Centrum voor Wiskunde en Informatica
Centre for Mathematics and Computer Science

F.A. van der Duyn Schouten, P. Wartenhorst

Time to failure, time to repair and availability of a two-unit standby system with Markovian degrading units

Department of Operations Research, Statistics, and System Theory Report BS-R9007 April

The Centre for Mathematics and Computer Science is a research institute of the Stichting Mathematisch Centrum, which was founded on February 11, 1946, as a nonprofit institution aiming at the promotion of mathematics, computer science, and their applications. It is sponsored by the Dutch Government through the Netherlands Organization for the Advancement of Research (N.W.O.).

Time to Failure, Time to Repair and Availability of a Two-Unit Standby System with Markovian Degrading Units

Frank A. van der Duyn Schouten
Centre for Mathematics and Computer Science
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands
Faculty of Economics, Tilburg University
P.O. Box 90153, 5000 LE Tilburg, The Netherlands

Pieter Wartenhorst
Centre for Mathematics and Computer Science
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

A two-unit cold standby system with Markovian degrading units and one repair facility is considered. Two types of repair are possible: preventive and corrective, where the latter is supposed to be more time consuming than the first. Repair of the units is controlled by a control limit rule. By probabilistic arguments we obtain explicit expressions for the Laplace transforms of the "time to system failure" and "time to system repair". Furthermore we show how to obtain the steady state availability and approximations for the interval availability distribution. Iterative schemes to numerically compute the quantities derived in this paper in case of generalized Erlangian distributed repair times are presented together with some numerical examples.

1980 Mathematics Subject Classification: 60K10, 90B25.
Key Words & Phrases: repairable systems, single repair facility, preventive maintenance, corrective maintenance, cold standby, Markovian degradation, time to system failure, time to system repair, (interval) availability.

1. Introduction

In studying the performance of a system one is often not only interested in steady state measures like the fraction of time the system will be available (cf. Kawai [7]), because such measures don't always provide enough information for practical purposes. A proper question in this context would be whether the system will be down seven and a half minute every hour or one hour every working day. For a gas production platform, for instance, the information that on average it is available 360 days a year is not sufficient in sales contracts. Whereas short interruptions in production may be covered by the inventory (buffer), a loss of production for several days may cause serious problems (van Rijn & Schornagel [10]). That is why one is often interested in transient measures like the cumulative operational time distribution over finite intervals (cf. de Souza e Silva & Gail [11]), or even in the characteristics of the alternating "time to system failure" and "time to system repair".

In this paper we consider a two-unit standby system with Markovian degrading units and one repair facility. If not both in failed condition, the units are in operation one at a time. Two types of repair can be applied (either preventive or corrective), with durations which have probability distributions of general type. This model has been introduced by Kawai [7]. Kawai shows that under certain regularity conditions the maintenance policy which maximizes the system availability is a Control Limit Rule (CLR), i.e. a preventive repair on the working unit is carried out if and only if the repair facility is free and the condition of the working unit is less than or equal to a prespecified critical level. Moreover using probabilistic arguments Kawai [7] derives an explicit expression for the

Report BS-R9007
Centre for Mathematics and Computer Science
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

availability of the system under a given CLR. Van der Duyn Schouten & Ronner [1] apply the embedding technique from Markov decision theory to obtain another more rigorous derivation of the availability under a CLR. They furthermore propose an iterative computational scheme to numerically compute the availability in case of Erlangian distributed repair times.

For the model we present the Laplace transforms of both the time to system failure and the time to system repair from which all moments can be obtained. From the first moments we obtain the availability result of Kawai [7]. We indicate how further knowledge about the distributions of the length of the alternating up and down periods can be used to obtain approximations for the interval availability distribution applying an approach developed by van der Heijden [5] and van der Heijden & Schornagel [6]. Employing similar ideas as van der Duyn Schouten & Ronner [1] we propose a more general iterative scheme to compute the required expressions in case of generalized Erlangian distributed repair times.

The remainder of this paper is organized as follows. In section 2 we give a detailed model description. In section 3 we derive the probability distribution and the Laplace transform of the "time to system failure" and we show how to compute the moments of this variable. Section 4 is devoted to the analysis of the "time to system repair". In section 5 we show that the results of sections 3 and 4 are in accordance with the long term availability result of Kawai [7]. Moreover we indicate how our results can be used to obtain approximations for the interval availability distribution. Finally in section 6 we show how all quantities mentioned in this paper can be calculated in case of generalized Erlangian distributions and we present some numerical examples.

To close this section we introduce some standard notation. Let $f_\tau(t)$ be the probability density function of an arbitrary nonnegative stochastic variable τ. We define

$$F_\tau(x) = \int_0^x f_\tau(t)\,dt \qquad \text{and} \qquad \phi_\tau(s) = \int_0^\infty e^{-st}[1 - F_\tau(t)]\,dt, \quad s \geqslant 0, \tag{1.1}$$

the cumulative distribution function and the Laplace transform of τ respectively. The moments of τ can be obtained from:

$$E[\tau^{n+1}] = (-1)^n\,(n+1)\,\phi_\tau^{(n)}(0), \quad n = 0,1,\ldots . \tag{1.2}$$

Here $\phi_\tau^{(n)}(s)$ denotes the nth derivative of $\phi_\tau(s)$; $\phi_\tau^{(0)}(s)$ equals $\phi_\tau(s)$.

2. Model description

We consider a two-unit standby system consisting of two identical Markovian degrading units (cf. Gertsbakh [2], ch. 3) and one repair facility. At most one of the units is working at a time. The other unit is either in repair or in standby position. The units may be in any of the states $\{i \mid i \in S = \{0,1,\ldots,n,n+1\}\}$ (0 : perfect state; $1 \ldots n$: degraded states; $n+1$: failed state). Under the absence of repair the condition of the working unit deteriorates according to a continuous time Markov process with absorbing state $n+1$ and infinitesimal generator $Q = (q_{ij}),\ i,j \in S$. We assume that $q_{ij} = 0$ for $j<i,\ i,j \in S$, which means that a unit cannot improve without being repaired. A unit in standby state neither fails nor degrades (cold standby).

In applications a Markovian degrading unit may be used for instance to represent a sub-system consisting of $n+1$ components which are subject to breakdown. During operation of the unit, one or more components could fail. The ability of the unit to recover (with a specified "coverage") to a fault-free configuration consisting of the remaining non-failed components can be incorporated easily (cf. Meyer [8]).

There are two types of repair. As long as a working unit has not entered state $n+1$ the option of a preventive (type 1) repair exists provided the repair facility is free. When the working unit enters state $n+1$ a corrective (type 2) repair is required. The type 1 and type 2 repair times form two mutually independent sequences of i.i.d. random variables with distribution functions G_1 and G_2 with finite means μ_1 and μ_2 and $G_k(0) = 0,\ k = 1,2$ respectively. Preventive repair times are independent of the actual state in which repair is carried out. The repair times are also independent of the sojourn times of the working unit in the different states. Repair on a unit is continued until it is completed (no preemption). A repaired unit is like new (state 0).

When a unit enters repair the other one takes the working position in state 0, and vice versa. At repair completion the repaired unit takes the cold standby position if the other unit is still functioning and not demanding preventive maintenance. Otherwise the repaired unit takes the working position immediately. A system down time starts when the working unit enters state $n+1$ while the other unit is still under repair and ends as soon as this repair is finished. We do not consider setup times.

The system is observed continuously in time. In this paper we consider a maintenance policy of control limit type. The Control Limit Rule (CLR(m)) is defined as follows. Maintain a unit as soon as the repair facility is free and the unit is in any of the states $m, m+1, \ldots, n+1$. The state m is called the control limit. Under some regularity conditions the optimal maintenance policy which maximizes the system long run availability is of CLR-type (Kawai [7]).

3. Time to system failure

Our analysis is mainly built upon the observation that the process describing the system's behaviour regenerates itself at the beginning of *every* system up period.

At first we note that after a system down time a system up time always starts with one unit entering operation and the other unit entering the repair facility for a corrective (type 2) repair. We call this system-state A and denote the time to system failure starting from this state A by τ_{SFA}. Analogously we define τ_{SFB} as the time to system failure starting in state B, where B is the system-state in which one unit enters operation and the other unit enters the repair facility for a preventive (type 1) repair. Note that states A and B are regeneration points for the system. They may be entered for several times during one system up period.

3.1. Probability distribution of τ_{SFA}

The computation of the probability distribution of τ_{SFA} will be reduced to the computation of the probability of occurrence of five mutually exclusive events.

We introduce some additional notation:

- $\{X(t),\ t \geq 0\}$: a continuous time Markov chain on $S = \{0, \ldots, n+1\}$ with infinitesimal matrix Q; $X(0) := 0$,
- $\sigma_m := inf\{t \geq 0 \mid m \leq X(t) \leq n\}$,
- $\tau_{n+1} := inf\{t \geq 0 \mid X(t) = n+1\}$,
- $R_k :=$ type k repair time; $P(R_k \leq t) = G_k(t),\ k = 1,2$.

Assume that after a system down time the system starts operating again at time 0. Now for t fixed we make the following important observation. Under CLR(m), the event $\{\tau_{SFA} > t\}$ occurs if and only if one of the following five mutually exclusive events occurs:

$E_1'(t)$: at time t the unit working at time 0 is still working and no maintenance is desired ($X(t) < m$).

$E_2'(t)$: at time t the unit working at time 0 is still working. Maintenance is desired ($m \leq X(t) \leq n$), but not allowed because the type 2 repair that started at time 0 hasn't been finished yet.

$E_3'(t)$: the unit working at time 0 enters $\{m, \ldots, n\}$ before the type 2 repair that started at time 0 has been finished. Repair is finished before t and before the working unit fails. The system enters state B (which is a regeneration point!) at repair completion. The system doesn't fail during the residual time to t.

$E_4'(t)$: repair is finished before the unit working at time 0 enters $\{m, \ldots, n+1\}$. The working unit enters $\{m, \ldots, n\}$ before time t. At that moment the system enters state B. The system doesn't fail during the residual time to t.

$E_5'(t)$: repair is finished before the unit working at time 0 enters $\{m, \ldots, n+1\}$. Before time t the working unit fails and the system enters state A (which is also a regeneration point!). The system doesn't fail during the residual time to t.

These events have the same probability of occurrence as:

$E_1(t)$: $\{ \min\{ \sigma_m, \tau_{n+1} \} > t \}$, (3.1.1)

$E_2(t)$: $\{ \sigma_m \leqslant t < \tau_{n+1} ;\ \ R_2 > t \}$, (3.1.2)

$E_3(t)$: $\{ \sigma_m \leqslant R_2 < \tau_{n+1} ;\ \ R_2 \leqslant t ;\ \ \tau_{SFB} > t - R_2 \}$, (3.1.3)

$E_4(t)$: $\{ R_2 \leqslant \sigma_m \leqslant t ;\ \ \tau_{SFB} > t - \sigma_m \}$, (3.1.4)

$E_5(t)$: $\{ R_2 \leqslant \tau_{n+1} < \sigma_m ;\ \ \tau_{n+1} \leqslant t ;\ \ \tau_{SFA} > t - \tau_{n+1} \}$. (3.1.5)

Note that a similar reasoning holds for the event $\{ \tau_{SFB} > t \}$ by replacing in (3.1.1)-(3.1.5) all indices "2" by "1".

For $j \in S$, and $k = 1,2$, let

$$H_j(t) := P(\, X(t) = j \mid X(0) = 0 \,), \tag{3.1.6}$$

$$\bar{G}_k(t) := 1 - G_k(t) = P(R_k > t), \tag{3.1.7}$$

$$S(j) := \{ X(\tau_{n+1}^-) = j \}, \tag{3.1.8}$$

that is, $S(j)$ is the event that the entrance into state $n+1$ takes place by a jump from state j.
The following lemma which has been proved by van der Duyn Schouten & Ronner [1] provides an expression for the probability that the Markov process $\{X(t),\ t \geqslant 0\}$ has entered state $n+1$ before time t.

LEMMA 3.1.

$$F_{\tau_{n+1}\,;\,S(j)}(t) := P(\, \tau_{n+1} \leqslant t\ ;\ S(j)\,) = q_{j,n+1} \int_0^t H_j(x)\, dx. \qquad \square$$

The next lemma provides an expression for the probability that the Markov process $\{X(t),\ t \geqslant 0\}$ has entered one of the states $\{m,\ldots,n\}$ before time t.

LEMMA 3.2.

$$F_{\sigma_m}(t) := P(\, \sigma_m \leqslant t\,) = \sum_{j=m}^{n} [\, H_j(t) + q_{j,n+1} \int_0^t H_j(x)\, dx \,].$$

PROOF:

$$\begin{aligned} P(\, \sigma_m \leqslant t\,) &= P(\, m \leqslant X(t) \leqslant n \vee \{ X(t) = n+1\ ;\ m \leqslant X(\tau_{n+1}^-) \leqslant n \}\,) \\ &= \sum_{j=m}^{n} H_j(t) + P(\, \tau_{n+1} \leqslant t\ ;\ \bigcup_{j=m}^{n} S(j)\,) \\ &= \sum_{j=m}^{n} H_j(t) + \sum_{j=m}^{n} q_{j,n+1} \int_0^t H_j(x)\, dx. \end{aligned}$$

The last equality is based on lemma 3.1. $\square$

From (3.1.1)-(3.1.5), using (3.1.6)-(3.1.8) and lemma's 3.1 and 3.2, $P(E_1(t))$-$P(E_5(t))$ can be obtained as follows:

$$\begin{aligned} P(E_1(t)) &= P(\, 0 \leqslant X(t) \leqslant m-1\,) \\ &= \sum_{j=0}^{m-1} H_j(t), \end{aligned}$$

$$
\begin{aligned}
P(E_2(t)) &= P(\, m \leqslant X(t) \leqslant n \;;\; R_2 > t \,) \\
&= \sum_{j=m}^{n} H_j(t)\overline{G}_2(t), \\
P(E_3(t)) &= P(\, m \leqslant X(R_2) \leqslant n \;;\; R_2 \leqslant t \;;\; \tau_{SFB} > t - R_2 \,) \\
&= \sum_{j=m}^{n} \int_0^t H_j(x) P(\, \tau_{SFB} > t - x)\, dG_2(x), \\
P(E_4(t)) &= P(\, \sigma_m \leqslant t \;;\; R_2 \leqslant \sigma_m \;;\; \tau_{SFB} > t - \sigma_m \,) \\
&= \int_0^t G_2(x) P(\, \tau_{SFB} > t - x)\, dF_{\sigma_m}(x) \\
&= \sum_{j=m}^{n} \int_0^t [\, H_j'(x) + q_{j,n+1} H_j(x) \,] G_2(x) P(\, \tau_{SFB} > t - x \,)\, dx, \\
P(E_5(t)) &= \int_0^t P(\, R_2 \leqslant x \;;\; \tau_{SFA} > t - x \,)\, dP(\, \tau_{n+1} \leqslant x \;;\; \bigcup_{j=0}^{m-1} S(j) \,) \\
&= \sum_{j=0}^{m-1} \int_0^t P(\, R_2 \leqslant x \;;\; \tau_{SFA} > t - x \,)\, dF_{\tau_{n+1}\,;\, S(j)}(x) \\
&= \sum_{j=0}^{m-1} q_{j,n+1} \int_0^t H_j(x) G_2(x) P(\, \tau_{SFA} > t - x \,)\, dx.
\end{aligned}
$$

In the last equality concerning $P(E_4(t))$ lemma 3.2 is used and in the last equality concerning $P(E_5(t))$ lemma 3.1 is used. $P(\tau_{SFA} > t)$ is obtained by summation of $P(E_1(t))$ to $P(E_5(t))$.

3.2. Laplace transform of τ_{SFA}

For notational convenience we introduce for $j \in S$ and $k = 1,2$:

$$P_{jk}(s) := \int_0^\infty e^{-st} H_j(t)\overline{G}_k(t)\, dt, \tag{3.2.1}$$

$$p_{jk}(s) := \int_0^\infty e^{-st} H_j(t)\, dG_k(t), \tag{3.2.2}$$

$$R_{jk}(s) := \int_0^\infty e^{-st} H_j(t) G_k(t)\, dt, \tag{3.2.3}$$

$$h_j(s) := \int_0^\infty e^{-st} H_j(t)\, dt. \tag{3.2.4}$$

Using the identity:

$$\int_0^\infty f(t)\, dG_k(t) = - \int_0^\infty f(t)\, d\overline{G}_k(t),$$

and Kolmogorov's forward differential equations:

$$
\begin{aligned}
H_j'(t) &= -\, q_j H_j(t) + \sum_{i=0}^{j-1} H_i(t) q_{ij} \\
&= \sum_{i=0}^{j} H_i(t) q_{ij},
\end{aligned}
$$

we find by partial integration of (3.2.2)-(3.2.3) for $k = 1,2$:

$$p_{0k}(s) = 1 - (s+q_0)P_{0k}(s), \tag{3.2.5}$$

$$p_{jk}(s) = \sum_{i=0}^{j} q_{ij}P_{ik}(s) - sP_{jk}(s), \quad 1 \leq j \leq n+1, \tag{3.2.6}$$

$$R_{jk}(s) = h_j(s)-P_{jk}(s), \quad 0 \leq j \leq n, \tag{3.2.7}$$

$$h_0(s) = \frac{1}{s + q_0}, \tag{3.2.8}$$

$$h_j(s) = \frac{1}{s + q_j}\sum_{i=0}^{j-1} q_{ij}h_i(s), \quad 1 \leq j \leq n. \tag{3.2.9}$$

So, $h_j(s)$, $j=0,...,n$, can be computed using (3.2.8) and (3.2.9). From (3.2.5)-(3.2.7) we see that if $P_{jk}(s)$ is known for $j \in S$, $k=1,2$, then all $R_{jk}(s)$ and $p_{jk}(s)$ are known. Actually one can show that $h_j(s)$ can be obtained from $P_{jk}(s)$ as well. Bearing this in mind one should realize that all quantities mentioned in the remainder of this section (except for (3.3.1)) can be computed if $P_{jk}(s)$ is known for $j \in S$, $k=1,2$. The computation of $P_{jk}(s)$ for generalized Erlangian distributed repair times is dealt with in section 6.

Now we shall derive closed form expressions for $\phi_A(s)$ and $\phi_B(s)$, the Laplace transforms of τ_{SFA} and τ_{SFB} respectively (cf. (1.1)).
For $m \in \{1,...,n+1\}$, and $k = 1,2$ let:

$$D_{mk}(s) := 1 - \sum_{j=m}^{n} (s+q_{j,n+1})R_{jk}(s), \tag{3.2.10}$$

$$A_{mk}(s) := \sum_{j=0}^{m-1} h_j(s) + \sum_{j=m}^{n} P_{jk}(s), \tag{3.2.11}$$

$$K_{mk}(s) := \sum_{j=0}^{m-1} q_{j,n+1}R_{jk}(s). \tag{3.2.12}$$

THEOREM 3.1.

$$\phi_A(s) = \frac{(1-D_{m2}(s))\,A_{m1}(s) + D_{m1}(s)\,A_{m2}(s)}{D_{m1}(s)\,(1-K_{m2}(s)) - (1-D_{m2}(s))\,K_{m1}(s)}, \tag{3.2.13}$$

and

$$\phi_B(s) = \frac{(1-K_{m2}(s))\,A_{m1}(s) + K_{m1}(s)\,A_{m2}(s)}{D_{m1}(s)\,(1-K_{m2}(s)) - (1-D_{m2}(s))\,K_{m1}(s)}. \tag{3.2.14}$$

PROOF:

$$\int_0^\infty e^{-st}P(E_1(t))\,dt = \sum_{j=0}^{m-1} h_j(s), \tag{3.2.15}$$

$$\int_0^\infty e^{-st}P(E_2(t))\,dt = \sum_{j=m}^{n} P_{j2}(s), \tag{3.2.16}$$

$$\int_0^\infty e^{-st}P(E_3(t))\,dt = \sum_{j=m}^{n} \int_{x=0}^{\infty} H_j(x) \int_{t=x}^{\infty} e^{-st}P(\tau_{SFB} > t-x)\,dt\,dG_2(x)$$

$$= \phi_B(s)\sum_{j=m}^{n} \int_{x=0}^{\infty} e^{-sx}H_j(x)\,dG_2(x)$$

$$= \phi_B(s) \sum_{j=m}^{n} p_{j2}(s), \tag{3.2.17}$$

$$\int_0^\infty e^{-st} P(E_4(t))\, dt = \sum_{j=m}^{n} \int_{x=0}^{\infty} e^{-sx}[\, H_j'(x) + q_{j,n+1} H_j(x)\,] G_2(x) \int_{t=x}^{\infty} e^{-s(t-x)} P(\tau_{SFB} > t-x)\, dt\, dx$$

$$= \phi_B(s) \sum_{j=m}^{n} [\, q_{j,n+1} R_{j2}(s) + \int_{x=0}^{\infty} e^{-sx} H_j'(x) G_2(x)\, dx\,]$$

$$= \phi_B(s) \sum_{j=m}^{n} [\, q_{j,n+1} R_{j2}(s) + s \int_0^\infty e^{-sx} H_j(x) G_2(x)\, dx - \int_0^\infty e^{-sx} H_j(x)\, dG_2(x)\,]$$

$$= \phi_B(s) \sum_{j=m}^{n} [\, (s + q_{j,n+1}) R_{j2}(s) - p_{j2}(s)\,], \tag{3.2.18}$$

$$\int_0^\infty e^{-st} P(E_5(t))\, dt = \sum_{j=0}^{m-1} q_{j,n+1} \int_{t=0}^{\infty} e^{-st} \int_{x=0}^{t} H_j(x) G_2(x) P(\tau_{SFA} > t-x)\, dx\, dt$$

$$= \sum_{j=0}^{m-1} q_{j,n+1} \int_{x=0}^{\infty} e^{-sx} H_j(x) G_2(x) \int_{t=x}^{\infty} e^{-s(t-x)} P(\tau_{SFA} > t-x)\, dt\, dx$$

$$= \phi_A(s) \sum_{j=0}^{m-1} q_{j,n+1} R_{j2}(s). \tag{3.2.19}$$

According to (1.1) and (3.1.1)-(3.1.5) $\phi_A(s)$ is obtained by summation of (3.2.15)-(3.2.19):

$$\phi_A(s) = \sum_{j=0}^{m-1} h_j(s) + \sum_{j=m}^{n} P_{j2}(s) + \phi_B(s) \sum_{j=m}^{n} (s + q_{j,n+1}) R_{j2}(s) + \phi_A(s) \sum_{j=0}^{m-1} q_{j,n+1} R_{j2}(s), \tag{3.2.20}$$

and replacing all indices "2" by "1":

$$\phi_B(s) = \sum_{j=0}^{m-1} h_j(s) + \sum_{j=m}^{n} P_{j1}(s) + \phi_B(s) \sum_{j=m}^{n} (s + q_{j,n+1}) R_{j1}(s) + \phi_A(s) \sum_{j=0}^{m-1} q_{j,n+1} R_{j1}(s). \tag{3.2.21}$$

After some simple algebra (3.2.13) and (3.2.14) are obtained from (3.2.20) and (3.2.21). □

3.3. Moments of τ_{SFA}

Using (1.2) we obtain the expected time to system failure $E[\tau_{SFA}]$ by setting $s := 0$ in (3.2.13). Here we present the calculation of the second moment $E[\tau_{SFA}^2]$. One could obtain an explicit expression directly by taking the derivative of $\phi_A(s)$ in (3.2.13). A more efficient way is to use (3.2.20) and (3.2.21) and compute $\phi_A(s)$ and $\phi_A'(s)$ subsequently.
From (3.2.20) we obtain:

$$\phi_A'(s) = \sum_{j=0}^{m-1} h_j'(s) + \sum_{j=m}^{n} P_{j2}'(s) + \phi_B'(s) \sum_{j=m}^{n} (s + q_{j,n+1}) R_{j2}(s) + \phi_B(s) \sum_{j=m}^{n} \{ R_{j2}(s) + (s + q_{j,n+1}) R_{j2}'(s) \}$$

$$+ \phi_A'(s) \sum_{j=0}^{m-1} q_{j,n+1} R_{j2}(s) + \phi_A(s) \sum_{j=0}^{m-1} q_{j,n+1} R_{j2}'(s). \tag{3.3.1}$$

Replacing all indices "2" by "1" at the right hand side of the latter equation gives $\phi_B'(s)$. Now again we have a system of two equations and two unknowns from which $\phi_A'(s)$ can be solved easily. Note that we can express $p_{jk}'(s)$ and $R_{jk}'(s)$ in terms of $P_{jk}'(s)$ analogous to the way we expressed $p_{jk}(s)$ and $R_{jk}(s)$ in terms of $P_{jk}(s)$ in (3.2.5)-(3.2.7). Starting with $h_0'(s)$ we obtain $h_j'(s)$ iteratively, cf. (3.2.8) and (3.2.9).

4. Time to system repair

The time to system repair, which is defined as the length of an arbitrary system down time, is denoted by τ_{SR}. In this section we shall subsequently derive the distribution function of τ_{SR} and its Laplace transform from which all moments of τ_{SR} can be obtained by straightforward differentiation.

4.1. Probability distribution of τ_{SR}

We use the observation that τ_{SR} equals the residual repair time of the unit under repair at the moment the system goes down because of failure of the working unit. Using the theory of regenerative processes (Ross [9]) we note that at the beginning of an arbitrary down period there is a type k repair going on with a probability, which does not depend on the particular regeneration cycle. We denote this probability by p_k, $k = 1,2$ ($p_1+p_2=1$).

Let

- $D_k :=$ length of an arbitrary down period given that a type k repair is going on at the beginning of that down period.

Conditioning on the type of repair to be completed during the time to system repair we obtain:

$$P(\tau_{SR} > t) = \sum_{k=1}^{2} p_k P(D_k > t). \tag{4.1.1}$$

First we shall derive an expression for p_k, $k = 1,2$. For this purpose we study an embedded Markov chain of the process describing the state of the system.

For $l \geqslant 0$ let

- $T^l :=$ the lth epoch at which one unit enters operation and the other one goes under repair (either preventive or corrective),
- $Z^l :=$ type of repair that started at time T^l,
- $\rho^l :=$ time to finish the repair that started at time T^l,
- $\tau_{n+1}^l :=$ time to failure of the unit that entered operation at time T^l (where $\tau_{n+1}^l = \infty$ if at T^{l+1} a preventive repair is applied).

We assume that the process starts at time T^0. Also we assume that $Z^0 = 2$. Then $\{Z^l, l \geqslant 0\}$ is a positive recurrent Markov chain on the state space $\{1,2\}$.

For $k = 1,2$ we define the steady state probabilities:

$$\pi_k := \lim_{l\to\infty} P(Z^l = k),$$

which are known to be (van der Duyn Schouten & Ronner [1]):

$$\pi_1 = \frac{1 - D_{m2}(0)}{D_{m1}(0) + 1 - D_{m2}(0)}\ ; \quad \pi_2 = \frac{D_{m1}(0)}{D_{m1}(0) + 1 - D_{m2}(0)}\ , \tag{4.1.2}$$

where $D_{m1}(0)$ and $D_{m2}(0)$ are defined in (3.2.10).

Using Bayes' formula we prove the following lemma.

LEMMA 4.1.

$$p_k = \frac{p_{n+1,k}(0)\pi_k}{\sum_{j=1}^{2} p_{n+1,j}(0)\pi_j}\ , \ k = 1,2. \tag{4.1.3}$$

PROOF:

$$p_k = \lim_{l\to\infty} P(Z^l = k \mid \tau_{n+1}^l < \rho^l)$$

$$= \lim_{l\to\infty} \frac{P(\tau_{n+1}^l < \rho^l \mid Z^l = k)\, P(Z^l = k)}{\sum_{j=1}^{2} P(\tau_{n+1}^l < \rho^l \mid Z^l = j)\, P(Z^l = j)}$$

$$= \frac{P(\tau_{n+1} < R_k)\pi_k}{\sum_{j=1}^{2} P(\tau_{n+1} < R_j)\pi_j} . \qquad \square$$

REMARK 4.1: If we assume that $\{\overline{G}_1(t) \leqslant \overline{G}_2(t)$, for all $t \geqslant 0\}$, then $P(\tau_{n+1} < R_1) \leqslant P(\tau_{n+1} < R_2)$ and so we see from lemma 4.1 that $p_1 \leqslant \pi_1$ and $p_2 \geqslant \pi_2$.

The distribution function of τ_{SR} can now be obtained easily.

LEMMA 4.2.

$$P(\tau_{SR} > t) = \sum_{k=1}^{2} p_k \left[\frac{\int_0^\infty \overline{G}_k(t+y)\, dH_{n+1}(y)}{p_{n+1,k}(0)} \right] . \tag{4.1.4}$$

PROOF:

$$P(D_k > t) = P(R_k - \tau_{n+1} > t \mid R_k > \tau_{n+1}) \tag{4.1.5}$$

$$= \frac{P(R_k - \tau_{n+1} > t\, ;\, R_k > \tau_{n+1})}{P(R_k > \tau_{n+1})}$$

$$= \frac{\int_0^\infty \overline{G}_k(t+y)\, dH_{n+1}(y)}{\int_0^\infty \overline{G}_k(y)\, dH_{n+1}(y)} .$$

By partial integration and using (4.1.1) we obtain (4.1.4). □

REMARK 4.2: Due to the memoryless property we see from (4.1.1) and (4.1.5) that in case of exponentially distributed repair times the length of an arbitrary down period has a hyperexponential distribution.

4.2. Laplace transform of τ_{SR}

Define for $k = 1,2$:

$$V_{n+1,k}(s) := \int_{z=0}^{\infty} \int_{y=0}^{z} e^{-s(z-y)} H_{n+1}(y) \overline{G}_k(z)\, dy\, dz, \tag{4.2.1}$$

and let $\phi_{\tau_{SR}}(s)$ be the Laplace transform of τ_{SR} (cf. (1.1)).

THEOREM 4.1.

$$\phi_{\tau_{SR}}(s) = \sum_{k=1}^{2} p_k \left[\frac{P_{n+1,k}(0) - sV_{n+1,k}(s)}{p_{n+1,k}(0)} \right] . \tag{4.2.2}$$

PROOF:

Using lemma 4.2 we obtain:

$$\phi_{D_k}(s) = \int_0^\infty e^{-st} P(D_k > t)\, dt$$

$$= \frac{1}{p_{n+1,k}(0)} \int_{y=0}^{\infty} \int_{t=0}^{\infty} e^{-st}\overline{G}_k(t+y)\, dt\, dH_{n+1}(y)$$

$$= \frac{1}{p_{n+1,k}(0)} \int_{z=0}^{\infty} \int_{y=0}^{z} e^{-s(z-y)}\overline{G}_k(z)\, dH_{n+1}(y)\, dz$$

$$= \frac{1}{p_{n+1,k}(0)} \int_{z=0}^{\infty} \left[e^{sz}H_{n+1}(z) - s\int_{y=0}^{z} e^{sy}H_{n+1}(y)\, dy \right] e^{-sz}\overline{G}_k(z)\, dz$$

$$= \frac{1}{p_{n+1,k}(0)} \left[\int_{z=0}^{\infty} H_{n+1}(z)\overline{G}_k(z)\, dz - s\int_{z=0}^{\infty}\int_{y=0}^{z} e^{-s(z-y)}H_{n+1}(y)\overline{G}_k(z)\, dy\, dz \right]$$

$$= \frac{1}{p_{n+1,k}(0)} \left[P_{n+1,k}(0) - sV_{n+1,k}(s) \right].$$

Combining this last result with (4.1.1) gives (4.2.2). □

As in the previous section we note that all expressions mentioned in this section can be calculated if $P_{jk}(s)$ and $V_{n+1,k}(s)$ are known for $j=0,...,n+1$, $k=1,2$. In general $P_{jk}(s)$ and $V_{n+1,k}(s)$ have to be evaluated by numerical integration techniques. In case of generalized Erlangian distributed repair times however, they can be obtained iteratively using lemma's 6.1 and 6.2 in section 6.

5. Availability

5.1. Average availability

The two-unit standby system as defined in section 2 constitutes an alternating renewal process (Ross [9]). That is, the process starts over again after a complete cycle consisting of an up and down interval.

We define

- $g(m) :=$ long run proportion of time the system is down when CLR(m) is used,

that is, $1-g(m)$ is the availability of the system. From the theory of regenerative processes it is known that (Ross [9]):

$$g(m) = \frac{E[\tau_{SR}]}{E[\tau_{SFA}] + E[\tau_{SR}]}. \tag{5.1.1}$$

Using (3.2.13) and (4.2.2), we obtain after some tedious calculations from (5.1.1):

$$g(m) = \frac{(1-D_{m2}(0))\, P_{n+1,1}(0) + D_{m1}(0)\, P_{n+1,2}(0)}{(1-D_{m2}(0))\, H_{m1} + D_{m1}(0)\, H_{m2}}, \quad 1 \leqslant m \leqslant n+1, \tag{5.1.2}$$

where

$$H_{mk} := \mu_k + \sum_{j=0}^{m-1} R_{jk}(0), \quad k = 1,2,$$

$\mu_k :=$ mean type k repair time, $k = 1,2$ (cf. section 2).

Formula (5.1.2) has been found already by Kawai [7]. Kawai uses intuitive probabilistic arguments to derive this formula. Van der Duyn Schouten & Ronner [1] proved (5.1.2) using the embedding technique from Markov decision theory.

5.2. Interval availability distribution

In production planning one often requires the production system to be available for more than a certain fraction of time over a finite observation period. Steady state measures like the average availability $(1-g(m))$ mentioned above don't always give enough guarantee of meeting this required level of availability. In such a case the evaluation of the probability of meeting this interval availability requirement becomes necessary (e.g. Goyal & Tantawi [4]).

If the behaviour of a system is modeled as a homogeneous continuous-time Markov process a very elegant way to obtain the interval availability distribution is to make use of the uniformization technique (de Souza e Silva & Gail [11]). If the state space grows too large however this technique becomes computationally infeasible. Aggregation techniques become necessary to analyze the system.

Promising results in this area are obtained by van der Heijden [5] and van der Heijden & Schornagel [6]. In the following we will shortly present some basic elements of their approach applied to our model.

We approximate our two-unit standby system by a two state single component system which alternates between up- and down-state. Sojourn times in up- and down-state constitute two sequences of i.i.d. random variables which are assumed to be independent of each other. Appropriate probability distribution functions ($G(.)$ and $H(.)$) of the length of the alternating up and down periods have to be extracted from our analysis in section 3 and 4 respectively.

For example in case of exponentially distributed repair times arbitrary down times turn out to be hyperexponentially distributed (remark 4.2). In case of highly reliable components unit lifetimes last considerable longer than unit repair times. During fault-free operation the system may reenter system states A and B for several times (section 3). In such a case the process describing the system's behaviour during system up time is called a 'regenerating process with rare events' (rare event: repair of one unit). The exponential distribution with mean to be obtained from (3.2.13) might be a proper approximation for the distribution of system up time in this case (cf. Gertsbakh [3] and section 6).

Assuming that we start observing the system at the beginning of an arbitrary up period, we define:

$$\beta(t) := \text{total sojourn time in down state during } [0,t],$$

$$\Omega(t,x) := P(\beta(t) \leqslant x).$$

Using the results of Takács [12] we obtain:

$$\Omega(t,x) := \sum_{n=0}^{\infty} H_n(x) \left[G_n(t-x) - G_{n+1}(t-x) \right], \quad 0 \leqslant x \leqslant t, \tag{5.2.1}$$

where we denote by $G_n(.)$ and $H_n(.)$ the n-fold convolution of $G(.)$ and $H(.)$ respectively; $G_0(z) = H_0(z) = 1$, $0 \leqslant z \leqslant t$. A similar result can be obtained if one starts observing the system at the beginning of an arbitrary down period.

Let $IA(t)$ be the interval availability during $[0,t]$, (i.e. the fraction of time during $[0,t]$ that the system is functioning). We easily see that:

$$P(IA(t) \geqslant z) = \Omega(t,(1-z)t), \quad 0 \leqslant z \leqslant 1.$$

There are some difficulties in using this approach. The first problem that one encounters is the fact that in general the first observed period is not an arbitrary complete up- or down period. When at the beginning of the first period both units are new the first up period will be stochastically larger than an arbitrary one (recall that an arbitrary up period starts when one unit takes the working position and the other one enters the repair facility for a corrective repair). Another possibility is that one starts observing the system at some point in time so that the first observed period can either be a residual up period or a residual down period.

The second difficulty is that the length of a particular down period depends on the length of the up period immediately preceding it, so the independency assumption needed for (5.2.1) to hold does not apply here. For example in case some particular up period has very short length, there is a high probability that the corrective repair started at the beginning of this system up period has not been completed at system failure. So the probability that at the beginning of the down period immediately following this particular up period a corrective (type 2) repair is going on exceeds p_2 (4.1.3). We also know that this repair still has to be performed almost completely, because it was going on only for a very short while when the system broke down. As it seems natural to assume corrective repair times to be stochastically larger than preventive ones, we conclude from both observations above that the system down time immediately following a very short system up time is stochastically larger than an arbitrary one (cf. (4.1.1)). In case of highly reliable units however, this dependency does not seem to influence the results considerably (van der Heijden & Schornagel [6]).

Finally we mention the fact that using (5.2.1) involves evaluation of an infinite summation, consisting of multiple convolutions, which has to be truncated properly. For further computational aspects of this approximative method we refer to van der Heijden [5], van der Heijden & Schornagel [6] and van Rijn & Schornagel [10].

6. Computational aspects and numerical examples

To compute $\phi_A(s)$ from (3.2.13) and (3.2.14) using (3.2.5)-(3.2.7), the only difficulty is the computation of $P_{jk}(s)$. In order to compute $\phi_{T_{SR}}(s)$ from (4.2.2) we need to compute $V_{n+1,k}(s)$. Generally $P_{jk}(s)$ and $V_{n+1,k}(s)$ have to be computed by numerical integration techniques. In case of generalized Erlangian distributed repair times however, $P_{jk}(s)$ and $V_{n+1,k}(s)$ can be computed iteratively using lemma 6.1 and lemma 6.2 below respectively.

For notational convenience we omit the index k in the following expressions. So, in case of *Erlang*(p,λ) distributed repair times we denote:

$$\overline{G}^p(t) := \sum_{l=0}^{p-1} e^{-\lambda t}\frac{(\lambda t)^l}{l!}, \quad p \in \{1,2,...\},\ \lambda > 0, \tag{6.1}$$

$$P_j^p(s) := \int_0^\infty e^{-st} H_j(t)\overline{G}^p(t)\, dt, \tag{6.2}$$

$$V_{n+1}^p(s) := \int_{z=0}^\infty \int_{y=0}^z e^{-s(z-y)} H_{n+1}(y)\overline{G}^p(z)\, dy\, dz. \tag{6.3}$$

LEMMA 6.1.

Let $\overline{G}^p(t)$ and $P_j^p(s)$ be as defined in (6.1) and (6.2) respectively. Then

$$P_0^p(s) = \frac{1}{s+q_0+\lambda}\sum_{l=0}^{p-1}\left[\frac{\lambda}{s+q_0+\lambda}\right]^l,$$

$$P_j^p(s) = \frac{1}{s+q_j+\lambda}\sum_{i=0}^{j-1} q_{ij}\sum_{l=0}^{p-1}\left[\frac{\lambda}{s+q_j+\lambda}\right]^l P_i^{p-l}(s), \quad 1 \leq j \leq n+1.$$

PROOF:
For $j = 0$ we have:

$$\begin{aligned} P_0^p(s) &= \int_0^\infty e^{-st}e^{-q_0 t}\overline{G}^p(t)\, dt \\ &= \sum_{l=0}^{p-1}\frac{\lambda^l}{l!}\int_0^\infty e^{-(s+q_0+\lambda)t}t^l\, dt \\ &= \sum_{l=0}^{p-1}\lambda^l\frac{1}{(s+q_0+\lambda)^{l+1}}\, . \end{aligned}$$

For $1 \leq j \leq n+1$ using Kolmogorov's forward integral equations:

$$H_j(t) = e^{-q_j t} \int_{x=0}^{t} \sum_{i=0}^{j-1} q_{ij} H_i(x) e^{q_j x}\, dx,$$

we find from (6.1) and (6.2):

$$\begin{aligned}
P_j^p(s) &= \sum_{i=0}^{j-1} q_{ij} \int_{t=0}^{\infty} e^{-st}\overline{G}^p(t) e^{-q_j t} \int_{x=0}^{t} H_i(x) e^{q_j x}\, dx \\
&= \sum_{i=0}^{j-1} q_{ij} \int_{x=0}^{\infty} H_i(x) e^{q_j x} \int_{t=x}^{\infty} e^{-(s+q_j)t}\overline{G}^p(t)\, dt\, dx \\
&= \sum_{i=0}^{j-1} q_{ij} \int_{x=0}^{\infty} H_i(x) e^{q_j x} \sum_{l=0}^{p-1} \frac{\lambda^l}{l!} \int_{t=x}^{\infty} e^{-(s+q_j+\lambda)t} t^l\, dt\, dx \\
&= \sum_{i=0}^{j-1} q_{ij} \int_{x=0}^{\infty} H_i(x) e^{q_j x} \sum_{l=0}^{p-1} \lambda^l e^{-(s+q_j+\lambda)x} \sum_{m=0}^{l} \frac{x^m}{m!\,(s+q_j+\lambda)^{l-m+1}}\, dx \\
&= \frac{1}{s+q_j+\lambda} \sum_{i=0}^{j-1} q_{ij} \sum_{l=0}^{p-1} \left[\frac{\lambda}{s+q_j+\lambda}\right]^l \int_{x=0}^{\infty} e^{-sx} H_i(x) \sum_{m=0}^{p-l-1} e^{-\lambda x}\frac{(\lambda x)^m}{m!}\, dx. \qquad \square
\end{aligned}$$

In a similar way we can prove:

LEMMA 6.2.

Let $\overline{G}^p(t)$, $P_j^p(s)$ and $V_{n+1}^p(s)$ be as defined in (6.1), (6.2) and (6.3) respectively. Then

$$V_{n+1}^p(s) = \frac{1}{s+\lambda} \sum_{l=0}^{p-1} \left[\frac{\lambda}{s+\lambda}\right]^l P_{n+1}^{p-l}(0). \qquad \square$$

Note that lemma's 6.1 and 6.2 can easily be extended to the case of generalized Erlangian distributed repair times. Such distributions can be used to approximate arbitrarily closely the distribution of any nonnegative stochastic variable (Tijms [13], pp. 397-399). For example if $G(.)$ is a mixture of an $Erlang(n_1,\lambda)$ distribution and an $Erlang(n_2,\lambda)$ distribution then

$$G(t) := p\left[1 - \sum_{l=0}^{n_1-1} e^{-\lambda t}\frac{(\lambda t)^l}{l!}\right] + (1-p)\left[1 - \sum_{l=0}^{n_2-1} e^{-\lambda t}\frac{(\lambda t)^l}{l!}\right], \quad 0 \leq p \leq 1,$$

$$P_j(s) = p\, P_j^{n_1}(s) + (1-p)\, P_j^{n_2}(s), \quad 0 \leq p \leq 1,$$

$$V_{n+1}(s) = p\, V_{n+1}^{n_1}(s) + (1-p)\, V_{n+1}^{n_2}(s), \quad 0 \leq p \leq 1,$$

where we obtain $P_j^{n_1}(s)$, $P_j^{n_2}(s)$, $V_{n+1}^{n_1}(s)$ and $V_{n+1}^{n_2}(s)$ from lemma 6.1 and lemma 6.2 respectively.

Finally we shall illustrate the results of this paper by some numerical examples. Let $n = 7$ and let the Q-matrix be given by:

$$
Q = \begin{array}{c|ccccccccc|}
 & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\
0 & -1 & 0.98 & & & & & & & 0.02 \\
1 & & -2 & 1.95 & & & & & & 0.05 \\
2 & & & -3 & 2.90 & & & & & 0.10 \\
4 & & & & -4 & 3.80 & & & & 0.20 \\
4 & & & & & -5 & 4.70 & & & 0.30 \\
5 & & & & & & -6 & 5.50 & & 0.50 \\
6 & & & & & & & -7 & 6.30 & 0.70 \\
7 & & & & & & & & -8 & 8 \\
8 & & & & & & & & & 0
\end{array}
$$

In table 1 the following repair time distributions are considered:

- $G_1(.)$ is $Erlang(2,\mu_1)$ and $Erlang(4,2\mu_1)$ respectively with $\mu_1 = 2.2, 2.6, 3.0, 4.0$, and 8.0,
- $G_2(.)$ is $Erlang(2,2)$ and $Erlang(4,4)$ respectively.

Let c_x^2 denote the coefficient of variation of a stochastic variable X with expectation E_x and variance σ_x^2. Then $c_x^2 := \sigma_x^2 \;/\; (E_x)^2$, where $\sigma_x^2 := E(X-E_x)^2$. In table 1 below we give m^*, the optimal value of m; $1-g(m^*)$, the availability; E_u, c_u^2 and E_d, c_d^2, the expected up- and down time with their coefficients of variation respectively.

G_2		*Erl.* (2,2.0)						*Erl.* (4,4.0)					
G_1	μ_1	m^*	$1-g(m^*)$	E_u	c_u^2	E_d	c_d^2	m^*	$1-g(m^*)$	E_u	c_u^2	E_d	c_d^2
	2.2	7	0.9736	23.13	1.05	0.63	0.88	8	0.9813	25.39	1.04	0.48	0.78
	2.6	5	0.9801	27.59	1.10	0.56	0.91	7	0.9839	30.50	1.07	0.50	0.84
Erl. $(2,\mu_1)$	3.0	4	0.9849	33.17	1.14	0.51	0.95	5	0.9875	36.20	1.11	0.46	0.84
	4.0	3	0.9913	50.44	1.20	0.44	1.10	3	0.9930	53.37	1.16	0.38	0.88
	8.0	1	0.9968	101.44	1.25	0.32	1.75	1	0.9976	107.13	1.20	0.26	1.27
	2.2	5	0.9801	25.01	1.09	0.51	0.90	7	0.9839	28.09	1.07	0.46	0.78
	2.6	4	0.9856	31.72	1.14	0.46	0.97	5	0.9882	34.68	1.11	0.41	0.80
Erl. $(4,2\mu_1)$	3.0	3	0.9892	38.61	1.18	0.42	1.04	4	0.9911	42.52	1.14	0.38	0.83
	4.0	2	0.9935	56.83	1.22	0.37	1.24	3	0.9949	64.75	1.17	0.33	0.92
	8.0	1	0.9972	110.00	1.25	0.31	1.97	1	0.9979	116.52	1.20	0.24	1.45

Table 1: optimal control limits, availability and moments of system up- and down times

A general conclusion from table 1 is that a decreasing coefficient of variation of the corrective repair time distribution yields a higher value of the optimal availability and a higher preventive repair limit; preventive repair is postponed for a while, because the probability of occurrence of a very long corrective repair gets smaller. However, a decreasing coefficient of variation of the preventive repair time distribution again yields a higher value of the optimal availability, but a lower preventive repair limit; preventive repair is applied earlier, because excessively long preventive repair times will occur less frequently.

Another remarkable result is that c_u^2 increases in case of decreasing coefficient of variation of the preventive repair time distribution ($G_1(.)$).

Recall that an $Erlang(n,\lambda)$ distributed variable X has expectation $E_x = n/\lambda$, and $c_x^2 = 1/n$. Table 2 and table 3 concern repair times which are twice as short on average as in table 1. Note that the coefficients of variation of repair time distributions in table 2 are twice as large as they are in table 3. Thus from our conclusions concerning table 1 we would expect the optimal availability in table 3 to be higher than in table 2.

Comparison of the first lines of table 1, 2 and 3 reveals that dividing the mean repair times by a factor 2 gives rise to a multiplication of the mean time to system failure by a factor 3 to 4; the corresponding mean times to system repair differ considerably as well. These differences may have substantial consequences in case of high set up costs or in case finite buffers are used to cover system down times.

G_2		*Erl.* (1,2.0)						*Erl.* (2,4.0)					
G_1	μ_1	m^*	$1-g(m^*)$	E_u	c_u^2	E_d	c_d^2	m^*	$1-g(m^*)$	E_u	c_u^2	E_d	c_d^2
	2.2	7	0.9934	71.32	1.03	0.47	1.00	8	0.9964	92.57	1.02	0.34	0.87
	2.6	5	0.9952	86.13	1.04	0.42	1.03	7	0.9964	101.39	1.02	0.37	0.97
Erl. $(1,\mu_1)$	3.0	4	0.9964	103.39	1.05	0.38	1.08	7	0.9972	121.10	1.03	0.33	0.94
	4.0	3	0.9979	151.91	1.07	0.32	1.24	4	0.9984	172.95	1.04	0.28	0.98
	8.0	1	0.9992	256.60	1.08	0.21	2.12	2	0.9995	348.34	1.05	0.19	1.41
	2.2	5	0.9958	86.99	1.04	0.36	1.04	7	0.9969	102.22	1.02	0.32	0.87
	2.6	3	0.9970	101.95	1.06	0.31	1.08	5	0.9977	123.19	1.03	0.28	0.88
Erl. $(2,2\mu_1)$	3.0	3	0.9977	127.04	1.06	0.29	1.21	4	0.9982	145.49	1.04	0.26	0.91
	4.0	2	0.9985	168.28	1.07	0.25	1.49	3	0.9989	201.65	1.05	0.22	1.02
	8.0	1	0.9993	273.79	1.08	0.18	2.68	1	0.9995	294.03	1.05	0.13	1.72

Table 2: optimal control limits, availability and moments of system up- and down times

G_2		*Erl.* (2,4.0)						*Erl.* (4,8.0)					
G_1	μ_1	m^*	$1-g(m^*)$	E_u	c_u^2	E_d	c_d^2	m^*	$1-g(m^*)$	E_u	c_u^2	E_d	c_d^2
	2.2	7	0.9969	102.22	1.02	0.32	0.87	8	0.9978	119.98	1.02	0.27	0.73
	2.6	5	0.9977	123.19	1.03	0.28	0.88	7	0.9981	139.60	1.02	0.27	0.80
Erl. $(2,2\mu_1)$	3.0	4	0.9982	145.49	1.04	0.26	0.91	6	0.9985	164.04	1.03	0.24	0.80
	4.0	3	0.9989	201.65	1.05	0.22	1.02	4	0.9991	223.68	1.03	0.20	0.84
	8.0	1	0.9995	294.03	1.05	0.13	1.72	2	0.9997	397.84	1.04	0.14	1.28
	2.2	5	0.9977	114.67	1.03	0.27	0.81	7	0.9981	131.36	1.02	0.25	0.72
	2.6	4	0.9983	138.56	1.04	0.24	0.85	6	0.9985	157.36	1.03	0.23	0.73
Erl. $(4,4\mu_1)$	3.0	3	0.9986	157.27	1.04	0.22	0.88	5	0.9988	182.82	1.03	0.21	0.75
	4.0	3	0.9991	218.97	1.05	0.20	1.11	3	0.9993	232.86	1.04	0.17	0.82
	8.0	1	0.9996	302.10	1.05	0.12	2.02	2	0.9997	407.82	1.04	0.13	1.46

Table 3: optimal control limits, availability and moments of system up- and down times

In case of highly reliable components one often assumes the system up times to be exponentially distributed ($c_u^2 = 1$) as this assumption will simplify further analysis a lot (cf. section 5.2 and van der Heijden [5]). From table 1 we see that the system up times are not exponentially distributed in general ($c_u^2 \neq 1$). In table 2 and table 3 (repairs twice as fast) the coefficients of variation approach one quite closely, so the assumption seems to be justified in these cases. Relatively speaking one could say that in table 1 components are not reliable enough for the assumption to hold. Note that system down times are definitely non-exponential in general (cf. (4.1.1) and remark 4.2). The coefficient of variation of system down time turns out to be quite sensitive to changes in the preventive repair rates.

Although mean repair times are equal, comparison of table 2 with table 3 reveals that there are considerable differences between the corresponding availability characteristics of the system. So in applications one may often need more information about the repair times than just the mean. In table 4 below we present some numerical examples concerning different repair time distributions having equal first two moments. The repair time distribution functions are obtained by fitting the first two moments to an Erlang distribution (*Erl.*), a mixture of exponential and Erlang distributions with the same scale parameters ($E_{1,k}$), a hyperexponential distribution with the gamma normalization

(*GN*), and a hyperexponential distribution with balanced means (*BM*) respectively (Tijms [13], pp. 397-400). Repair time distributions are presented in table 4 by their mean (E_k), coefficient of variation (c_k^2), and skewness ($sk_k := E(X-E_x)^3 \ / \ \sigma_x^3$), $k = 1,2$.

fit	E_1	c_1^2	sk_1	E_2	c_2^2	sk_2	m^*	$1-g(m^*)$	E_u	c_u^2	E_d	c_d^2
Erl.	0.5	0.2	0.89	1.0	0.2	0.89	3	0.9954	68.73	1.17	0.31	0.87
$E_{1,k}$	0.5	0.2	0.58	1.0	0.2	0.58	3	0.9955	69.28	1.17	0.31	0.84
Erl.	0.5	0.5	1.41	1.0	0.5	1.41	3	0.9913	50.44	1.20	0.44	1.10
$E_{1,k}$	0.5	0.5	1.10	1.0	0.5	1.10	3	0.9917	50.72	1.20	0.43	1.04
GN	0.5	1.1	2.10	1.0	1.1	2.10	3	0.9806	35.19	1.20	0.70	1.26
BM	0.5	1.1	2.28	1.0	1.1	2.28	3	0.9805	35.69	1.19	0.71	1.33
$E_{1,k}$	0.5	1.1	1.25	1.0	1.1	1.25	3	0.9821	33.13	1.24	0.60	1.00
GN	0.5	2.0	2.83	1.0	2.0	2.83	3	0.9631	27.59	1.16	1.06	1.31
BM	0.5	2.0	3.89	1.0	2.0	3.89	4	0.9639	32.76	1.14	1.23	1.61
$E_{1,k}$	0.5	2.0	2.30	1.0	2.0	2.30	3	0.9614	25.63	1.14	1.03	1.02
GN	0.5	5.0	4.47	1.0	5.0	4.47	4	0.9214	26.74	1.05	2.28	1.34
BM	0.5	5.0	6.62	1.0	5.0	6.62	4	0.9325	35.24	1.07	2.55	1.82

Table 4: optimal control limits, availability and moments of up- and down times; repair time distributions obtained by two moments fit

The results in table 4 and further computational experiments indicate the first two moments of repair time distributions to provide sufficient information for practical purposes if coefficients of variation are not too large (in accordance with the results of van der Heijden [5]). We also note that the transient measures (E_u, c_u^2, E_d, c_d^2) are more sensitive to the distributional form of repair times than the steady state measure ($1-g(m^*)$).

References

[1] F. A. van der Duyn Schouten & T. Ronner (1989), "Calculation of the availability of a two-unit parallel system with cold standby: an illustration of the embedding technique", *Probability in the Engineering and Informational Sciences,* vol. 3, pp. 341-353.

[2] I. B. Gertsbakh (1977), *Models of preventive maintenance,* North-Holland Publishing Company, Amsterdam.

[3] I. B. Gertsbakh (1984), "Asymptotic methods in reliability theory: a review", *Advances in Applied Probability,* vol. 16, pp. 147-175.

[4] A. Goyal & A. N. Tantawi (1988), "A measure of guaranteed availability and its numerical evaluation", *IEEE Transactions on Computers,* vol. 37, no. 1, pp. 25-32.

[5] M. C. van der Heijden (1987), "Interval availability distribution for a 1-out-of-2 reliability system with repair", *Probability in the Engineering and Informational Sciences,* vol. 1, pp. 211-224.

[6] M. C. van der Heijden & A. Schornagel (1988), "Interval uneffectiveness distribution for a k-out-of-n multistate reliability system with repair", *European Journal of Operational Research,* vol. 36, pp. 66-77.

[7] H. Kawai (1981), "An optimal maintenance policy of a two-unit standby system", *Transactions of the IECE of Japan,* vol. E 64, no. 9, pp. 579-582.

[8] J. F. Meyer (1982), "Closed-form solutions of performability", *IEEE Trans. on Computers,* vol. C-31, no. 7, pp. 648-657.

[9] S.M. Ross (1985), *Introduction to probability models,* Academic Press, Inc., Orlando, Florida.

[10] C. F. H. van Rijn & A. Schornagel (1988), "STAMP - A new technique to assess the probability distribution of system effectiveness", *The SRS Quarterly Digest, supplement : Proceedings of the Seventeenth Annual Meeting of the Systems Reliability Service Associate Members 29-30 October 1987,* pp. 50-62.

[11] E. de Souza e Silva & H. R. Gail (1986), "Calculating cumulative operational time distributions of repairable computer systems", *IEEE Transactions on Computers,* vol. C-35, no. 4, pp. 322-332.

[12] L. Takács (1957), "On certain sojourn time problems in the theory of stochastic processes", *Acta Mathematica Academiae Scientiarum Hungaricae,* vol. 8, pp. 169-191.

[13] H. C. Tijms (1986), *Stochastic modelling and analysis: a computational approach,* Wiley, Chichester.