing systems with feedback loops are natural models of many computing, communications, and manufacturing systems. In communication systems for instance feedback arises in packet transmission in error-prone channels and in segmented (batch form) message transmission. In computing systems the round-robin service discipline results in feedback queues. In manufacturing systems feedback occurs when rework of a part is required in a processing station, for instance when modeling the performance of a manufacturing cell consisting of several CNC machines and inspection facilities. $M/G/1$ queues with Bernoulli feedback were first analyzed in the seminal paper of Takács [27]. Bernoulli feedback queues with vacations were studied by Takagi [28] and Wortman and Disney [33]. A batch arrival queueing system with feedback and optional server vacations under single vacation policy was studied by Madan and Al-Rawwash [24]. A system using the multiple vacation policy for server and consider feedback and two phase service for customers was studied by Thangaraj and Vanitha [29]. A general retrial queue with balking and feedback together with a modified vacation policy was studied by Ke and Chang [14]. Recent research work, including Bernoulli vacations, retrials. optional services and feedbacks given by Jain and Kaur [11]. Liu et al. [21] studied a system with retrials, feedback, and server breakdowns due to "negative customers". Also, Ke et al. [16] studied retrial systems with feedback and disasters. For $M/G/1$ feedback queues with two classes of customers and gated vacations, Boxma and Yechiali [3] and Choi et al. [6] derived the distributions of the queue length and the waiting time. Kim et. al. [18] analyze a multiclass queueing system with markovian feedback and its application to weighted round robin policies.

The present paper is the first to study the interplay of vacations, Bernoulli feedback, and disasters with subsequent repairs.

2 Model description and notations

Consider a single server queue with customers arriving in batches according to a Poisson process with rate λ. Let χ_n denote the size of the nth batch. $\{\chi_n\}$, $n = 1, 2, \ldots$, is assumed to be an i.i.d. sequence. Let

$\chi(z) = \sum_{n=1}^{\infty} \mathbb{P}(\chi_1 = n) z^n$ denote the corresponding probability generating function. We do not assume the mean to be finite. When this assumption is necessary it will be stated explicitly and the mean batch size will be denoted by m_χ. The mean batch size, when finite will be denoted by $m_\chi := \mathbb{E}\chi$. Service times are assumed to be i.i.d. random variables with common distribution S, corresponding density S' and hazard rate function $\mu(x) = \frac{S'(x)}{1-S(x)}$, $x \geq 0$. The Laplace transform of S will be denoted by $\hat{S}(s) := \int_0^\infty e^{-sx} dS(x)$. Again, its mean will not be assumed finite. When this assumption becomes necessary it will be stated clearly and the mean will be denoted by $m_S := \mathbb{E}S$. When a customer completes service, he leaves the system with probability $1-r$ (independently of everything else) or returns to the tail of the queue as a feedback customer.

The system is subject to disasters when the server is busy. These occur according to a Poisson process with rate δ, independently of all other processes in the system. When a disaster occurs all customers present, including the one in service, are removed from the system. Immediately after a disaster a repair period begins. These periods have i.i.d. durations with distribution function R, density function R', hazard rate $r(x) = \frac{R'(x)}{1-R(x)}$, $x \geq 0$, and Laplace transform $\hat{R}(s) = \int_0^\infty e^{-sx} dR(x)$. The mean of the repair period, *which will be assumed finite throughout the paper* is $m_R := \mathbb{E}R < \infty$. Any customers that may arrive during a repair period wait in line.

At the end of a busy period or at the end of a repair period if no customers are present, the server takes a vacation. We assume that the system implements a multiple vacation policy, according to which the server takes repeated vacations until at the end of a vacation there is at least one customer present, waiting to be served. At this point a new busy period begins. Vacation durations are i.i.d. random variables with common distribution function U, density U', and hazard rate $u(x) = \frac{U'(x)}{1-U(x)}$, $x \geq 0$. The corresponding Laplace transform will be denoted by $\hat{U}(s) = \int_0^\infty e^{-sx} dU(x)$ and the mean $m_U := \mathbb{E}U < \infty$ *is assumed finite*. Service times, vacation and repair durations, batch sizes, the Poisson arrival process, and the feedback decision are assumed to be independent.

3 Supplementary Variables Analysis

In this section we derive the steady-state differential-difference equations for the system by treating the elapsed service time, the elapsed repair time, and the elapsed vacation time as supplementary variables. Consider the processes

$$\begin{aligned}
\mathcal{N}_t &: \text{ number of customers in the system at time } t \\
\mathcal{S}_t &: \text{ elapsed service time at time } t \text{ (if server is busy, otherwise 0)} \\
\mathcal{R}_t &: \text{ elapsed repair time at time } t \text{ (if server under repair, otherwise 0)} \\
\mathcal{U}_t &: \text{ elapsed vacation time at time } t \text{ (if server on vacation otherwise 0)}
\end{aligned}$$

and the process $\{\xi_t\}$ taking values in the set $\{\boldsymbol{s}, \boldsymbol{r}, \boldsymbol{v}\}$, whose elements correspond to the server being busy *serving* customers, being under *repair*, and being on *vacation* respectively. Due to the presence of disasters the system has a regenerative structure with the epochs of consecutive disasters acting as regeneration points. In section 6 the Laplace transform of the time between two consecutive disasters is derived and it is shown to have a finite mean. Therefore a stationary version of the process exists by virtue of standard results on regenerative processes (e.g. see [1]). Suppose the process is stationary under the probability measure $\mathbb{P}$ and define the densities

$$
\begin{aligned}
P_n(x) &:= \lim_{h\downarrow 0}\frac{1}{h}\mathbb{P}(\mathcal{N}_0=n,\xi_0=\boldsymbol{s};\, x<\mathcal{S}_0\le x+h),\\
W_n(x) &:= \lim_{h\downarrow 0}\frac{1}{h}\mathbb{P}(\mathcal{N}_0=n,\xi_0=\boldsymbol{r};\, x<\mathcal{R}_0\le x+h),\\
V_n(x) &:= \lim_{h\downarrow 0}\frac{1}{h}\mathbb{P}(\mathcal{N}_0=n,\xi_0=\boldsymbol{v};\, x<\mathcal{U}_0\le x+h).
\end{aligned} \tag{1}
$$

The balance equations satisfied by the stationary distribution are

$$\frac{d}{dx}P_n(x)+(\lambda+\delta+\mu(x))\,P_n(x)=\lambda\sum_{k=1}^{n-1}\chi_k P_{n-k}(x),\quad x>0,\ n\ge 1 \tag{2}$$

$$\frac{d}{dx}W_0(x)+(\lambda+r(x))W_0(x)=0 \tag{3}$$

$$\frac{d}{dx}W_n(x)+(\lambda+r(x))W_n(x)=\lambda\sum_{k=1}^{n}\chi_k W_{n-k}(x),\quad x>0, n\ge 1 \tag{4}$$

$$\frac{d}{dx}V_0(x)+(\lambda+u(x))V_0(x)=0, \tag{5}$$

$$\frac{d}{dx}V_n(x)+(\lambda+u(x))V_n(x)=\lambda\sum_{k=1}^{n}\chi_k V_{n-k}(x),\quad x>0,\ n\ge 1 \tag{6}$$

The boundary conditions of the above system of differential equations are

$$
\begin{aligned}
P_n(0) &= (1-r)\int_0^\infty P_{n+1}(x)\mu(x)dx + r\int_0^\infty P_n(x)\mu(x)dx\\
&\quad + \int_0^\infty V_n(x)u(x)dx + \int_0^\infty W_n(x)r(x)dx,\qquad n\ge 1 \qquad (7)\\
V_0(0) &= (1-r)\int_0^\infty P_1(x)\mu(x)dx + \int_0^\infty V_0(x)u(x)dx + \int_0^\infty W_0(x)r(x)dx \qquad (8)\\
V_n(0) &= 0,\quad n\ge 1, \qquad (9)\\
W_0(0) &= \delta\sum_{n=1}^{\infty}\int_0^\infty P_n(x)dx, \qquad (10)
\end{aligned}
$$

with normalization condition

$$\sum_{n=1}^{\infty}\int_0^\infty P_n(x)dx+\sum_{n=0}^{\infty}\left(\int_0^\infty W_n(x)dx+\int_0^\infty V_n(x)dx\right)=1. \tag{11}$$

Define the generating functions

$$P(x;z):=\sum_{n=1}^{\infty}z^nP_n(x),\quad W(x;z):=\sum_{n=0}^{\infty}z^nW_n(x),\quad V(x;z):=\sum_{n=0}^{\infty}z^nV_n(x).$$

The partial probability generating functions (pgf) for the number of customers in the system in stationarity regardless of the value of the supplementary variables are then given by

$$P(z):=\int_0^\infty P(x;z)dx,\quad W(z):=\int_0^\infty W(x;z)dx,\quad V(z):=\int_0^\infty V(x;z)dx. \tag{12}$$

Proposition 1. *The partial pgf for the number of customers in the system when the server is busy is given by*

$$P(z) \;=\; P(0;z)\,\frac{1-\hat{S}(\delta+\alpha(z))}{\delta+\alpha(z)} \tag{13}$$

where

$$P(0;z) = z\,\frac{\delta P(1)\hat{R}(\alpha(z)) - V_0(0)\left(1-\hat{U}(\alpha(z))\right)}{z-(1-r+rz)\hat{S}(\delta+\alpha(z))}. \tag{14}$$

and $\alpha(z) := \lambda(1-\chi(z))$.

Proof. Multiplying (2), (4), (6), and (7) by z^n and adding we obtain the linear first order PDE's

$$\begin{aligned}
\frac{\partial P(x;z)}{\partial x} + (\alpha(z)+\delta+\mu(x))P(x;z) &= 0,\\
\frac{\partial W(x;z)}{\partial x} + (\alpha(z)+r(x))W(x;z) &= 0,\\
\frac{\partial V(x;z)}{\partial x} + (\alpha(z)+u(x))V(x;z) &= 0,
\end{aligned} \tag{15}$$

and the equation

$$\begin{aligned}
\sum_{n=1}^{\infty} z^n P_n(0) \;=\;& \int_0^\infty \sum_{n=1}^{\infty} z^n V_n(x)u(x)dx + (1-r)\int_0^\infty \sum_{n=1}^{\infty} z^n P_{n+1}(x)\mu(x)dx\\
&+\int_0^\infty \sum_{n=1}^{\infty} z^n W_n(x)r(x)dx + r\int_0^\infty \sum_{n=1}^{\infty} z^n P_n(x)\mu(x)dx
\end{aligned} \tag{16}$$

which, taking into account that $\sum_{n=1}^{\infty} V_n(x)z^n = V(x;z)-V_0(x)$, $\sum_{n=1}^{\infty} W_n(x)z^n = W(x;z)-W_0(x)$, and $\sum_{n=1}^{\infty} P_{n+1}(x)z^n = z^{-1}P(x;z)-P_1(x)$, gives

$$\begin{aligned}
P(0;z) = z^{-1}(1-r)\int_0^\infty P(x;z)\mu(x)dx - (1-r)\int_0^\infty P_1(x)\mu(x)dx + r\int_0^\infty P(x;z)\mu(x)dx\\
+\int_0^\infty V(x;z)u(x)dx - \int_0^\infty V_0(x)u(x)dx + \int_0^\infty W(x;z)r(x)dx - \int_0^\infty W_0(x)r(x)dx.
\end{aligned} \tag{17}$$

The PDE's (15) have the solution

$$\begin{aligned}
P(x;z) &= P(0;z)(1-S(x))e^{-(\delta+\alpha(z))x},\\
W(x;z) &= W(0;z)(1-R(x))e^{-\alpha(z)x},\\
V(x;z) &= V(0;z)(1-U(x))e^{-\alpha(z)x}.
\end{aligned} \tag{18}$$

Since no customers are present when a vacation or a repair period starts, it holds that

$$V(0;z) = V_0(0) \qquad \text{and} \qquad W(0;z) = W_0(0). \tag{19}$$

Also, from (3) and (5) we obtain

$$W_0(x) = W_0(0)(1-R(x))e^{-\lambda x}, \qquad V_0(x) = V_0(0)(1-U(x))e^{-\lambda x}. \tag{20}$$

From (18) we obtain

$$\int_0^\infty P(x;z)\mu(x)dx = \int_0^\infty P(0;z)(1-S(x))e^{-(\delta+\alpha(z))x}\mu(x)dx = P(0;z)\hat{S}(\delta+\alpha(z)), \tag{21}$$

$$\int_0^\infty W(x;z)r(x)dx = \int_0^\infty W(0;z)(1-R(x))e^{-\alpha(z)x}r(x)dx = W(0;z)\hat{R}(\alpha(z)), \tag{22}$$

$$\int_0^\infty V(x;z)u(x)dx = \int_0^\infty V(0;z)(1-U(x))e^{-\alpha(z)x}u(x)dx = V(0;z)\hat{U}(\alpha(z)). \tag{23}$$

Similarly from (20) we have

$$\int_0^\infty W_0(x)r(x)dx = W_0(0)\int_0^\infty e^{-\lambda x}dR(x) = W_0(0)\hat{R}(\lambda), \tag{24}$$

$$\int_0^\infty V_0(x)u(x)dx = V_0(0)\int_0^\infty e^{-\lambda x}dU(x) = V_0(0)\hat{U}(\lambda). \tag{25}$$

Using (21)–(25) in (17) we obtain

$$\begin{aligned} P(0;z) = {} & V(0;z)\hat{U}(\alpha(z)) + z^{-1}(1-r)P(0;z)\hat{S}(\delta+\alpha(z)) + W(0;z)\hat{R}(\alpha(z)) \\ & -V_0(0) + rP(0;z)\hat{S}(\delta+\alpha(z)). \end{aligned}$$

Using (18) we write (10) as

$$W_0(0) = \delta\sum_{n=1}^{\infty}\int_0^\infty P_n(x)dx = \delta\int_0^\infty P(x;1)dx = \delta P(1) \tag{26}$$

whence we obtain (14). □

We can prove using Rouche's theorem (see the Appendix) that the denominator of $P(0;z)$ (given in (14)) has a unique root, z_0, in the open unit disk $|z|<1$. However, for $\delta>0$, the system is stable for all values of the other parameters due to the presence of disasters. This implies that the power series that defines $P(0;z)$ converges uniformly on the closed unit disk $|z|\le 1$ and defines an analytic function there. Therefore, the numerator of $P(0;z)$ must also vanish at z_0 which implies that the, as of yet, unknown constants $P(0)$ and $V_0(0)$ are connected by the relationship

$$V_0(0)\left(1-\hat{U}(\alpha(z_0))\right) = \delta P(1)\hat{R}(\alpha(z_0)).$$

Setting

$$\gamma := \frac{\hat{R}(\alpha(z_0))}{1-\hat{U}(\alpha(z_0))} \tag{27}$$

and taking into account (26) we thus have

$$V_0(0) = \gamma W_0(0). \tag{28}$$

Since from the above, γ is the ratio of the rate of vacation initiation to the rate of repair initiation. This means that γ gives the relative frequency of vacation periods to repairs. Also, It will be shown in section 4.2 (see (50)) that z_0 can be interpreted as the probability that a busy period of the system, *starting with a single customer,* will be completed without the occurrence of a disaster. Also, as can be readily seen from (28), γ is the relative frequency of occurrence of vacation periods relative to that of repairs.

3.1 Partial pgf's

Here we derive the stationary partial pgf's for the number of customers in the system according to the state of the server (working, under repair, or on vacation).

Server is working. From Proposition 1, (26), (28),

$$P(z) \;=\; \delta P(1)\, z\, \frac{\hat{R}(\alpha(z)) - \gamma\left(1 - \hat{U}(\alpha(z))\right)}{z - (1 - r + rz)\hat{S}(\delta + \alpha(z))}\, \frac{1 - \hat{S}(\delta + \alpha(z))}{\delta + \alpha(z)}. \tag{29}$$

Server is under repair. Taking into account (12), (18), and (26) we obtain

$$W(z) \;=\; \delta \frac{1 - \hat{R}(\alpha(z))}{\alpha(z)} P(1). \tag{30}$$

The probability that the server is under repair in stationarity can be obtained from the above using de l' Hôpital's rule as

$$W(1) \;=\; P(1)\delta m_R. \tag{31}$$

Server is on vacation. Taking into account (12), (18), and (28) we obtain

$$V(z) = \delta\gamma \frac{1 - \hat{U}(\alpha(z))}{\alpha(z)} P(1). \tag{32}$$

The probability that the server is on vacation, again using de l' Hôpital's rule is

$$V(1) \;=\; \delta P(1) m_U \gamma. \tag{33}$$

The above expression (with $P(1)$ given by (34) below) gives in practice the fraction of time the system will be available to perform the auxiliary tasks performed during the vacation time. In §7 this fraction of time is one of the factors taken into account in the optimization problem examined.

The probability that the server is busy. The probability that the server is busy, $P(1)$, is determined by the normalization condition $P(1) + W(1) + V(1) = 1$ which gives

$$P(1) = (1 \;+\; \gamma\delta m_U \;+\; \delta m_R)^{-1}. \tag{34}$$

(The dependence of $P(1)$ on the service time distribution and λ comes through z_0 and γ which depends on z_0.)

The pgf of the number of customers in the system in stationarity. The pgf of the number of customers in the system in stationarity, $\Phi(z) := \sum_{n=0}^{\infty} z^n \mathbb{P}(N_0 = n)$, is obtained by adding the above marginal pgf's. Setting

$$K(z) := \hat{R}(\alpha(z)) - \gamma\left(1 - \hat{U}(\alpha(z))\right) \tag{35}$$

and using (29), (30), and (32), we have $\Phi(z) = P(z) + W(z) + V(z)$ which gives

$$\begin{aligned}\Phi(z) \;=\; & \frac{\delta}{1 + \gamma\delta m_U + \delta m_R} \\ & \times \left[\frac{1 - \hat{S}(\delta + \alpha(z))}{z - (1 - r + rz)\hat{S}(\delta + \alpha(z))}\, \frac{zK(z)}{\delta + \alpha(z)} + m_R \hat{R}_e(\alpha(z)) + \gamma m_U \hat{U}_e(\alpha(z))\right].\end{aligned} \tag{36}$$

In the above, $\hat{U}_e(s) := \frac{1-\hat{U}(s)}{sm_U}$ and $\hat{R}_e(s) := \frac{1-\hat{R}(s)}{sm_R}$ denote the Laplace transform of the corresponding equilibrium (integrated tail) distributions of the vacation and repair periods. If the mean of the batch size distribution, m_χ, and the second moments of the repair and vacation distributions are finite, the expected number of customers in stationarity is finite and is given by

$$\Phi'(1) = \frac{\lambda m_\chi}{\delta} + \frac{1}{1+\delta m_R + \gamma\delta m_U}\left[-\frac{(1-r)\hat{S}(\delta)}{1-\hat{S}(\delta)} + \frac{\lambda m_\chi \delta}{2}(\mathbb{E}R^2 + \gamma\mathbb{E}U^2)\right]. \tag{37}$$

3.2 The pgf of the system size at departure and service completion epochs

Let $\{d_n\}$ denote the point process of departure epochs, corresponding to service completions that are followed by customer departures (as opposed to feedbacks). Let $\mathcal{N}_{d_n-}$ denote the number of customers in the system just before the customer's departure. Such a departing customer will leave behind l customers in the system if and only if there are $l+1$ customers in the system just before the departure i.e. if $\mathcal{N}_{d_n-} = l+1$. Denote by ϕ_l^+ the probability that, in steady state, a customer completing service leaves behind l customers in the system. Then

$$\phi_l^+ = \lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}\mathbf{1}(\mathcal{N}_{d_k-} = l+1) \quad \text{w.p. 1.} \tag{38}$$

Also,

$$\lim_{n\to\infty}\frac{n}{d_n} = D_0, \tag{39}$$

where D_0 is the customer departure rate (counting only normal departures and not departures due to disasters). The rate of customer departures that leave behind in the system l customers is equal to

$$\lim_{n\to\infty}\frac{1}{d_n}\sum_{k=1}^{n}\mathbf{1}(\mathcal{N}_{d_k-} = l+1) = (1-r)\int_0^\infty P_{l+1}(x)\mu(x)dx. \tag{40}$$

The right hand side of the above equation gives this rate in terms of the stationary solution we have obtained in this section. On the other hand since (w.p. 1) the limits exist, the left hand side can be written as

$$\lim_{n\to\infty}\frac{1}{d_n}\sum_{k=1}^{n}\mathbf{1}(\mathcal{N}_{d_k-} = l+1) = \lim_{n\to\infty}\frac{n}{d_n}\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}\mathbf{1}(\mathcal{N}_{d_k-} = l+1) = D_0\,\phi_l^+.$$

The type of argument sketched above is a ratio of rates argument (expounded for instance in [32]) that allows us to obtain *event stationary* probabilities such as ϕ_l^+.

$$\phi_l^+ = D_0^{-1}(1-r)\int_0^\infty P_{l+1}(x)\mu(x)dx, \qquad l = 0,1,2,\ldots, \tag{41}$$

where D_0^{-1}, the inverse of the total rate of departures, is the normalizing constant. If $\Phi^+(z) := \sum_{l=0}^\infty \phi_l^+ z^l$ denotes the corresponding pgf, then

$$\begin{aligned}
\Phi^+(z) &= D_0^{-1}(1-r)\int_0^\infty\sum_{l=0}^\infty z^l P_{l+1}(x)\mu(x)dx \\
&= D_0^{-1}(1-r)z^{-1}\int_0^\infty P(z;x)\mu(x)dx = C_0(1-r)z^{-1}P(z;0)\hat{S}(\delta+\alpha(z)) \\
&= D_0^{-1}(1-r)\hat{S}(\delta+\alpha(z))\delta P(1)\,\frac{\hat{R}(\alpha(z)) - \gamma\left(1-\hat{U}(\alpha(z))\right)}{z-(1-r+rz)\hat{S}(\delta+\alpha(z))}
\end{aligned}$$

where, in the last equation, we have used (14). F_0 can be determined from the condition $\Phi^+(1) = 1$ whence we obtain $D_0^{-1} = \frac{1-\hat{S}(\delta)}{\delta(1-r)(\gamma-1)P(1)\hat{S}(\delta)}$. Thus, taking into account (35),

$$\Phi^+(z) = \frac{\hat{S}(\delta+\alpha(z))}{z-(1-r+rz)\hat{S}(\delta+\alpha(z))} \frac{1-\hat{S}(\delta)}{\hat{S}(\delta)} K(z). \tag{42}$$

It is interesting that if we consider *all* departures, regardless of whether they are permanent or correspond to feedbacks, we obtain the same result. Indeed, if $\tilde{\phi}_l^+$ denotes the probability that a departing customer leaves behind l customers in steady state, then the same ratio of rates argument would give $\tilde{\phi}_l^+ = \tilde{D}_0^{-1} \int_0^\infty P_{l+1}(x)\mu(x)dx$.

3.3 The system without repairs

Here we suppose that the repair time following a disaster is negligible and thus following the occurrence of a disaster the server immediately takes a vacation immediately. The stationary pgf of the number of customers in the system can be obtained by setting $\hat{R}(s) = 1$ and $m_R = 0$ in (36). Thus

$$\Phi(z) = \frac{\delta\gamma}{1+\delta\gamma m_U} \left[z \frac{1-\hat{S}(\delta+\alpha(z))}{(1-r+rz)\,\hat{S}(\delta+\alpha(z))-z} \frac{\hat{U}(\alpha(z_0))-\hat{U}(\alpha(z))}{\delta+\alpha(z)} + \frac{1-\hat{U}(\alpha(z))}{\alpha(z)} \right]. \tag{43}$$

In the above expression, and in (44) and (45) as well, $\gamma = \left(1-\hat{U}(\alpha(z_0))\right)^{-1}$. When $m_\chi < \infty$ and $\mathbb{E}U^2 < \infty$ the corresponding mean number of customers in the system is finite and is given by

$$\Phi'(1) = \frac{\lambda m_\chi}{\delta} + \frac{1}{1+\gamma\delta m_U} \left(-\frac{(1-r)\hat{S}(\delta)}{1-\hat{S}(\delta)} + \frac{1}{2}\lambda m_\chi \delta\gamma \mathbb{E}U^2 \right). \tag{44}$$

If, furthermore, the mean service time, m_S, is finite then, as $\delta \downarrow 0$, $z_0 \uparrow 1$ and $m_U\delta\gamma \to \frac{1-r-\lambda m_S m_\chi}{\lambda m_S m_\chi}$ and therefore (43) becomes

$$\Phi(z) = \hat{U}_e(\alpha(z)) \left(1-\lambda \frac{m_S}{1-r} m_\chi\right) \frac{(1-z)(1-r)\hat{S}(\alpha(z))}{(1-r+rz)\hat{S}(\alpha(z))-z}, \tag{45}$$

thus recovering the well-known decomposition formula for queues with vacations, in the absence of disasters.

4 The Busy Period

4.1 The pgf of the System Size at a Busy Period Initiation Epoch

Let $\{t_l\}$, $l = 1, 2, \ldots$, be the busy period initiation epochs and $\mathcal{N}_{t_l}$ the number of customers in the system at the initiation epoch of the lth busy period. Let the probability that the typical busy period in stationarity starts with n customers be denoted by $\psi_n := \lim_{k\to\infty} \frac{1}{k} \sum_{l=1}^k \mathbf{1}(\mathcal{N}_{t_l} = n)$ and $\Psi(z) := \sum_{n=1}^\infty \psi_n z^n$ denote the corresponding pgf. By a ratio of rates argument, similar to that used in §3.2 we obtain

$$\psi_n = F^{-1} \left(\int_0^\infty W_n(x)r(x)dx + \int_0^\infty V_n(x)u(x)dx \right), \qquad n = 1, 2, \ldots \tag{46}$$

where F^{-1} is a normalization constant given by the total rate of busy period initiation

$$F \;=\; \sum_{n=1}^{\infty}\left(\int_0^{\infty} W_n(x)r(x)dx + \int_0^{\infty} V_n(x)u(x)dx\right).$$

Since $\sum_{n=1}^{\infty} z^n W_n(x) = W(x;z) - W_0(x)$, and $\sum_{n=1}^{\infty} z^n V_n(x) = V(x;z) - V_0(x)$, we conclude that

$$\Psi(z) = \frac{W_0(0)\left(\hat{R}(\alpha(z)) - \hat{R}(\lambda)\right) + V_0(0)\left(\hat{U}(\alpha(z)) - \hat{U}(\lambda)\right)}{W_0(0)\left(1 - \hat{R}(\lambda)\right) + V_0(0)\left(1 - \hat{U}(\lambda)\right)}$$

and, taking into account (28) and (35),

$$\Psi(z) \;=\; \frac{\hat{R}(\alpha(z)) - \hat{R}(\lambda) + \gamma\left(\hat{U}(\alpha(z)) - \hat{U}(\lambda)\right)}{1 - \hat{R}(\lambda) + \gamma\left(1 - \hat{U}(\lambda)\right)} \;=\; \frac{K(z) - K(0)}{1 - K(0)}. \tag{47}$$

4.2 The Laplace Transform of the Busy Period Length

The basis of our analysis is the remark that the length of the busy period of a queue with a work-conserving priority policy does not depend on the order in which customers are served. Therefore, we may suppose here that each customer, upon service completion, returns immediately to the server with probability $1 - r$ for additional service instead of joining the end of the queue. Thus the length of the busy period in our system is the same as the length of the busy period in a system without feedback in which each customer has a service time requirement whose Laplace transform is

$$\hat{S}_r(s) \;:=\; \sum_{n=1}^{\infty}(1-r)r^{n-1}\hat{S}(s)^n \;=\; \frac{(1-r)\hat{S}(s)}{1 - r\hat{S}(s)}. \tag{48}$$

In this framework, *for a system operating without disasters*, a standard argument gives the Laplace transform of the length of a busy period *starting with a single customer*, $\hat{\Gamma}_0(s)$, as the unique solution of the Takács equation $\hat{\Gamma}_0(s) = \hat{S}_r(s + \lambda - \lambda\chi(\hat{\Gamma}_0(s)))$ satisfying $|\hat{\Gamma}_0(s)| < 1$ for $s \in (0, \infty)$. Equivalently, taking into account (48),

$$\hat{\Gamma}_0(s) = \left(1 - r + r\hat{\Gamma}_0(s)\right)\, \hat{S}\left(s + \lambda - \lambda\chi(\hat{\Gamma}_0(s))\right). \tag{49}$$

In particular, comparing (49) with the fixed point problem of Proposition 5 of the Appendix, we see that

$$z_0 = \hat{\Gamma}_0(\delta). \tag{50}$$

Theorem 2. *The Laplace transform of the length of a busy period, $\hat{B}(s)$, in our system is given by*

$$\hat{B}(s) \;=\; 1 - \frac{s}{s+\delta}\,\frac{1 - \hat{R}(\alpha(\hat{\Gamma}_0(s+\delta))) + \gamma\left[1 - \hat{U}(\alpha(\hat{\Gamma}_0(s+\delta)))\right]}{1 - \hat{R}(\lambda) + \gamma\left[1 - \hat{U}(\lambda)\right]} \;=\; 1 - \frac{s}{s+\delta}\,\frac{1 - K(\hat{\Gamma}_0(s+\delta))}{1 - K(0)}. \tag{51}$$

The probability q_1 that a busy period ends normally (by the departure of a customer leaving the system empty), and q_2, that it ends by a disaster, are given by

$$q_1 \;=\; \frac{-\hat{R}(\lambda) + \gamma\left(1 - \hat{U}(\lambda)\right)}{1 - \hat{R}(\lambda) + \gamma\left(1 - \hat{U}(\lambda)\right)} \;=\; \frac{-K(0)}{1 - K(0)}, \tag{52}$$

$$q_2 \;=\; \frac{1}{1 - \hat{R}(\lambda) + \gamma\left(1 - \hat{U}(\lambda)\right)} \;=\; \frac{1}{1 - K(0)}. \tag{53}$$

Finally, the conditional Laplace transform of a busy period given that it ends normally, $\hat{B}_1(s)$*, or* given that it ends by a disaster, $\hat{B}_2(s)$ *are*

$$\begin{aligned}
\hat{B}_1(s) &= \frac{\hat{R}(\alpha(\hat{\Gamma}_0(s+\delta))) - \hat{R}(\lambda) + \gamma\left[\hat{U}(\alpha(\hat{\Gamma}_0(s+\delta))) - \hat{U}(\lambda)\right]}{-\hat{R}(\lambda) + \gamma\left[1 - \hat{U}(\lambda)\right]} \\
&= \frac{K(\hat{\Gamma}_0(s+\delta))) - K(0)}{-K(0)}, \qquad (54) \\
\hat{B}_2(s) &= \left[1 - \hat{R}(\alpha(\hat{\Gamma}_0(s+\delta))) + \gamma\left[1 - \hat{U}(\alpha(\hat{\Gamma}_0(s+\delta)))\right]\right] \frac{\delta}{\delta+s} \\
&= \frac{\delta}{\delta+s}\left[1 - K(\hat{\Gamma}_0(s+\delta))\right]. \qquad (55)
\end{aligned}$$

where

$$\hat{\Gamma}(s) = \Psi\left(\hat{\Gamma}_0(s)\right) \qquad (56)$$

and $\Psi(z)$ *is given by (47).*

Proof. Let us first imagine that the system operates in stationarity and that at the moment when a busy period is initiated, the disaster mechanism becomes inactive. Let Γ denote the duration of the busy period in such a situation. At an initiation epoch of a typical busy period the distribution of the number of customers present has pgf given by (47). Then the Laplace transform of Γ is given by

$$\hat{\Gamma}(s) = \sum_{n=1}^{\infty}\left(\hat{\Gamma}_0(s)\right)^n \psi_n \;=\; \Psi\left(\hat{\Gamma}_0(s)\right).$$

Let Δ be an exponential random variable with rate δ, independent of Γ. Then, q_1, the probability that a busy period terminates normally is given by $q_1 := \mathbb{P}(\Delta > \Gamma) = \hat{\Gamma}(\delta)$ and the complementary probability q_2 that it terminates by a disaster is $q_2 := \mathbb{P}(\Delta \le \Gamma) = 1 - \hat{\Gamma}(\delta)$. This establishes (52) and (53). A simple conditioning argument gives the Laplace transforms of typical busy period given that it terminates normally, $\hat{B}_1(s)$, or by a disaster, $\hat{B}_1(s)$,

$$\begin{aligned}
\hat{B}_1(s) &= \mathbb{E}[e^{-s(\Gamma\wedge\Delta)} \mid \Gamma < \Delta] = \frac{\hat{\Gamma}(s+\delta)}{\hat{\Gamma}(\delta)}, \\
\hat{B}_2(s) &= \mathbb{E}[e^{-s(\Gamma\wedge\Delta)} \mid \Gamma > \Delta] = \frac{1 - \hat{\Gamma}(s+\delta)}{1 - \hat{\Gamma}(\delta)} \frac{\delta}{s+\delta}.
\end{aligned}$$

Taking into account (27), (35), (47), (50), and (56) we obtain, after some algebraic manipulations, and taking also into account that $K(z_0) = 0$, (54) and (55). Finally, (51) follows from $\hat{B}(s) = q_1\hat{B}_1(s) + q_2\hat{B}_2(s)$. □

5 Performance Measures

C1. *Actual rate of disasters.* The rate of busy period terminations due to disasters is equal to the rate of repair initiations given by $W_0(0) = \delta P(1) = \frac{\delta}{1+\gamma\delta m_U+\delta m_R}$ where $P(1)$ is given by (34). This is of course the rate of actual disaster occurrence.

C2. *Expected length of a vacation string.* A vacation string consists of a number of vacations during which there were no arrivals and a final vacations during which at least one batch arrived. Since vacations are i.i.d. and batches arrive according to an independent Poisson process, if U denotes a generic vacation duration and Λ an independent exponential random variable with rate λ then $\beta := \hat{U}(\lambda) = \mathbb{P}(\Lambda > U)$. The expected number of vacations without arrivals is $\beta/(1-\beta)$ and the expected length of each such vacation is $\mathbb{E}[U|U < \Lambda]$. Thus the total expected length of the vacation string is

$$\frac{\beta}{1-\beta}\mathbb{E}[U|U < \Lambda] + \mathbb{E}[U|U \geq \Lambda],$$

the second term above corresponding to the length of the final vacation during which arrivals occur. Taking into account that $\beta\mathbb{E}[U|U < \Lambda] + (1-\beta)\mathbb{E}[U|U \geq \Lambda] = m_U$, we see that the expected length of the vacation string is

$$m_U(1-\beta)^{-1} \tag{57}$$

C3. *Rate of initiation of vacations.* This is clearly given by $V_0(0) = \gamma W_0(0)$ (see equation 28). Each vacation *string* following a busy period, or a repair period during which there were no arrivals, consists on the average of $(1-\beta)^{-1}$ vacations, each starting with no customers present. (Of course the last one ends with customers present.) Hence the rate of *vacation string* initiations is $V_0(0)(1-\beta)$.

C4. *Rate of normal busy period terminations, θ_1 and busy period terminations by disasters, θ_2.* Here, normal is meant to signify busy period terminations caused by a customer completing service and departing for good, leaving the system empty. Thus $\theta_1 = (1-r)\int_0^\infty P_1(x)\mu(x)dx$. Setting $b := \hat{R}(\lambda)$ and using (8) together with (24), (25),(26), and (28) we obtain

$$\begin{aligned}
\theta_1 &= (1-r)\int_0^\infty P_1(x)\mu(x)dx \;=\; V_0(0) - \int_0^\infty V_0(x)u(x)dx - \int_0^\infty W_0(x)r(x)dx \\
&= V_0(0) - V_0(0)\hat{U}(\lambda) - W_0(0)\hat{R}(\lambda) \;=\; V_0(0)(1-\beta) - W_0(0)b \\
&= \delta P(1)(\gamma(1-\beta) - b).
\end{aligned}$$

This makes perfect sense and can also be obtained by a different argument: Each normal busy period ending initiates a string of vacations. From the rate of vacation string initiations however we have to subtract the rate of repair initiations corresponding to repairs with no arrivals.

By a conditional PASTA argument we see that the rate of busy periods terminated by disasters is

$$\theta_2 \;=\; P(1)\delta \;=\; \frac{\delta}{1 + \delta m_R + \delta\gamma m_U}.$$

From these two rates we can obtain the probabilities q_1, q_2, obtained in (52), (53) by a ratio of rates argument:

$$q_1 \;:=\; \frac{\theta_1}{\theta_1+\theta_2} \;=\; \frac{\gamma(1-\beta)-b}{1-b+\gamma(1-\beta)}, \qquad q_2 \;:=\; \frac{\theta_2}{\theta_1+\theta_2} \;=\; \frac{1}{1-b+\gamma(1-\beta)}.$$

C5. *Expected length of an inactive period, $\mathbb{E}I$.* If the inactive period follows a busy period that terminates normally then its expected length is simply the expected length of a vacation string obtained in P6. If on the other hand the inactive period follows a busy period that terminates as a result of a disaster then its expected length is $m_R + b\frac{m_U}{1-\beta}$, then the second term is the sum corresponding to the probability that there are no arrivals during the repair period, b, multiplied by the expected length of a vacation string. Thus

$$\mathbb{E}I \;=\; q_1\frac{m_U}{1-\beta} + q_2\left(m_R + b\frac{m_U}{1-\beta}\right) \;=\; \frac{\gamma m_U + m_R}{1-b+\gamma(1-\beta)}.$$

C6. *Expected length of a busy period, $\mathbb{E}B$.* This could of course be obtained from (51). However it is easier to use instead renewal theoretic argument from which it follows that $P(1) = \frac{\mathbb{E}B}{\mathbb{E}B+\mathbb{E}I}$ whence we get

$$\mathbb{E}B \;=\; \mathbb{E}I\,\frac{1}{P(1)^{-1}-1} \;=\; \frac{\gamma m_U + m_R}{1-b+\gamma(1-\beta)}\,\frac{1}{\delta(\gamma m_U + m_R)} \;=\; \frac{1}{\delta}\,\frac{1}{1-b+\gamma(1-\beta)}. \tag{58}$$

This expression is striking enough to merit some commentary. Note that the service time distribution, or the feedback probability r for that matter, do not appear explicitly in the formula. They are buried in the parameter $\gamma = \frac{\hat{R}(z_0)}{1-\hat{U}(z_0)}$ through its dependence on $z_0(\delta)$ which is the unique solution of modulus less than 1 of equation (125) of the Appendix which we repeat here:

$$z_0 = \hat{S}(\delta + \lambda - \lambda\chi(z_0))(1 - r + rz_0). \tag{59}$$

In particular, $z_0(0) = 1$. Let us examine for simplicity the case where the repair time is negligible (i.e. $\hat{R}(s) \equiv 1$). Then $b = \hat{R}(\lambda) = 1$ and denoting the mean busy period $\mathbb{E}B$ by $m_B(\delta)$ we have

$$m_B(\delta) = \frac{1 - \hat{U}(\lambda - \lambda z_0(\delta))}{\delta} \frac{1}{1 - \hat{U}(\lambda)}. \tag{60}$$

Thus

$$\begin{aligned} m_B(0) &= \frac{1}{1 - \hat{U}(\lambda)} \lim_{\delta \to 0} \frac{1 - \hat{U}(\lambda - \lambda z_0(\delta))}{\delta} = \frac{1}{1 - \hat{U}(\lambda)} \left(-\hat{U}'(0)\right) \left(-\lambda z_0'(0)\right) \\ &= -\frac{\lambda m_U}{1 - \beta} z_0'(0). \end{aligned} \tag{61}$$

In the above we have taken into account that $z_0(0) = 1$ and $\hat{U}'(0) = m_U$. $z_0'(\delta)$ can be obtained by differentiating the implicit equation (59) with respect to δ: Thus $z_0' = \hat{S}'(\delta + \lambda - \lambda\chi(z_0)) (1 - \lambda\chi'(z_0)z_0') + \hat{S}(\delta + \lambda - \lambda\chi(z_0))rz_0'$ and setting $\delta = 0$ in this equation we obtain

$$z_0'(0) = -\frac{m_S}{1 - r - \lambda m_\chi m_S}.$$

Substitute this into (61) to obtain

$$m_B(0) = \frac{\lambda m_U}{1 - \beta} \frac{m_S}{1 - r - \lambda m_\chi m_S}.$$

This is indeed the mean duration of the busy period in a queue with no disasters, non-terminating vacations, and Bernoulli feedback.

In practice one may be interested in the length of the busy period for the following reason: Suppose that certain maintenance operations must be carried out during the vacation periods. Suppose also that, for the system to operate properly, the periods between maintenance (corresponding to busy periods) must not be too long. Results such as those given by (54), (55), (51), may be used to make sure, when designing the system, that the maintenance operations may be carried out at appropriate times.

C7. *The number of customers removed by a typical disaster.* Denote the corresponding pgf by $\phi_d(z)$. Then

$$\phi_d(z) = \frac{P(z)}{P(1)}, \tag{62}$$

i.e. it is equal to the stationary number of customers in the system, *conditional on the server being busy, and thus subject to disasters*. A rigorous proof of this can be given by appealing to Papangelou's theorem (see [2]). Heuristically, this follows from a PASTA type of argument, taking into account the fact that the intensity of the disaster process is δ, when the server is busy, and equal to zero when the server is under repair or on vacation. Thus the number of customers seen by a disaster is the stationary distribution of the number of customers conditional on the server being busy. Therefore, from (29),

$$\phi_d(z) = \delta z \frac{\hat{R}(\alpha(z)) - \gamma\left(1 - \hat{U}(\alpha(z))\right)}{z - (1 - r + rz)\hat{S}(\delta + \alpha(z))} \frac{1 - \hat{S}(\delta + \alpha(z))}{\delta + \alpha(z)}. \tag{63}$$

The mean number of customers eliminated by a disaster occurrence is given by

$$\phi_d'(1) = \frac{\lambda m_\chi}{\delta}\left(1 + \gamma\delta m_U + \delta m_R\right) - (1-r)\frac{\hat{S}(\delta)}{1-\hat{S}(\delta)}. \tag{64}$$

C8. *Rate of customer departures after completing service.* This is the rate of service completions that correspond to customers who leave the system for good (as opposed to being fed back). This is given by

$$\begin{aligned}(1-r)\int_0^\infty \sum_{n=1}^\infty P_n(x)\mu(x)dx &= (1-r)\int_0^\infty P(x;1)\mu(x)dx \;=\; (1-r)P(0;1)\hat{S}(\delta) \\ &= (1-r)\delta P(1)\frac{\hat{S}(\delta)}{1-\hat{S}(\delta)}. \end{aligned} \tag{65}$$

C9. *Rate of customer removals by disasters.* This is obtained by subtracting (65) from the customer arrival rate, λm_χ, which gives

$$\lambda m_\chi - (1-r)\delta P(1)\frac{\hat{S}(\delta)}{1-\hat{S}(\delta)}. \tag{66}$$

C10. *Fraction of customers which complete service.* This is obtained by a ratio of rates argument by dividing the rate of service completions that result to departures from the system, given by (65) by the rate of customer arrivals, λm_χ. We obtain

$$p \;=\; \frac{1}{\lambda m_\chi}\left((1-r)\delta P(1)\frac{\hat{S}(\delta)}{1-\hat{S}(\delta)}\right) \;=\; \frac{\delta(1-r)}{\lambda m_\chi(1+\gamma\delta m_U + \delta m_R)}\frac{\hat{S}(\delta)}{1-\hat{S}(\delta)}. \tag{67}$$

6 The Laplace transform of the time between two consecutive disasters

Here we provide a renewal theoretic analysis of the cycles of the system with multiple vacations. The sample path of the system consists of *cycles* which we define as the segments of the sample path of the system between consecutive disasters. These cycles consist in turn of *sub-cycles*: There is an *initial inactive-active sub-cycle* which consists of a repair time, possibly followed by a string of vacations (if there are no arrivals during the repair period). The active part of this sub-cycle is the busy period that ensues. If a disaster does not happen during this busy period it ends by a customer departure leaving the system empty. Then an *ordinary inactive-active sub-cycle* begins. It consists of a string of vacations and the ensuing busy period. Again, the busy period either ends as a result of a disaster, in which case the whole cycle ends, or as a result of a customer departure leaving the system empty. In this second case, a new ordinary inactive-active sub-cycle begins. Thus, a cycle consists of an *initial active-inactive sub-cycle* plus an number (possible zero) of *ordinary active-inactive sub-cycles*. By analyzing the structure of these sub-cycles we obtain the following

Proposition 3. *The Laplace Transform of the time between two consecutive disasters is given by*

$$\hat{\Omega}(s) \;=\; \frac{\delta}{s+\delta}\left(\hat{R}(s) - \hat{R}(s+\eta(s+\delta))\frac{1-\hat{U}(s)}{1-\hat{U}(s+\eta(s+\delta))}\right). \tag{68}$$

where

$$\eta(s) \;:=\; \alpha(\hat{\Gamma}_0(s)) \;=\; \lambda - \lambda\chi(\hat{\Gamma}_0(s)). \tag{69}$$

In particular, when m_U and m_R are finite, so is the mean time between two consecutive disasters, and it is given by

$$-\hat{\Omega}'(0) \;=\; \frac{1}{\delta} + m_R + m_U\gamma. \tag{70}$$

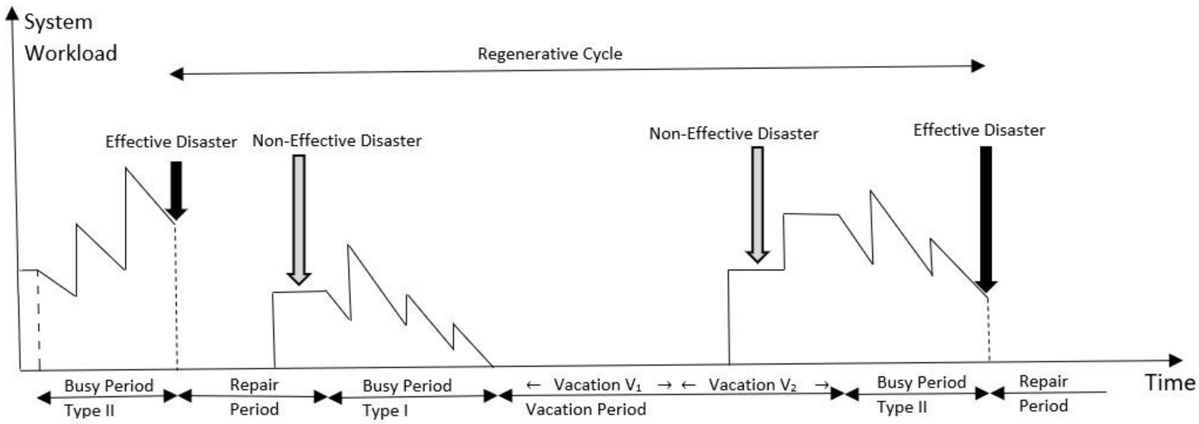

Figure 1: A sample path of the system is shown. The distribution of the time between two consecutive disasters (solid arrows) is obtained in Proposition 3.

This establishes the claim that a stationary version of the process exists, made in the beginning of section 3, on which the analysis there is based.

Intuitively, (70) is clear in view of the interpretation of γ as the relative frequency of vacations to repairs implied by (28). Each cycle consists of a single repair period, γ vacation periods on average, and a number of busy periods, the last of which ends by a disaster. Thus the total length of time in a cycle during which the server is busy serving customers is exponential with mean $\frac{1}{\delta}$. A renewal-reward argument then also shows that the fraction of time the server is busy is $P(0) = \frac{\delta^{-1}}{\delta^{-1}+m_R+m_U\gamma}$ which gives an intuitive explanation of (34).

Proof. 1. Analysis of the structure of an ordinary active-inactive sub-cycle.
Let τ_1 denote the length of a string of vacations until an arrival occurs. If $\{U_i\}$ is an i.i.d. sequence or random variables with the vacation distribution then the length of a vacation string is $\tau_1 = U_1 + U_2 + \cdots + U_M$ where $M = \min\{n : U_1 + \cdots + U_n > \Lambda\}$ and Λ is an exponential random variable with rate λ, independent of the vacation lengths, corresponding to the next arrival time. The probability that during a vacation period there are no arrivals is given by $\int_0^\infty e^{-\lambda t} dU(t) = \hat{U}(\lambda)$. The length of a vacation period, given that there are no arrivals in its duration, has Laplace transform given by $\frac{\hat{U}(s+\lambda)}{\hat{U}(\lambda)}$. Finally, the joint distribution of the length of a vacation, *given that there were arrivals in its duration,* and the number of customers present in the system when the vacation ends (and the busy period begins) is given by

$$\frac{1}{1-\hat{U}(\lambda)} \int_0^\infty \sum_{n=1}^{\infty} \chi(z)^n \frac{(\lambda t)^n}{n!} e^{-\lambda t} e^{-st} dU(t) \;=\; \frac{\hat{U}(s+\lambda-\lambda\chi(z)) - \hat{U}(\lambda+s)}{1-\hat{U}(\lambda)}. \tag{71}$$

The probability that a vacation string consists of n vacations, the first $n-1$ without arrivals, and the last containing arrivals, is $\hat{U}(\lambda)^{n-1}(1-\hat{U}(\lambda))$. Thus the joint distribution of the length, τ_1, of the string of vacations constituting the inactive period and the number of customers, ν_1, present in the system when it ends is

$$\begin{aligned}\mathbb{E}[e^{-s\tau_1}z^{\nu_1}] &= \sum_{n=1}^{\infty}(1-\hat{U}(\lambda))\hat{U}(\lambda)^{n-1}\left(\frac{\hat{U}(s+\lambda)}{\hat{U}(\lambda)}\right)^{n-1}\frac{\hat{U}(s+\lambda-\lambda\chi(z))-\hat{U}(\lambda+s)}{1-\hat{U}(\lambda)}\\ &= \frac{\hat{U}(s+\lambda-\lambda\chi(z))-\hat{U}(\lambda+s)}{1-\hat{U}(s+\lambda)}. \end{aligned} \tag{72}$$

Next we consider the second part of the sub-cycle, namely the busy period that ensues. To this end, *we imagine first that the disaster mechanism has been deactivated.* Denote by τ_2 the length of the busy period that follows. Its Laplace transform, conditional on the fact that it starts with ν_1 customers, is given by $\mathbb{E}[e^{-s_2\tau_2} \mid \nu_1] = \hat{\Gamma}_0(s_2)^{\nu_1}$ where $\hat{\Gamma}_0(s_2)$ is the Laplace transform of the length of a busy period of the system *without disasters* and *starting with a single customer*. (See section 4.2 for further discussion.) Thus we obtain the joint Laplace transform of the duration of an inactive period τ_1 and the busy period that follows it, τ_2, (without disasters) as

$$\mathbb{E}[e^{-s_1\tau_1-s_2\tau_2}] = \frac{\hat{U}(s_1+\eta(s_2))-\hat{U}(\lambda+s_1)}{1-\hat{U}(s_1+\lambda)}. \tag{73}$$

At this point we can take into account the possible occurrence of disasters. Let Δ be an exponential random variable with rate δ (the rate of occurrence of disasters), independent of all other random variables. Then the joint distribution of the inactive and the busy part of the sub-cycle, given that a disaster does not occur during the busy part, is

$$\begin{aligned}\mathbb{E}[e^{-s_1\tau_1-s_2\tau_2} \mid \Delta>\tau_2] &= \frac{\mathbb{E}[e^{-s_1\tau_1-s_2\tau_2}\mathbf{1}(\Delta>\tau_2)]}{\mathbb{P}(\Delta>\tau_2)} = \frac{\mathbb{E}[e^{-s_1\tau_1-(s_2+\delta)\tau_2}]}{\mathbb{E}[e^{-\delta\tau_2}]}\\ &= \frac{\hat{U}(s_1+\eta(s_2+\delta)))-\hat{U}(\lambda+s_1)}{1-\hat{U}(s_1+\lambda)}\,\frac{1-\hat{U}(\lambda)}{\hat{U}(\eta(\delta)))-\hat{U}(\lambda)}.\end{aligned}$$

Thence, the length of a sub-cycle starting with a vacation, given that a disaster does not occur during its busy period is

$$\psi_1(s) := \frac{\hat{U}(s+\eta(s+\delta)))-\hat{U}(\lambda+s)}{1-\hat{U}(s+\lambda)}\,\frac{1-\hat{U}(\lambda)}{\hat{U}(\eta(\delta)))-\hat{U}(\lambda)}. \tag{74}$$

The probability that such a sub-cycle is completed without a disaster happening is

$$a_1 := \mathbb{P}(\Delta>\tau_2) = \mathbb{E}[e^{-\delta\tau_2}] = \frac{1-\hat{U}(\eta(\delta))}{1-\hat{U}(\lambda)} \tag{75}$$

and the probability that a disaster occurs during such a sub-cycle is

$$1-a_1 = \frac{\hat{U}(\eta(\delta))-\hat{U}(\lambda)}{1-\hat{U}(\lambda)}. \tag{76}$$

Using (73) we can also obtain the joint distribution of the inactive and busy period given that the latter ends by a disaster. In that case the length of the busy period is equal to Δ and the corresponding joint Laplace transform is

$$\mathbb{E}[e^{-s_1\tau_1-s_2\Delta} \mid \Delta\le\tau_2] = \frac{1}{1-\mathbb{E}[e^{-\delta\tau_2}]}\,\frac{\delta}{\delta+s_2}\left(\mathbb{E}[e^{-s_1\tau_1}]-\mathbb{E}[e^{-s_1\tau_1-(\delta+s_2)\tau_2}]\right). \tag{77}$$

Therefore, the Laplace transform of the length of the sub-cycle, given that its busy period ends by a disaster, which we will denote by $\overline{\psi}_1(s)$, is obtained by setting $s_1=s_2=s$ in the above to obtain

$$\overline{\psi}_1(s) = \frac{\delta}{\delta+s}\,\frac{1-\hat{U}(\lambda)}{1-\hat{U}(\eta(\delta))}\,\frac{\hat{U}(s)-\hat{U}(s+\eta(s+\delta))}{1-\hat{U}(s+\lambda)}. \tag{78}$$

2. The initial active-inactive sub-cycle.

The inactive period here consist of a repair period, followed possibly by a string of vacations if there are no arrivals during the repair period. The length of the initial inactive period is $\tau_1^0 := R + U_1 + \cdots + U_M$ where $M = \inf\{n \geq 0 : R + U_1 + \cdots + U_n \leq \Delta\}$. We denote the number of customers present at the end of this inactive period by ν_1^0. There are two cases: If there are customer arrivals during the repair period, an event occurring with probability $1 - \hat{R}(\lambda)$, then the joint distribution of (τ_1^0, ν_1^0), conditional on this event, is

$$\frac{\hat{R}(s_1 + \lambda - \lambda\chi(z)) - \hat{R}(s_1 + \lambda)}{1 - \hat{R}(\lambda)}. \tag{79}$$

If there are no arrivals during the repair period, which happens with probability $\hat{R}(\lambda)$, then a string of vacations begins and, by the same arguments that lead to (72), the conditional joint distribution of (τ_1^0, ν_1^0) is

$$\begin{aligned}\frac{\hat{R}(s_1+\lambda)}{\hat{R}(\lambda)} \sum_{n=0}^{\infty} \left(\frac{\hat{U}(s_1+\lambda)}{\hat{U}(\lambda)}\right)^n \frac{\hat{U}(s_1+\lambda-\lambda\chi(z)) - \hat{U}(s_1+\lambda)}{1-\hat{U}(\lambda)} \left(1-\hat{U}(\lambda)\right)\hat{U}(\lambda)^n \\ = \frac{\hat{R}(s_1+\lambda)}{\hat{R}(\lambda)} \left(\hat{U}(s_1+\lambda-\lambda\chi(z)) - \hat{U}(s_1+\lambda)\right)\frac{1}{1-\hat{U}(s_1+\lambda)}. \end{aligned} \tag{80}$$

By multiplying (79) by $1 - \hat{R}(\lambda)$ and the right hand side of (80) by $\hat{R}(\lambda)$ and adding we obtain

$$\mathbb{E}[e^{-s_1\tau_1^0} z^{\nu_1^0}] = \hat{R}(s_1+\lambda-\lambda\chi(z)) - \hat{R}(s_1+\lambda)\frac{1-\hat{U}(s_1+\lambda-\lambda\chi(z))}{1-\hat{U}(s_1+\lambda)}.$$

Arguing as in (73), the joint Laplace transform of the duration of the initial inactive period τ_1^0 and the busy period that follows it, τ_2^0, (without disasters) is

$$\mathbb{E}[e^{-s_1\tau_1^0 - s_2\tau_2^0}] = \hat{R}(s_1+\eta(s_2)) - \hat{R}(s_1+\lambda)\frac{1-\hat{U}(s_1+\eta(s_2))}{1-\hat{U}(s_1+\lambda)}$$

and the Laplace transform of the initial inactive-active sub-cycle, given that its busy period does not suffer a disaster, $\psi_0(s)$, is given by

$$\psi_0(s) = \frac{\mathbb{E}[e^{-s_1\tau_1^0-(s_2+\delta)\tau_2^0}]}{\mathbb{E}[e^{-\delta\tau_2^0}]} = \frac{\hat{R}(s+\eta(s+\delta)) - \hat{R}(s+\lambda)\frac{1-\hat{U}(s+\eta(s+\delta))}{1-\hat{U}(s+\lambda)}}{\hat{R}(\eta(\delta)) - \hat{R}(\lambda)\frac{1-\hat{U}(\eta(\delta))}{1-\hat{U}(\lambda)}}. \tag{81}$$

The probability that the initial sub-cycle is completed without a disaster happening is

$$a_0 := \hat{R}(\eta(\delta)) - \hat{R}(\lambda)\frac{1-\hat{U}(\eta(\delta))}{1-\hat{U}(\lambda)}. \tag{82}$$

The complementary probability, $1 - a_0$, is of course the probability that during the busy period of the initial sub-cycle a disaster occurs. Again, using (77) in this case, the Laplace transform of the length of the initial sub-cycle, given that a disaster occurs, $\overline{\psi}_0(s)$, can be obtained as follows:

$$\begin{aligned}&\mathbb{E}[e^{-s_1\tau_1^0 - s_2\Delta} \mid \Delta \leq \tau_2^0] \\ &= \frac{\delta}{\delta+s_2}\frac{1}{1-a_0}\left(\hat{R}(s_1) - \hat{R}(s_1+\eta(s_2+\delta)) - \hat{R}(s_1+\lambda)\frac{\hat{U}(s_1+\eta(s_2+\delta)) - \hat{U}(s_1)}{1-\hat{U}(s_1+\lambda)}\right)\end{aligned} \tag{83}$$

Setting $s_1 = s_2 = s$ in (83) we obtain

$$\overline{\psi}_0(s) \;=\; \frac{\delta}{\delta+s}\frac{1}{1-a_0}\left(\hat{R}(s) - \hat{R}(s+\eta(s+\delta)) - \hat{R}(s+\lambda)\frac{\hat{U}(s+\eta(s+\delta)) - \hat{U}(s)}{1-\hat{U}(s+\lambda)}\right) \quad (84)$$

The Laplace Transform of the time between two consecutive disasters is

$$\begin{aligned}\hat{\Omega}(s) \;&:=\; (1-a_0)\overline{\psi}_0(s) + a_0(1-a_1)\psi_0(s)\overline{\psi}_1(s)\sum_{n=0}^{\infty} a_1^n \psi_1(s)^n \\ &=\; (1-a_0)\overline{\psi}_0(s) + a_0\psi_0(s)\,\frac{(1-a_1)\overline{\psi}_1(s)}{1-a_1\psi_1(s)}.\end{aligned}$$

Substituting in the above (75), (74), (78), (82), (81), (84), we obtain (68).

Differentiating (68) and evaluating at $s = 0$ we obtain $\hat{\Omega}'(0) = \frac{1}{\delta} - \hat{R}'(0) - \hat{U}'(0)\,\frac{\hat{R}(\eta(\delta))}{1-\hat{U}(\eta(\delta))}$. Since $\eta(\delta) = \alpha(\hat{\Gamma}_0(\delta))$ and $\Gamma_0(\delta) = z_0$, $\hat{\Omega}'(0) = \frac{1}{\delta} + m_R + m_U\,\frac{\hat{R}(\alpha(z_0))}{1-\hat{U}(\alpha(z_0))}$. This, together with the definition (27), establishes (70). □

Corollary 4. *When the repair time is negligible it suffices to take $R = U$ in order to obtain the corresponding Laplace transform which is*

$$\hat{\Omega}(s) \;=\; \frac{\delta}{s+\delta}\,\frac{\hat{U}(s) - \hat{U}(s+\eta(s+\delta))}{1-\hat{U}(s+\eta(s+\delta))}.$$

In particular, when the duration of each vacation is exponentially distributed, i.e. $\hat{U}(s) = \frac{\omega}{\omega+s}$, the Laplace transform becomes

$$\hat{\Omega}(s) \;=\; \frac{\delta}{\delta+s}\frac{\omega}{\omega+s}\;\frac{\lambda - \lambda\chi(\hat{\Gamma}_0(s+\delta))}{s+\lambda-\lambda\chi(\hat{\Gamma}_0(s+\delta))}.$$

7 Optimizing the system performance with respect to the service rate under a Quality of Service constraint

In this section we show how the results obtained can be used to provide decision support in designing and/or operating this system. We discuss the problem of choosing the optimal service rate while maintaining a minimum level of service defined as the percentage of customers who complete service and depart without being affected by a disaster. Suppose that the cost/reward structure in the system we discussed is determined by the following factors:

1. Each time a disaster occurs, a repair cost of C_R is incurred. (Recall that the rate of disaster occurrences is $P(1)\delta$.)
2. For each customer removed from the system as a result of a disaster a cost C_D is incurred. (The rate of customer removals due to disasters is $\lambda m_\chi - \delta P(1)\frac{(1-r)\hat{S}(\delta)}{1-\hat{S}(\delta)}$.)
3. A reward with rate C_V per unit time is realized when the server is on vacation since the server is then performing some additional useful task. (The proportion of time the server is on vacation is $\gamma m_V P(1)$.)

4. Finally we assume that a server with server rate μ costs $\phi(\mu)$ per time unit (where ϕ is a given smooth, increasing function)

We do not include a reward term for customers served since it would only add a fixed term, given that we have already included a penalty for customers who do not complete their service. We could however add an additional cost related to the average number of customers in the system which however we do not do here. Problems involving the average waiting time for customers that complete service will be considered in a future paper.

The optimization problem we examine is to determine the optimal value of the service rate so as to minimize the overall cost rate under the Quality of Service (QoS) constraint guaranteeing a minimum for the fraction of customers who complete service (without suffering from a disaster).

For the sake of concreteness and (relative) simplicity we will make the assumption that the service time distribution is exponential with $\hat{S}(s) = \frac{\mu}{\mu+s}$. In our notation the probability that a customer leaves the system after receiving service (as opposed to be eliminated by a disaster) is

$$p = \frac{\delta P(1)}{\lambda m_\chi} \frac{(1-r)\hat{S}(\delta)}{1-\hat{S}(\delta)} = P(1)\mu\frac{1-r}{\lambda m_\chi}. \tag{85}$$

We thus have the minimization problem

$$\begin{aligned}&\min\left\{\phi(\mu) + C_R P(1)\delta + C_D\left(\lambda m_\chi - (1-r)\mu P(1)\right) - C_V \gamma m_V \delta P(1)\right\}\\ &\text{subject to } p \geq p_0.\end{aligned}$$

In the above problem the decision parameter is the service rate μ while the inequality constraint $p \geq p_0$, is a quality of service requirement ensuring there is a floor, p_0, on the number of customers who complete service successfully.

The inequality constraint can be rewritten as $P(1)\mu\frac{1-r}{\lambda m_\chi} \geq p_0$ or $P(1)\mu \geq p_0\frac{\lambda m_\chi}{1-r} =: c_0$. The minimization problem can thus be restated as

$$\begin{aligned}&\min\left\{\phi(\mu) + C_D\lambda m_\chi + P(1)\left[C_R\delta - C_D(1-r)\mu - C_V\gamma m_V\delta\right]\right\}\\ &\text{subject to } P(1)\mu \geq c_0.\end{aligned} \tag{86}$$

In the above expression, $P(1)$ depends on μ through γ. In fact we have

$$P(1) = \frac{1}{1+\delta m_R + \delta m_V\gamma} \quad \text{where} \quad \gamma = \frac{\hat{R}(\alpha(z_0))}{1-\hat{U}(\alpha(z_0))} \tag{87}$$

In turn, when $\hat{S}(s) = \frac{\mu}{\mu+s}$, the dependence of $z_0(\mu)$ on μ is given implicitly by

$$z_0 = \frac{\mu(1-r+rz_0)}{\mu+\delta+\lambda-\lambda\chi(z_0)}. \tag{88}$$

The above equation is equivalent to $\alpha(z_0) = \mu(1-r) + \delta - \mu(1-r)z_0^{-1}$. Assuming for instance that $\hat{R}(s) = (1+sm_R)^{-1}$ and $\hat{U}(s) = (1+sm_V)^{-1}$, and setting $\bar{\mu} := (1-r)\mu$,

$$\gamma(\mu) = \frac{1}{1+m_R\delta + m_R\bar{\mu}(1-z_0)^{-1}}\left(1 + \frac{1}{m_V\delta + m_V\bar{\mu}(1-z_0^{-1})}\right).$$

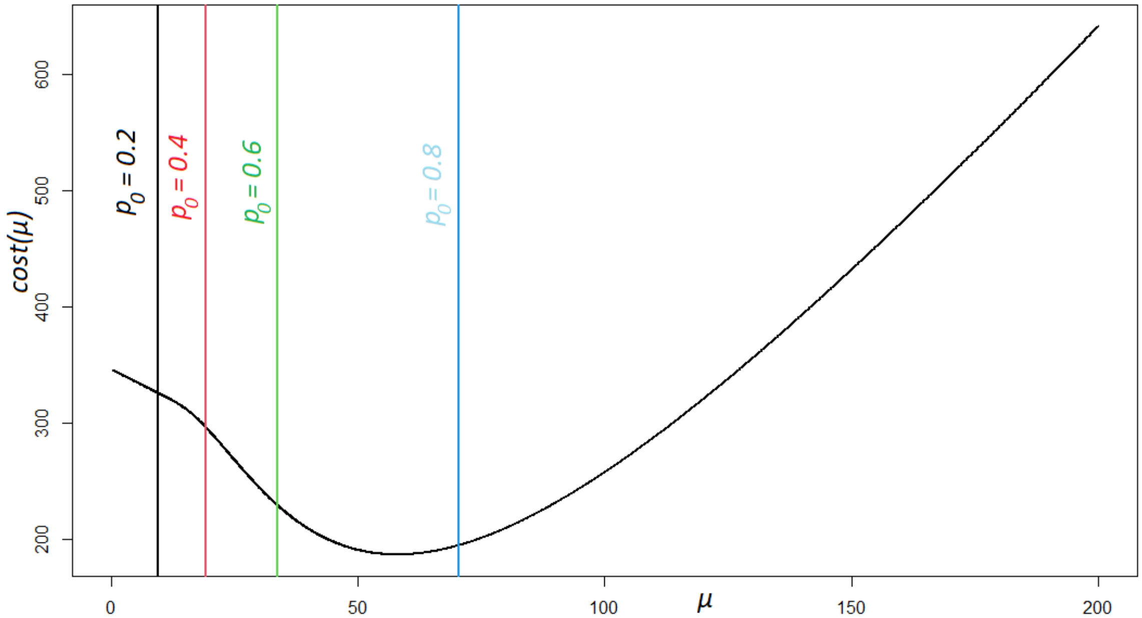

Figure 2: Minimizing the cost function. The cost function is plotted for $\lambda = 1$, $\delta = 0.02$, $r = 0.5$, $m_V = 200$, $m_R = 300$, and cost rates $C_V = 500$, $C_R = 1000$, and $C_D = 100$. Also, the batch size is geometric with $\chi(z) = \frac{0.3z}{1-0.7z}$. If for instance the QoS constraint is $p_0 = 0.6$ then we need to find the minimum cost to the right of the green line.

and therefore, the quantity to be minimized is

$$\begin{aligned} & \phi(\mu) + C_D\lambda m_\chi + P(1)\left[C_R\delta - C_D(1-r)\mu - C_V\gamma m_V\delta\right] \qquad (89) \\ = \; & \phi(\mu) + C_D\lambda m_\chi + \frac{\left[C_R\delta - C_D(1-r)\mu\right]\left[1 - \hat{U}(\alpha(z_0))\right] - \delta m_V C_V \hat{R}(\alpha(z_0))}{\left[1+\delta m_R\right]\left[1 - \hat{U}(\alpha(z_0))\right] + \delta m_V \hat{R}(\alpha(z_0))} \\ = \; & \phi(\mu) + C_D\lambda m_\chi \; + \; \frac{\left[C_R\delta - C_D\bar{\mu}\right]\alpha(z_0)\left[1 + m_R\alpha(z_0)\right] - \delta\left[1 + m_V\alpha(z_0)\right]}{\left[1+\delta m_R\right]\alpha(z_0)\left[1 + m_R\alpha(z_0)\right] + \delta\left[1 + m_V\alpha(z_0)\right]}. \end{aligned}$$

In general one would have to resort to non-linear programming techniques. However, in the problem formulated here the decision space is one dimensional and hence the solution may be obtained by plotting the cost and the QoS constraint vs. the decision parameter μ as shown Figure 2.

8 Unconditional Disasters: A Model with Disasters Affecting the System Regardless of Its State

In this section we consider a variation of the model in which disasters can occur at any time, regardless of the state of the server. This includes periods when the server is on vacation or under repair, following a disaster. In all cases, when a disaster occurs, a repair period begins causing the removal of all waiting customers. The repair period may itself be interrupted by a disaster as well. Such a model, which we will term the system with *unconditional disasters* may be appropriate when the disasters represent external factors such as power failures or attacks that do not require the server to be operational in order to be effective. Another example is given by a computer server which may be susceptible to DDoS attacks or to infection by viruses whether

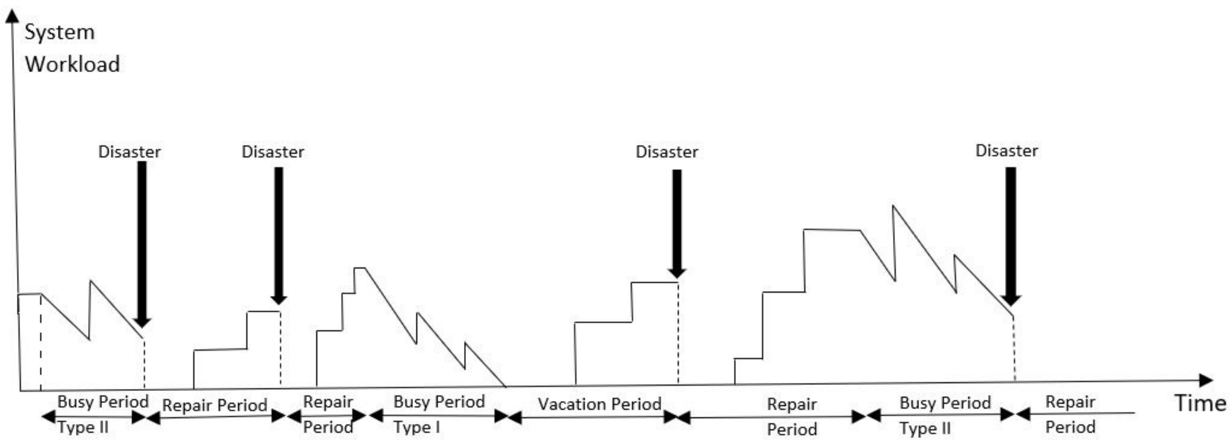

Figure 3: Sample path of a system with unconditional disasters

it performs the primary tasks (server busy) or it performs secondary tasks (on vacation). As far as we know this variation, while modeling realistic situations, has not been studied before. A model of a machine with set ups and breakdowns which may occur not only in the course of machine operation but also during the set up time has been studied in Bu and Liu [4].

8.1 Supplementary Variables Analysis

The balance equations for the unconditional disaster model are given by

$$\frac{d}{dx}P_n(x) + (\lambda + \delta + \mu(x))\,P_n(x) = \lambda \sum_{k=1}^{n-1} \chi_k P_{n-k}(x), \quad x > 0,\ n \geq 1 \tag{90}$$

$$\frac{d}{dx}W_0(x) + (\lambda + \delta + r(x))W_0(x) = 0 \tag{91}$$

$$\frac{d}{dx}W_n(x) + (\lambda + \delta + r(x))W_n(x) = \lambda \sum_{k=1}^{n} \chi_k W_{n-k}(x), \quad x > 0, n \geq 1 \tag{92}$$

$$\frac{d}{dx}V_0(x) + (\lambda + \delta + u(x))V_0(x) = 0, \tag{93}$$

$$\frac{d}{dx}V_n(x) + (\lambda + \delta + u(x))V_n(x) = \lambda \sum_{k=1}^{n} \chi_k V_{n-k}(x), \quad x > 0,\ n \geq 1. \tag{94}$$

Note that (90) is the same as (2) while (91)–(94) are modified, compared to (3)–(6) to include the effect of disasters during the vacation and repair periods. The boundary conditions of the above system of differential equations are given by (7), (8), (9), together with

$$W_0(0) \quad = \quad \delta, \tag{95}$$

which replaces (10). The normalization condition is again (11). Using the same methodology and arguments as in section 3 we obtain

$$P(0;z) = z\,\frac{\delta\hat{R}(\alpha(z)+\delta) - V_0(0)\left(1-\hat{U}(\alpha(z)+\delta)\right)}{z-(1-r+rz)\,\hat{S}(\delta+\alpha(z))}. \tag{96}$$

Equation (18) is modified here in what regards $V(x;z)$ and $W(x;z)$ as follows:

$$W(x;z) = W_0(0)[1-R(x)]e^{-(\delta+\alpha(z))x}, \qquad V(x;z) = V_0(0)[1-V(x)]e^{-(\delta+\alpha(z))x}. \tag{97}$$

The value of $V_0(0)$ is again determined using Rouche's theorem in (96) yielding

$$V_0(0) \;=\; \delta\gamma_\delta \tag{98}$$

where

$$\gamma_\delta := \frac{\hat{R}(\alpha(z_0)+\delta)}{1-\hat{U}(\alpha(z_0)+\delta)}. \tag{99}$$

Since δ is the rate of disasters (and therefore the repair initiation rate $W_0(0)$), and $V_0(0)$ is the rate of vacation initiations, it follows that $\gamma_\delta = V_0(0)\delta^{-1}$ is the frequency of vacation periods compared to repair periods. z_0 is again the unique root of the denominator of (96) within the unit disk. (See Proposition 5 on the Appendix.) Setting now

$$K_\delta(z) := \hat{R}(\delta+\alpha(z)) - \gamma_\delta\left(1-\hat{U}(\delta+\alpha(z))\right) \tag{100}$$

we rewrite (96) as

$$P(0;z) \;=\; \frac{\delta z\,K_\delta(z)}{z-(1-r+rz)\,\hat{S}(\delta+\alpha(z))}.$$

8.2 Probability Generating Functions and Performance Measures

U1. *The partial pgf of the number of customers in the system when the server is working.*

$$P(z) \;=\; \frac{\delta z\,K_\delta(z)}{z-(1-r+rz)\,\hat{S}(\delta+\alpha(z))}\,\frac{1-\hat{S}(\delta+\alpha(z))}{\delta+\alpha(z)} \tag{101}$$

with

$$P(1) = \hat{R}(\delta) - \gamma_\delta\left(1-\hat{U}(\delta)\right) \tag{102}$$

giving the steady-state probability that the server is busy serving customers.

U2. *The partial pgf of the number of customers in the system when the server is under repair.*

$$W(z) \;=\; \delta\,\frac{1-\hat{R}(\alpha(z)+\delta)}{\alpha(z)+\delta}. \tag{103}$$

The probability that the server is under repair is $W(1) = 1-\hat{R}(\delta)$.

U3. *The partial pgf of the number of customers in the system when the server is on vacation.*

$$V(z) = \delta\gamma_\delta\,\frac{1-\hat{U}(\alpha(z)+\delta)}{\alpha(z)+\delta}. \tag{104}$$

The probability that the server is on vacation is $V(1) = \gamma_\delta(1 - \hat{U}(\delta))$.

U4. *The pgf of the number of customers in the system in stationarity.*

$$\Phi(z) \;=\; \frac{zK_\delta(z)[1-\hat{S}(\delta+\alpha(z))]}{z-(1-r+rz)\,\hat{S}(\delta+\alpha(z))}\,\frac{\delta}{\delta+\alpha(z)} + \delta m_R \hat{R}_e(\alpha(z)+\delta) + \delta\gamma_\delta m_U \hat{U}_e(\alpha(z)+\delta).$$

The mean number of customers in stationarity is

$$\Phi'(1) \;=\; \frac{\lambda m_\chi}{\delta} - \left(\hat{R}(\delta) - \gamma_\delta\left(1-\hat{U}(\delta)\right)\right)\frac{(1-r)\hat{S}(\delta)}{1-\hat{S}(\delta)}. \tag{105}$$

U5. The pgf of the system size at a departure epoch.

$$\Phi^+(z) = D_0^{-1}(1-r)\delta\,\frac{K_\delta(z)}{z-(1-r+rz)\,\hat{S}(\delta+\alpha(z))} \tag{106}$$

where $D_0 = \frac{\delta(1-r)\left(\hat{R}(\delta)-\gamma_\delta(1-\hat{U}(\delta))\right)}{1-\hat{S}(\delta)}$.

U6. *The pgf of the system size at a busy period initiation epoch.* Using the ratio of rates argument that led to (46) we find in the same way that

$$\Psi(z) \;=\; \frac{\hat{R}(\alpha(z)+\delta) - \hat{R}(\lambda+\delta) + \gamma_\delta\left(\hat{U}(\alpha(z)+\delta) - \hat{U}(\lambda+\delta)\right)}{\hat{R}(\delta) - \hat{R}(\lambda+\delta) + \gamma_\delta\left(\hat{U}(\delta) - \hat{U}(\lambda+\delta)\right)} \;=\; \frac{K_\delta(z)-K_\delta(0)}{1-K_\delta(0)} \tag{107}$$

U7. *Rate of customer departures after completing service, f.* This is given by

$$f \;=\; (1-r)\int_0^\infty \sum_{n=1}^\infty P_n(x)\mu(x)dx. \tag{108}$$

Adding equations (7), (8) term by term, for all $n \geq 1$, and using (108) we obtain

$$\sum_{n=1}^\infty P_n(0) \;+\; V_0(0) \;=\; (1-r)f + rf + \int_0^\infty \sum_{n=0}^\infty V_n(x)u(x)dx + \int_0^\infty \sum_{n=0}^\infty W_n(x)r(x)dx \tag{109}$$

$$\sum_{n=1}^\infty P_n(0) = P(0;1) = \frac{\delta K_\delta(1)}{1-\hat{S}(\delta)} = \frac{\delta}{1-\hat{S}(\delta)}\left(\hat{R}(\delta) - \gamma_\delta(1-\hat{U}(\delta))\right).$$

Also

$$\sum_{n=0}^\infty W_n(x) = W(x;1) = W_0(0)[1-R(x)]e^{-\delta x}$$

and hence

$$\int_0^\infty \sum_{n=0}^\infty W_n(x)r(x)dx = W_0(0)\hat{R}(\delta).$$

Similarly $\int_0^\infty \sum_{n=0}^\infty V_n(x)u(x)dx = V_0(0)\hat{U}(\delta)$. Thus, from (109),

$$f = (1-r)P(0;1) + V_0(0)[1-\hat{U}(\delta)] - W_0(0)\hat{R}(\delta),$$

and hence,

$$f = (1-r)\frac{\delta\hat{S}(\delta)}{1-\hat{S}(\delta)}\left[\hat{R}(\delta) - \gamma_\delta(1-\hat{U}(\delta))\right]. \tag{110}$$

U8. *Fraction of customers that complete service without suffering a disaster, p.* The rate of departures from the system after a service completion is $(1-r)f$ where f is given by (110). The rate of arrivals to the system on the other hand is λm_χ. Hence, by a ratio of rates argument, the probability that a customer completes service and departs from the system without suffering a disaster is

$$p \;=\; \frac{\delta}{\lambda m_\chi} \frac{(1-r)\hat{S}(\delta)}{1-\hat{S}(\delta)} \left[\hat{R}(\delta) - \gamma_\delta(1-\hat{U}(\delta))\right]. \tag{111}$$

It is worth noting that the above expression, together with $\Phi'(1)$ gives the following interesting formula for the average number of customers in the system:

$$\Phi'(1) \;=\; \frac{\lambda m_\chi}{\delta}(1-p). \tag{112}$$

This expression has a simple explanation: Written as $\delta\Phi'(1) \;=\; \lambda m_\chi(1-p)$, the left hand side is the rate at which customers are removed due to disasters, as a result of PASTA. The right rand side is the rate at which customers arrive multiplied by the probability that a customer will suffer a disaster. These two rates must of course be equal.

U9. *Rate of ordinary busy period terminations, θ_1:*

$$\theta_1 = (1-r)\int_0^\infty P_1(x)\mu(x)dx$$

From (8) we have

$$V_0(0) = \theta_1 + \int_0^\infty V_0(x)u(x)dx + \int_0^\infty W_0(x)r(x)dx$$

hence

$$V_0(0) = \theta_1 + \int_0^\infty V_0(0)e^{-(\lambda+\delta)x}[1-U(x)]u(x)dx + \int_0^\infty W_0(0)e^{-(\lambda+\delta)x}[1-R(x)]r(x)dx$$

or

$$\theta_1 = \delta\left(\gamma_\delta[1-\hat{U}(\lambda+\delta)] - \hat{R}(\lambda+\delta)\right).$$

U10. *Rate of busy period terminations by disasters, θ_2:*

$$\theta_2 = \delta P(1) = \delta\frac{\delta K_\delta(1)}{1-\hat{S}(\delta)}\frac{1-\hat{S}(\delta)}{\delta} = \delta K_\delta(1) = \delta\left(\hat{R}(\delta) - \gamma_\delta[1-\hat{U}(\delta)]\right).$$

T11. *Probability that a busy period terminates by a departure, q_1 or by a disaster, q_2.* From a ratio of rates argument and the above results we have

$$q_1 \;=\; \frac{\theta_1}{\theta_1+\theta_2} \;=\; \frac{\gamma_\delta\left(1-\hat{U}(\delta+\lambda)\right) - \hat{R}(\delta+\lambda)}{\hat{R}(\delta) - \hat{R}(\lambda+\delta) + \gamma_\delta\left(\hat{U}(\delta) - \hat{U}(\lambda+\delta)\right)} \;=\; \frac{-K_\delta(0)}{K_\delta(1)-K_\delta(0)}, \tag{113}$$

$$q_2 \;=\; \frac{\theta_2}{\theta_1+\theta_2} \;=\; \frac{\hat{R}(\delta) - \gamma_\delta\left(1-\hat{U}(\delta)\right)}{\hat{R}(\delta) - \hat{R}(\lambda+\delta) + \gamma_\delta\left(\hat{U}(\delta) - \hat{U}(\lambda+\delta)\right)} \;=\; \frac{K_\delta(1)}{K_\delta(1)-K_\delta(0)}. \tag{114}$$

U12. According to the theorem for busy period when we have partial disasters the busy period in total disaster case given by

$$\hat{B}(s) = 1 - \frac{1}{s+\delta}\frac{(\delta K_\delta(1) - s - \delta)K_\delta(\hat{\Gamma}_0(s+\delta)) + sK_\delta(1)}{K_\delta(1)-K_\delta(0)}. \tag{115}$$

U13. *Rate of initiation of vacations.* This is given by $V_0(0) = \gamma_\delta W_0(0) = \delta\gamma_\delta$.

U14. *Probability that a disaster occurs during a busy period.* Due to PASTA this is simply $P(1)$, the probability that the server is busy.

U15. *Probability that a disaster removes no customer*. Again, due to PASTA, this probability is given by

$$W(0) + V(0) = \frac{\delta}{\delta+\lambda}\left(1 - \hat{R}(\lambda+\delta) + \gamma_\delta\left(1 - \hat{U}(\lambda+\delta)\right)\right).$$

8.3 Laplace transform of an inactive period.

We define as *inactive period* the time that elapses from the moment the server ceases serving customers (either because of a service completion leaving the system empty or because of a disaster) to the moment the next busy period begins (and thus the server begins serving customers again). We distinguish between two types of inactive periods, those that begin with a disaster, and those that begin with a vacation. If we denote by $\hat{I}(s)$ the Laplace transform of the duration of a typical inactive period, by $\hat{I}_1(s)$ that of an inactive period *that begins with a vacation*, and by $\hat{I}_2(s)$ that of an inactive period *that begins with a disaster*, then we have

$$\hat{I}(s) = q_1\hat{I}_1(s) + q_2\hat{I}_2(s) \tag{116}$$

where the probabilities q_1, q_2, are given by (113), (114). We evaluate first $\hat{I}_1(s)$, taking into account the fact that, in the model considered in this section, disasters may occur during vacation or repair periods. Let A_n, $n = 1, 2, \ldots$, denote the event that such an inactive period consists of a string of $n-1$ vacations during which neither a disaster nor an arrival occurs, ending by a vacation during which either a disaster occurs. (Note that in this last vacation a number of arrivals may occur before the disaster.) Suppose that $\{U_i\}$, $\{\Lambda_i\}$, $\{\Delta_i\}$, are three independent sequences of i.i.d. random variables. The elements of the first sequence are distributed according to the vacation distribution, while those of the second and the third are exponentially distributed with rates λ and δ respectively. Then, taking into account the memoryless property of the exponential distribution,

$$\mathbb{P}(A_n) = \mathbb{P}\left\{U_1 < (\Lambda_1 \wedge \Delta_1),\, \ldots,\, U_{n-1} < (\Lambda_{n-1} \wedge \Delta_{n-1}),\, \Delta_n < U_n\right\}.$$

Similarly, let B_n be the event that we have $n-1$ consecutive vacations during which neither a vacation nor an arrival occurs, followed by a vacation during at least one arrival and no disaster occur. Then

$$\mathbb{P}(B_n) = \mathbb{P}\left\{U_1 < \Lambda_1 \wedge \Delta_1,\, \ldots\, , U_{n-1} < \Lambda_{n-1} \wedge \Delta_{n-1}\, ,\, \Lambda_n < U_n < \Delta_n\right\}.$$

Clearly the events $\{A_n, B_n; n = 1, 2, \ldots\}$ are disjoint and their union is the whole sample space so that

$$\mathbb{E}[e^{-sI_1}] = \sum_{n=1}^{\infty} \mathbb{E}[e^{-sI_1}\mathbf{1}(A_n)] + \mathbb{E}[e^{-sI_1}\mathbf{1}(B_n)] \tag{117}$$

On A_n it holds that $I_1 = U_1 + \cdots + U_{n-1} + \Delta_n + \tilde{I}_2$ since, during the nth vacation a disaster occurs and after this an inactive period starting with a disaster begins. Here $\tilde{I}_2$ is an independent random variable with Laplace transform $\hat{I}_2(s)$. Thus

$$\begin{aligned}\mathbb{E}[e^{-sI_1}\mathbf{1}(A_n)] \;&=\; \mathbb{E}[e^{-s(U_1+\cdots+U_{n-1})}\mathbf{1}\left(U_1 < \Lambda_1 \wedge \Delta_1, \ldots, U_{n-1} < \Lambda_{n-1} \wedge \Delta_{n-1}\right)] \\ &\qquad \times\mathbb{E}[e^{-s\Delta_n}\mathbf{1}(\Delta_n < U_n)]\,\hat{I}_2(s).\end{aligned} \tag{118}$$

On the other hand, on B_n it holds that $I_1 = U_1 + \cdots + U_{n-1} + U_n$ and therefore

$$\begin{aligned}\mathbb{E}[e^{-sI_1}\mathbf{1}(B_n)] &= \mathbb{E}[e^{-s(U_1+\cdots+U_{n-1})}\mathbf{1}\left(U_1 < \Lambda_1 \wedge \Delta_1, \ldots, U_{n-1} < \Lambda_{n-1} \wedge \Delta_{n-1}\right)] \\ &\qquad \times \left\{\mathbb{E}[e^{-sU_n}\mathbf{1}(\Lambda_n < U_n < \Delta_n)]\right\}. \end{aligned} \tag{119}$$

Given that

$$\begin{aligned}\mathbb{E}[e^{-s\Delta_n}\mathbf{1}(\Delta_n < U_n)] &= \mathbb{E}\left[\int_0^{U_n} e^{-sx}\delta e^{-\delta x}dx\right] = \frac{\delta}{\delta+s}\left(1 - \hat{U}(s+\delta)\right), \\ \mathbb{E}[e^{-sU_n}\mathbf{1}(\Lambda_n < U_n < \Delta_n)] &= \hat{U}(s+\delta) - \hat{U}(s+\lambda+\delta), \end{aligned} \tag{120}$$

and

$$\begin{aligned}\mathbb{E}[e^{-s(U_1+\cdots+U_{n-1})}\mathbf{1}\left(U_1 < \Lambda_1 \wedge \Delta_1, \ldots, U_{n-1} < \Lambda_{n-1} \wedge \Delta_{n-1}\right)] &= \prod_{i=1}^{n-1} \mathbb{E}[e^{-sU_i}\mathbf{1}\left(U_i < \Lambda_i \wedge \Delta_i\right)] \\ &= \hat{U}(s+\lambda+\delta)^{n-1} \end{aligned} \tag{121}$$

we have

$$\hat{I}_1(s) = \sum_{n=1}^{\infty} \hat{U}(s+\lambda+\delta)^{n-1}\left[\frac{\delta}{\delta+s}[1-\hat{U}(s+\delta)]\hat{I}_2(s) + \hat{U}(s+\delta) - \hat{U}(s+\lambda+\delta)\right]$$

whence we obtain

$$\hat{I}_1(s) = \frac{\frac{\delta}{\delta+s}[1-\hat{U}(s+\delta)]\hat{I}_2(s) + \hat{U}(s+\delta) - \hat{U}(s+\delta+\lambda)}{1-\hat{U}(s+\delta+\lambda)}. \tag{122}$$

The corresponding computation for the Laplace transform of an inactive period starting with a disaster can be obtained by a similar argument, which we will not present at the same level of detail. We will consider again a partition of the sample space according to the number, $n = 1, 2, \ldots$ of consecutive repair times at the beginning of this inactive period. Of these, the first $n-1$ are all interrupted by disasters, whereas the n is completed without the occurrence of a disaster. If during this nth repair time customer arrivals occur then the end of the repair period is also the end of the inactive period, otherwise at the end of the repair period there follows an independent inactive period starting with a vacation with Laplace transform given by (122). With $\{R_i\}$ being an i.i.d. sequence of random variables with the distribution of the repair period we have the counterpart of (117), (118),

$$\hat{I}_2(s) = \sum_{n=1}^{\infty}\left(\begin{array}{l}\mathbb{E}[e^{-s(\Delta_1+\cdots+\Delta_{n-1})}\mathbf{1}(R_1 > \Delta_1, \ldots, R_{n-1} > \Delta_{n-1})] \\ \qquad \times\left\{\mathbb{E}[e^{-sR_n}\mathbf{1}(R_n < \Delta_n, R_n < \Lambda)]\,\hat{I}_1(s) + \mathbb{E}[e^{-sR_n}\mathbf{1}(\Lambda < R_n < \Delta)]\right\}\end{array}\right).$$

Evaluating the above quantities we obtain

$$\hat{I}_2(s) = \sum_{n=1}^{\infty}\left(\frac{\delta}{\delta+s}[1-\hat{R}(s+\delta)]\right)^{n-1}\left[\hat{R}(s+\lambda+\delta)\hat{I}_1(s) + \hat{R}(s+\delta) - \hat{R}(s+\lambda+\delta)\right]$$

whence we obtain

$$\hat{I}_2(s) = \frac{\hat{R}(s+\lambda+\delta)\hat{I}_1(s) + \hat{R}(s+\delta) - \hat{R}(s+\lambda+\delta)}{1-\frac{\delta}{\delta+s}[1-\hat{R}(s+\delta)]}. \tag{123}$$

From (122) and (123) we obtain

$$\hat{I}_1(s) = \frac{\left[s+\delta\hat{R}(s+\delta)\right]\left[1-\hat{U}(s+\lambda+\delta)\right]-\left[s+\delta\hat{R}(s+\lambda+\delta)\right]\left[1-\hat{U}(s+\delta)\right]}{\left[s+\delta\hat{R}(s+\delta)\right]\left[1-\hat{U}(s+\lambda+\delta)\right]-\delta\hat{R}(s+\lambda+\delta)\left[1-\hat{U}(s+\delta)\right]},$$

$$\hat{I}_2(s) = \frac{(s+\delta)\hat{R}(s+\delta)\left[1-\hat{U}(s+\lambda+\delta)\right]-(s+\delta)\hat{R}(s+\lambda+\delta)\left[1-\hat{U}(s+\delta)\right]}{\left[s+\delta\hat{R}(s+\delta)\right]\left[1-\hat{U}(s+\lambda+\delta)\right]-\delta\hat{R}(s+\lambda+\delta)\left[1-\hat{U}(s+\delta)\right]}.$$

From the above, together with (116) and (113), (114), we obtain the Laplace transform of the typical inactive period

$$\hat{I}(s) = \frac{\left[(\delta+q_2s)\hat{R}(s+\delta)+q_1s\right]\left[1-\hat{U}(s+\lambda+\delta)\right]-\left[(\delta+q_2s)\hat{R}(s+\lambda+\delta)+q_1s\right]\left[1-\hat{U}(s+\delta)\right]}{\left[s+\delta\hat{R}(s+\delta)\right]\left[1-\hat{U}(s+\lambda+\delta)\right]-\delta\hat{R}(s+\lambda+\delta)\left[1-\hat{U}(s+\delta)\right]}.$$

The corresponding mean values are given by

$$\mathbb{E}I_1 = \frac{1}{\delta}\frac{1-\hat{U}(\delta)}{\hat{R}(\delta)[1-\hat{U}(\lambda+\delta)]-\hat{R}(\lambda+\delta)[1-\hat{U}(\delta)]}$$

$$\mathbb{E}I_2 = \frac{1}{\delta}\frac{[1-\hat{R}(\delta)][1-\hat{U}(\lambda+\delta)]+\hat{R}(\lambda+\delta)[1-\hat{U}(\delta)]}{\hat{R}(\delta)[1-\hat{U}(\lambda+\delta)]-\hat{R}(\lambda+\delta)[1-\hat{U}(\delta)]}$$

$$\mathbb{E}I = \frac{1}{\delta}\frac{1-\hat{R}(\delta)+\gamma_\delta[1-\hat{U}(\delta)]}{\hat{R}(\delta)-\hat{R}(\lambda+\delta)+\gamma_\delta(\hat{U}(\delta)-\hat{U}(\lambda+\delta))} \qquad (124)$$

9 Numerical Results

Here we present numerical results both for the basic model of sections 2 through 6, which we will call here the *unconditional disasters* model, and for the conditional disasters model analyzed in section 8. Numerical results are plotted for two performance criteria, the average number of customers in the system (37) and (105) and the probability of a customer completing service and departing without suffering a disaster (67) and (111). In all cases presented, the Poisson arrival rate is $\lambda = 1$, the batch size is constant and equal to 1, and the feedback probability is $r = 0.5$. Regarding the distributions for the service time, repair, and vacation durations, two types of systems are also considered. In the first all these distributions are exponential and the in the second all are deterministic.

In Figure 4 an exponential system with conditional disasters is considered. The mean number of customers is plotted against the total mean service time for four cases, namely small and large mean repair and vacation times. (The total mean system time is mean service time divided by $1-r$, i.e. the total expected time the customer spends in the server.)

A salient feature of the qualitative behavior of the system is the following: For large mean vacation time, when the total mean service time becomes very small, the mean number of customers in the system converges to a limit. An exact analysis shows that this limit is $\lambda\frac{\mathbb{E}U^2}{2\mathbb{E}U}$. Intuitively this is clear because, when the mean service time is close to zero, all customers are instantaneously served and thus there is no chance

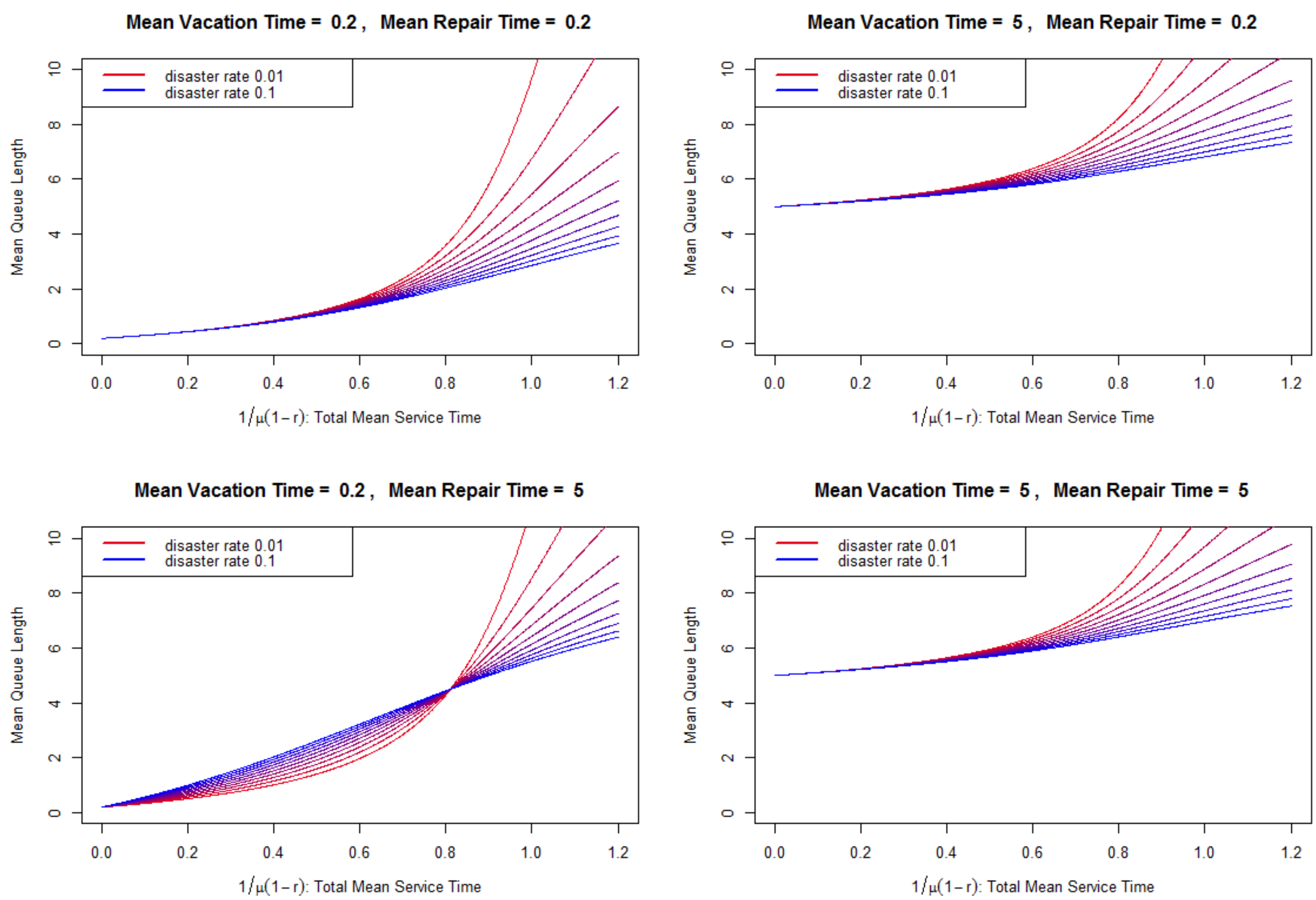

Figure 4: Unconditional Disasters - Exponential - Mean Queue Length as a function of the total mean service time $1/\mu(1-r)$. The disaster rate δ ranges from 0.01 (red) to 0.1 (blue). In general, lower disaster rates correspond to larger mean queue lengths since customers are not removed often from the system. However, on the lower left plot, corresponding to $m_U = 0.5$, $m_R = 5$, notice that for small values of the mean service time higher disaster rates (blue lines) correspond to larger mean queue length. This is due to the long repair periods, during which customer accumulate without being served.

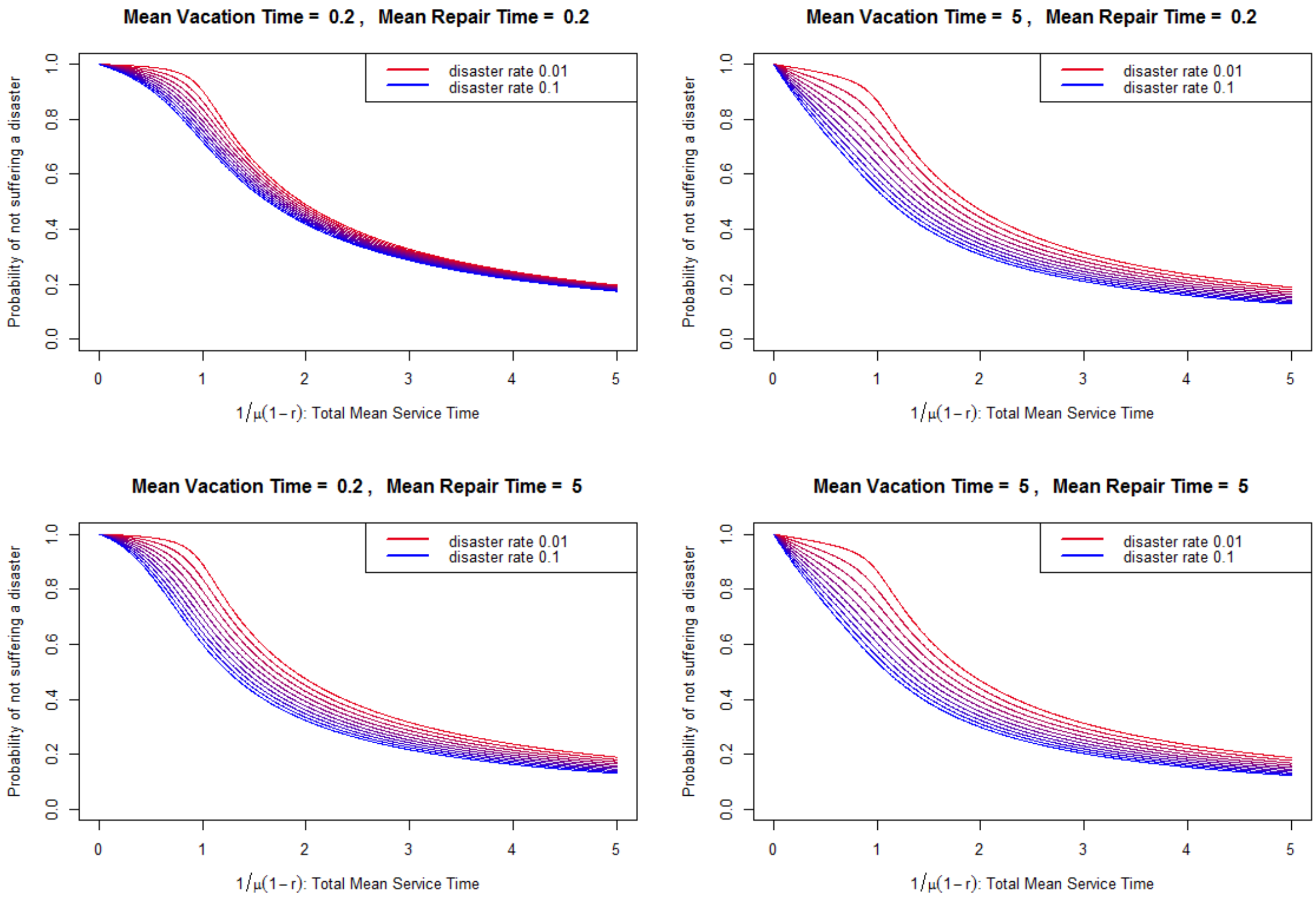

Figure 5: Unconditional Disasters - Exponential - Probability of Completing Service as a function of the total mean service time $1/\mu(1-r)$. The disaster rate δ ranges from 0.01 (red) to 0.1 (blue).

of a disaster occurring. As a result, arriving customers nearly always find the system in vacation. Therefore the mean system time for each of them is the mean residual life of the vacation distribution.

Figure 5 presents plots of the probability of a customer completing service for the same system as that of Figure 4. Such plots used in conjunction can provide support for deciding on the operational parameters of the system. Typically, as we have also seen in §7, the probability of completing service acts as a constraint. Notice the sigmoidal form of the plot, particularly when the disaster rate is low, which is the same qualitatively across values of the disaster rate and the values of the mean repair and vacation times.

In the system with unconditional disasters the mean number of customers in the system, $\Phi'(1)$, is given in (105) while the probability of completing service is given by (111). See also (112).

Figures 6 and 7 repeat the same for the system with unconditional disasters. Note that when the total mean service time exceeds 1 the systems would be unstable in the absence of disasters. In each plot the disaster rate δ takes values from 0.01 (red) to 0.1 (blue). Notice in particular the behavior of the mean queue length when the mean vacation time is small and the mean repair time is large. Also, from the above figures we see that, while there are of course quantitative differences there are no qualitative differences between the exponential and the deterministic system.

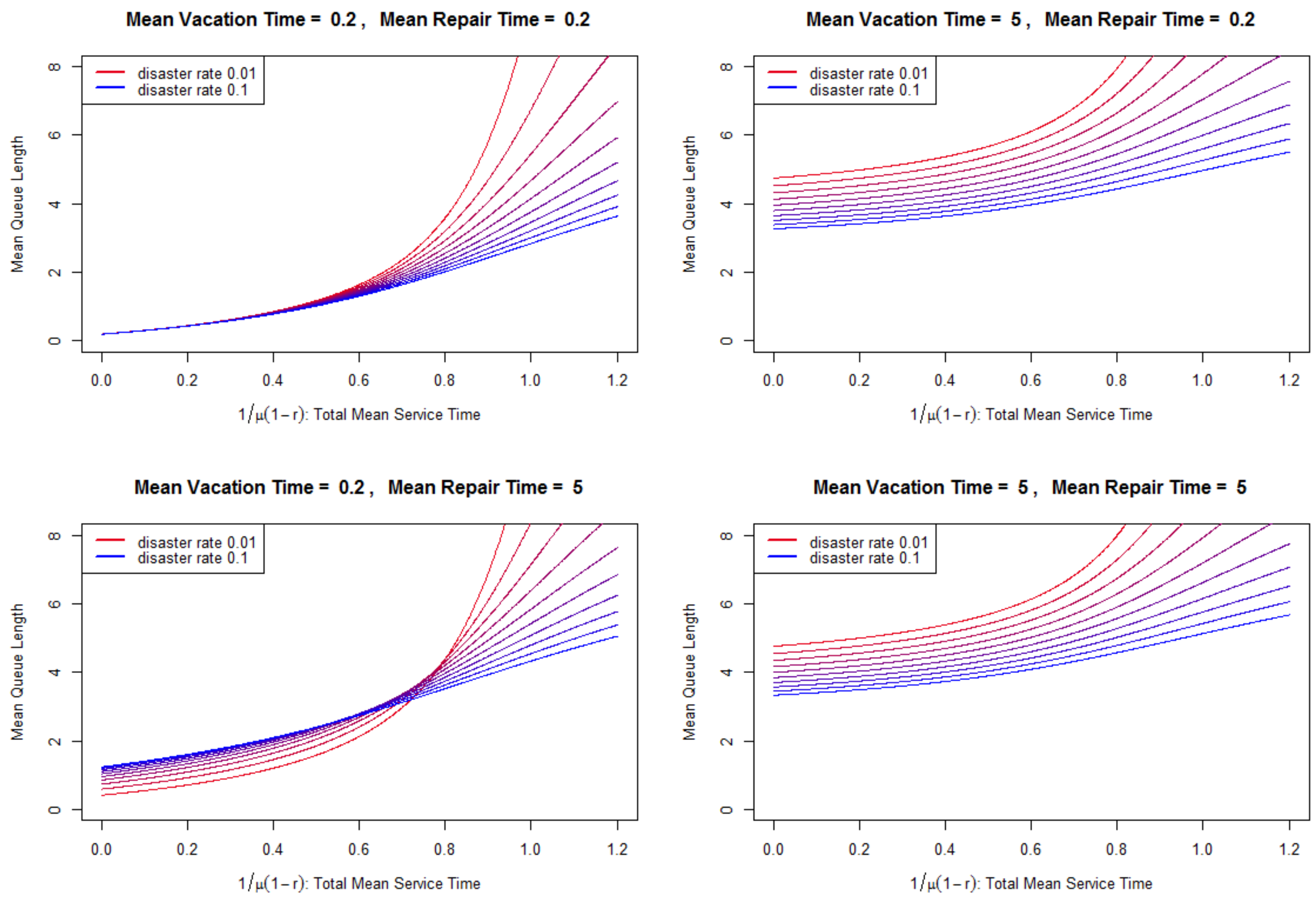

Figure 6: Conditional Disasters - Exponential - Mean Queue Length as a function of the total mean service time. The disaster rate δ ranges from 0.01 (red) to 0.1 (blue).

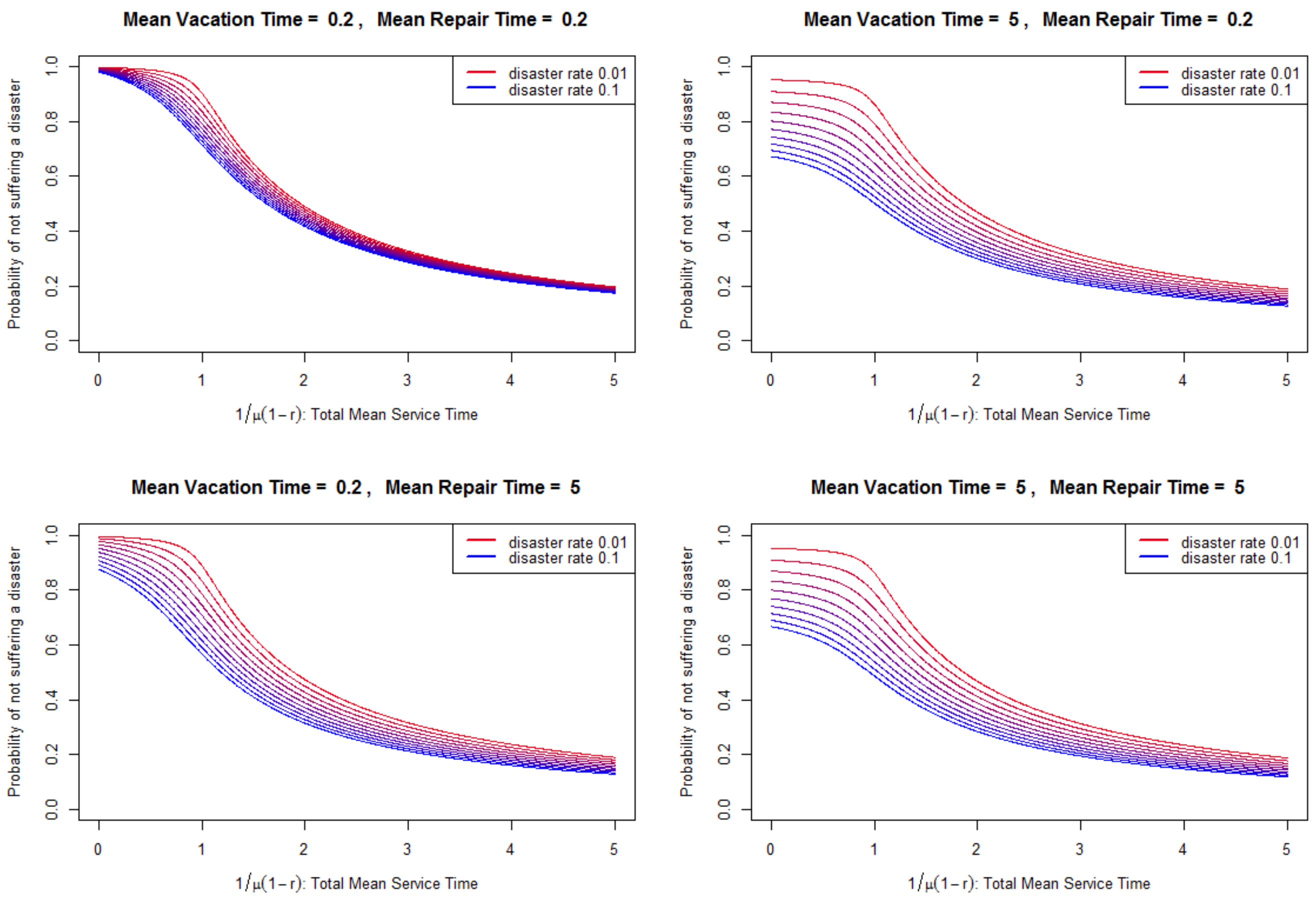

Figure 7: Conditional Disasters - Exponential - Probability of Completing Service as a function of the total mean service time. The disaster rate δ ranges from 0.01 (red) to 0.1 (blue).

10 Appendix – Determination of z_0

Proposition 5. *For $\delta > 0$ the equation*

$$z \;=\; (1-r+rz)\hat{S}(\delta+\alpha(z)) \tag{125}$$

has a single root, z_0, inside the unit disk $\{z \in \mathbb{C} : |z| < 1\}$ which is real and positive.

Proof. Let $h(z) := (1-r+rz)\hat{S}(\delta+\alpha(z))-z$. Since $h(0) = (1-r)\hat{S}(\delta+\lambda) > 0$ and $h(1) = \hat{S}(\delta)-1 < 0$ there exists $z_0 \in (0,1)$ such that $h(z_0) = 0$. This shows that (125) has a real and positive root within the unit disk. To complete the proof we must show that h has no other roots there. Let $f(z) := -z$ and $g(z) := (1-r+rz)\hat{S}(\delta+\alpha(z))$ which are both analytic in $|z| \le 1$. When $|z| \le 1$, $|1-r+rz| \le |1-r|+|rz| = 1-r+r|z| \le 1$. Thus

$$|g(z)| \;\le\; |1-r+rz| \int_0^\infty |\, e^{-(\delta+\alpha(z))x} \,|\, dS(x) \;\le\; \int_0^\infty e^{-\delta x}\, e^{-x\Re(\alpha(z))} dS(x).$$

The real part of $\alpha(z)$ when $|z| = 1$ (i.e. when $z = e^{i\theta}$ with $\theta \in [0,2\pi)$), is

$$\Re\left(\lambda\left(1-\textstyle\sum_{k=1}^\infty \chi_k e^{ik\theta}\right)\right) \;=\; \lambda \textstyle\sum_{k=1}^\infty \chi_k (1-\cos k\theta) \ge 0, \qquad \theta \in [0,2\pi)$$

and thus $|g(z)| \le \int_0^\infty e^{-\delta x} dS(x) < 1$ when $|z| = 1$. It follows by Rouché's theorem that $f(z)$ and $f(z)+g(z)$ will have the same number of zeros in the unit disk $|z| < 1$. Since $f(z)$ has only one zero in the unit disk (namely $z = 0$) $h(z)$ also has a single zero, z_0, there. □

References

[1] Asmussen S. *Applied Probability and Queues.* 2nd ed. New York: Springer Verlag; 2003.

[2] Baccelli F, Brémaud P. *Elements of Queueing.* 2nd ed. New York: Springer Verlag; 2003.

[3] Boxma, O. J., & Yechiali, U. (1997). An M/G/1 queue with multiple types of feedback and gated vacations. *Journal of Applied Probability*, 34(3), 773–784. https://doi.org/10.2307/3215102

[4] Bu Q. & Liu L. (2019). An $M/G/1$ clearing queueing system with setup time and multiple vacations for an unreliable server. *Communications in Statistics - Theory and Methods*, 48, 2810–2826. https://doi.org/10.1080/03610926.2018.1473426

[5] Chang, F. M., Liu, T.H., & Ke, J.C. (2018). On an unreliable-server retrial queue with customer feedback and impatience. *Applied Mathematical Modelling*, 55, 171-182. https://doi.org/10.1016/j.apm.2017.10.025

[6] Choi, S.D., Kim, B., & Choi, S.H. (2003). An $M/G/1$ queue with multiple types of feedback, gated vacations and FCFS policy. *Computers & Operations Research* 30, 1289–1309. https://doi.org/10.1016/S0305-0548(02)00071-0

[7] Choudhury, G. & Kalita C. (2018).An $M/G/1$ queue with two types of general heterogeneous service and optional repeated service subject to server's breakdown and delayed repair. *Quality Technology & Quantitative Management*, 15 (5), 622–654 http://dx.doi.org/10.1080/16843703.2017.1331499

[8] Choudhury, G., Tadj, L., & Ke, JC. (2011). A Two-Phase Service System with Bernoulli Vacation Schedule, Setup Time and N-policy for an Unreliable Server with Delaying Repair. *Quality Technology & Quantitative Management*, 8 (3), 271–284. http://dx.doi.org/10.1080/16843703.2011.11673259

[9] Jain M., & Bhargava. (2009). N-Policy Machine Repair System with Mixed Standbys and Unreliable Server. *Quality Technology & Quantitative Management*, 6 (2), 171–184 http://dx.doi.org/10.1080/16843703.2009.11673192

[10] Jain G, & Sigman K. (1996). A Pollaczek-Khintchine Formula for $M/G/1$ Queues with Disasters. *Journal of Applied Probability*,4, 1191–1200. https://doi.org/10.2307/3214996

[11] Jain M., & Kaur S. (2021). Bernoulli vacation model for $M^{[X]}/G/1$ Unreliable Server Retrial Queue with Bernoulli Feedback, Balking and Optional Service. *RAIRO*, 55, 2027–2053. https://doi.org/10.1051/ro/2020074

[12] Jain M., & Singh M. (2020). Transient Analysis of a Markov Queueing Model with Feedback, Discouragement and Disaster. *International Journal of Applied and Computational Mathematics*, 6 (31), https://doi.org/10.1007/s40819-020-0777-x

[13] Jain M., Kaur S. & Singh P. (2021). Supplementary variable technique (SVT) for non-Markovian single server queues with service interruption (QSI). *Operational Research*, 21, 2203–2246. https://doi.org/10.1007/s12351-019-00519-8

[14] Ke JC, & Chang FM. (2009). Modified vacation policy for $M/G/1$ retrial queue with balking and feedback. *Computers & Industrial Engineering*, 57(1), 433–443. https://DOI: 10.1016/j.cie.2009.01.002

[15] Ke JC, Wu CH, & Zhang ZG. (2010). Recent developments in vacation queueing models: A short survey. *International Journal of Operations Research*, 7(4), 3–8.

[16] Ke JC, Liu TH, Su S, & Zhang ZG. (2020). On retrial queue with customer balking and feedback subject to server breakdowns. *Communication in Statistics-Theory and Methods*, https://doi.org/10.1080/03610926.2020.1852432

[17] Ke JC, Liou & Yang (2018). Modeling of machine interference problem with unreliable repairman and standbys imperfect switchover. *Reliability Engineering and System Safety* 174, 12–18,

[18] Kim, J., Kim, B., & Luh, H. (2019). Analysis of a Markovian feedback queue with multi-class customers and its application to the weighted round-robin queuer. *Annals of Operations Research*, 277, 137–159. https://doi.org/10.1007/s10479-018-2917-9

[19] Kim, K., & Lee, H. (2014). The M/G/1 queue with disasters and working breakdowns. *Applied Mathematical Modelling*, 38, 1788–1798. https://doi.org/10.1016/j.apm.2013.09.016

[20] Kumar, BK., Krishnamoorthy, A., Madheswari, SP.,& Basha, SS. (2007). Transient analysis of a single server queue with catastrophes, failures and repairs. *Queueing Systems*, 56, 133–141. https://doi.org/10.1007/s11134-007-9014-0

[21] Liu, Z., Wua, J., & Yang, G. (2009). An $M/G/1$ retrial G-queue with preemptive resume and feedback under N-policy subject to the server breakdowns and repairs. *Computers and Mathematics with Applications*, 58, 1792–1807. https://DOI: 10.1016/j.camwa.2009.07.077

[22] Li K. & Wang J. (2020). Equilibrium balking strategies in the single server retrial queue with constant retrial rate and catastrophes. *Quality Technology & Quantitative Management*,18 (2), 156–178, https://doi.org/10.1080/16843703.2020.1760464

[23] Mytalas, G.,C., & Zazanis, M.A. (2015). An $M^{[X]}/G/1$ Queueing System with Disasters and Repairs under a Multiple Adapted Vacation Policy. *Nacal Research Logistics* 62(3), 171–189. https://doi.org/10.1002/nav.21621

[24] Madan K, Al-Rawwash M. (2005). On the $M^X/G/1$ queue with feedback and optional server vacations based on a single vacation policy. *Applied Mathematics and Computation*, 160, 909–919. https://doi.org/10.1016/j.amc.2003.11.037

[25] Semenova, O. V. (2007). Optimal Hysteresis Control for BMAP/SM/1 Queue with MAP-Input of Disasters. *Quality Technology & Quantitative Management*, 4(3), 395–405. http://dx.doi.org/10.1080/16843703.2007.11673159

[26] Sun W., Li S., Liu W., & Tian W. (2019). Comparisons of site-clearing and non-site-clearing Markovian queues with server failures and unreliable repairer. *Quality Technology & Quantitative Management*, 16, (2) 211–230. https://doi.org/10.1080/16843703.2017.1389144

[27] Takács L. (1963). A Single Server Queue with Feedback. *Bell System Technical Journal*, 42, 505–519. https://DOI: 10.1002/j.1538-7305.1963.tb00510.x

[28] Takagi H. *Queueing Analysis: A Foundation of Performance Evaluation, Vacation and Priority Systems. vol. 1*. Amsterdam: North Holland: 1991.

[29] Thangaraj V. & Vanitha S. (2009). A two phase $M/GI/1$ Feedback Queue with Multiple Server Vacations. *Stochastic Analysis and Applications*, 27, 1231–1245. https://doi.org/10.1080/07362990903259934

[30] Tian N, Ziang ZG. *Vacation Queueing Models, Theory and Applications.* New York: Springer Verlag; 2006.

[31] Rajadurai P, Saravanarajan MC, and Chandrasekaran VM. (2018). A study on $M/G/1$ feedback retrial queue with subject to server breakdown and repair under multiple working vacation policy. *Alexandria Engineering Journal*, 57 (2), 947–962. https://doi.org/10.1016/j.aej.2017.01.002

[32] Walrand J. *An Introduction to Queueing Networks.* Prentice Hall: 1988.

[33] Wortman MA, Disney R. (1991). The $M/GI/1$ Bernoulli feedback queue with vacations. *Queueing Systems*, 9, 353–364. https://doi.org/10.1007/BF01159222

[34] Yang D. & Ke JC (2014). Cost Optimization of a Repairable $M/G/1$ Queue with a Randomized Policy and Single Vacation. *Applied Mathematical Modelling*, 38, 21–22, 5113–5125, http://dx.doi.org/10.1016/j.apm.2014.03.012

[35] Yang D. & Wo C. (2015). Cost-minimization Analysis of a Working Vacation Queue with N-policy and Server Breakdowns. *Computers & Industrial Engineering*, 82, 151–158, http://dx.doi.org/10.1016/j.cie.2015.01.017

[36] Yashkov SF, Yashkova AS. (2009). On Busy Period and Sojourn Time Distributions in the $M/G/1$–EPS Queue with Catastrophes. *Automation and Remote Control*, 70, 2061–2072. https://doi.org/10.1134/S0005117909120133

[37] Yechiali U. (2007). Queues with system disasters and impatient customers when system is down. *Queueing Systems*, 56, 195–202. https://doi.org/10.1007/s11134-007-9031-z