EINDHOVEN UNIVERSITY OF TECHNOLOGY

Reliability analysis of a repairable dependent parallel system

Citation for published version (APA):
Pijnenburg, M. J. P. (1992). *Reliability analysis of a repairable dependent parallel system*. [Phd Thesis 1 (Research TU/e / Graduation TU/e), Industrial Engineering and Innovation Sciences]. Technische Universiteit Eindhoven. https://doi.org/10.6100/IR381326

DOI:
10.6100/IR381326

Document status and date:
Published: 01/01/1992

Document Version:
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

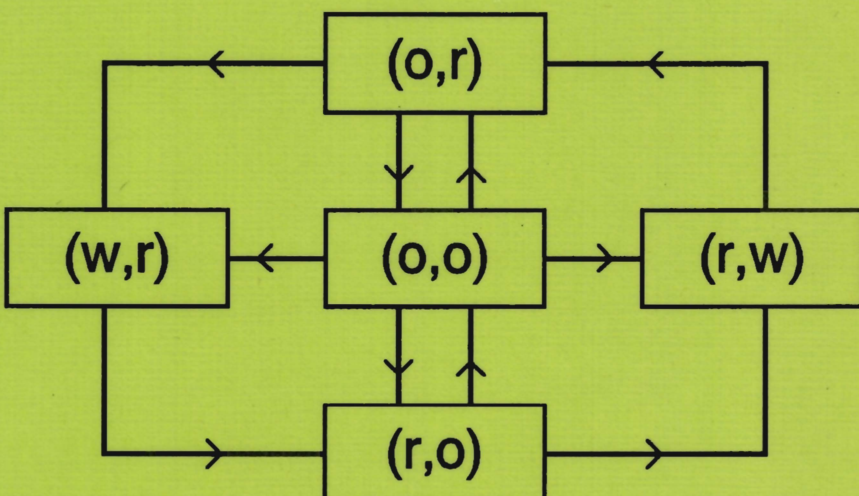
Take down policy
If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Download date: 16. Nov. 2023

Reliability Analysis of a Repairable Dependent Parallel System



Marc Pijnenburg

Reliability Analysis

of a

Repairable Dependent Parallel System

Reliability Analysis

of a

Repairable Dependent Parallel System

Proefschrift

ter verkrijging van de graad van doctor aan de Technische Universiteit Eindhoven, op gezag van de Rector Magnificus, prof. dr. J.H. van Lint, voor een commissie aangewezen door het College van Dekanen in het openbaar te verdedigen op dinsdag 22 september 1992 om 16.00 uur

door

Marcus Johannes Petrus Pijnenburg

geboren te Tilburg

Dit proefschrift is goedgekeurd door

de promotoren : prof. dr. P.C. Sander
prof. dr. J. Wessels

Ontwerp omslag : Fons Janssen

CONTENTS

1. INTRODUCTION

This monograph is a study of life and death. The study is carried out by developing a stochastic model to describe the life and death of a system. Although the term 'system', as well as the models considered, apply quite generally, the systems analysed are thought of as industrial machinery, equipment or products. Hence, a model is developed to describe the system's failure behaviour and, instead of life and death, the terms operating and down are used.

1.1 Motivation

Since breakdowns of a system increase costs and inconvenience or sometimes gravely threaten the public safety, the demand for systems that perform better and cost less increases. It is well known in the reliability field that providing redundancy, in part or all of a system, improves the performance of the system. The advantages of redundant systems include a reduction of the system down time and an enhancement of the reliability, within the technological constraints. For these reasons, much research has been reported on the analysis of redundant systems (Osaki *et al.*, 1976, Yearout *et al.*, 1986). Two basic redundancy configurations are parallel and standby. The analysis of a parallel redundant system with repairable units has an extensive literature. The fundamental and original contribution is due to Gaver (1963, 1964), who considered a two unit parallel redundant system with exponentially distributed life times and arbitrarily distributed repair times. Gaver used supplementary variables (Cox, 1955[a]) to derive the mean time to system failure and the stationary availability. Ever since this reported research there have been attempts to derive performance measures under relaxed assumptions on the life and repair time distributions of the units in the system. Some of the notable contributions are Liebowitz (1966), Kodama *et al.* (1974), Linton (1976), Subramanian *et al.* (1979), Ravichandran (1981) and Osaki (1985). Ravichandran (1990) reviewed the state of the art for a two–unit parallel system. However, an important role in a system's failure behaviour is played by factors such as maintenance, overhauls, the effects of repairs, dependence between units, intensity of use, stress situations, etc. Ascher and Feingold (1984) give a list of 18 'real world factors', which are hardly considered in existing models.

The present study concentrates on two of these real world factors: firstly, dependence between the units and secondly, maintenance. The subject of study is a two–unit parallel system, a basic redundancy configuration. However, in principle the methods and techniques also apply to standby redundancy or more complicated systems.

A frequently observed phenomenon which causes (statistical) dependence between units is the occurrence of common cause failures. Simple examples of common cause failures include situations where systems have shared (electric) connections or are subjected to common environmental stresses or shocks, etc. Nevertheless, dependence between units is ignored in the majority of reliability models. The main reason for this unsatisfactory state of affairs is that the assumption of independent units often considerably simplifies the analysis. However, Harris (1968) considered the situation of a two–unit parallel system with common cause failures and used a bivariate exponential distribution to model the life times of the units. He derived the mean time to system failure under the assumption of arbitrarily distributed repair times, using the supplementary variable technique. Osaki (1980) extended the analysis to obtain the availability of the system, using a variant of a semi–Markov process with some non–regeneration points. In the present study, not only the system reliability and availability are investigated, but also other important performance measures, well known in the reliability field, *viz.* the interval reliability, the joint availability, the system state probabilities and the stationary counterparts of these quantities. Apart from these performance measures, two quasi–stationary distributions are studied, namely the limiting residual life time distribution and the quasi–stationary system state probabilities. Both distributions are of particular interest when the system fails rarely or is not repairable. Quasi–stationarity is well known in the stochastic process literature (Darroch *et al.*, 1965, 1967, Seneta *et al.*, 1967, 1985, Cavender, 1978), but in the context of reliability modelling the use of quasi–stationary distributions is of recent origin. One of the few attempts that have been made in this direction is that of Kalpakam *et al.* (1983). The present study generalises and extends the results of Kalpakam *et al.*, in case of a two–unit dependent parallel system.

Most results in this study are obtained using phase type distributed life times (Neuts, 1981) and generally distributed repair times. The concept of phase type distributions is used since it allows a relatively simple analysis of the parallel system, which can be performed by studying an appropriate imbedded renewal process. Important operating characteristics follow from this imbedded renewal process in a direct and elegant way. Secondly, an important property of phase type distributions is that they are dense in the class of distribution functions and hence every distribution function can be approximated arbitrarily close by a phase type distribution.
A further extension of the analysis of the parallel system is obtained in the direction of maintenance. The importance of maintenance is beyond dispute. Consequently, the literature on maintenance and replacement models is extensive (see Pierskalla *et al.*, 1976, Sherif *et al.*, 1981, Valdez–Flores *et al.*, 1989). With respect to a two–unit parallel system, an important facet is that it can be advantageous to replace both units simultaneously when the system shows economic dependence, *i.e.* when the costs of a joint maintenance action are less than the costs of two separate maintenance actions. However, the overwhelming majority of papers deals with single–unit systems and hence the presence of economies of scale is not considered. Among the authors who investigate maintenance policies for two–unit systems, are Mine *et al.* (1974) and Berg (1978). Mine *et al.* consider a maintenance model for a two–unit parallel system under Markovian deterioration, including repair times in the analysis. Berg analyses an opportunistic replacement policy for a two–unit *series* system with arbitrarily distributed life times but ignores maintenance times. In the present study elements of the work of Berg and Mine *et al.* are combined and Berg's continuous time, opportunistic replacement policy is extended to a dependent parallel system, including the repair times in the analysis. Both the life and repair times are arbitrarily distributed and an optimisation problem is formulated in order to minimise the expected costs per unit of time.
Summarising the discussion, it appears that there are few studies of systems with dependent units and few studies on replacement policies for parallel systems: it is evident that a comprehensive and systematic analysis is necessary. This study attempts to provide such a systematic analysis.

1.2 The model

A system is defined here as a collection of one or more interconnected units, designed to perform one or more specified functions. A unit is a part of the system and is repaired or replaced every time it fails. The units under consideration are two-state units: they either work or fail. On the other hand, the system under consideration is a two-unit system and the units are connected in parallel. A system failure is defined as a state of the system in which it is not able to perform at least one of its functions satisfactorily. The system is repairable, *i.e.* after failure, it can be restored to fully satisfactory performance by *any* method other than replacement of the entire system.

A standard way to represent a two-unit parallel system with *dependent* units is to look at the life times of three components, say C_1, C_2 and C_3. The components C_1 and C_2 are connected in parallel and C_3 is an artificial component in series with the parallel configuration of C_1 and C_2, as in figure 1.1. Component C_3 is used to model the occurrences of common cause failures and in that way the dependence in the two-unit system. The components C_1 and C_2 form a two-unit parallel system with *independent* components. The main reason to decompose a two-unit dependent parallel system into a two-unit independent parallel system and a common cause component is that a decomposition allows a neat mathematical analysis. In the special case where the life times of the components are exponentially distributed, the technique is known as the β-factor technique (Lewis, 1987).

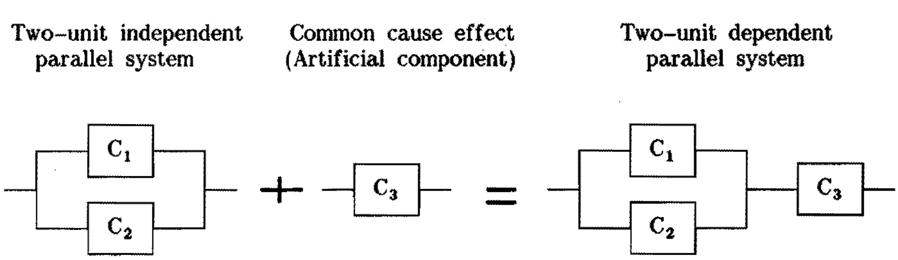


Fig. 1.1: Modelling a two-unit dependent parallel system

Note the terminology: the physical system consists of *two units* and is described by a model with ***three logical components***.

With respect to the model in figure 1.1, the following basic assumptions are made to perform the mathematical analysis in the subsequent chapters:

i. The two–unit parallel system requires at least one unit for successful operation.

ii. The units are repairable. On failure a unit is repaired by a single server repair facility with first–in–first–out repair policy. Repairs are perfect and restore the normal operational efficiency of the units.

iii. The identification of the operable and non–operable status of a unit is perfect. After failure, a repair is started immediately. Similarly, after repair completion a unit restarts operating immediately.

iv. For each component only one type of failure occurs. A failure of component C_i (i=1,2) destroys unit *i* and a failure of component C_3 destroys both units simultaneously.

v. The common cause effect is only present when the system is *up*. This means that when the system is down, caused by a failure of C_1 during a repair of C_2 (or *vice versa*), the life time of C_3 is ended. Subsequently, C_3 restarts operating immediately after repair completion of either C_1 or C_2, whichever occurs first. On the other hand, the life times of both C_1 and C_2 end on the occurrence of a common cause failure, *i.e.* with the failure of C_3.

vi. Since the common cause component C_3 is an artificial component, its 'repair' is assumed to be instantaneous: the repair time of C_3 is zero, in contrast to the repair times of C_1 and C_2. In case of a common cause failure the units queue for repair at random.

vii. At time t=0 two new units start operating.

Dependent upon the nature of the common cause effect, often two types are distinguished: internal and external common cause failures. An external effect is caused by the environment and usually modelled by a homogeneous Poisson process. Simple examples are failures caused by common electric connections, by fire or by vibration. On the other hand, internal common cause failures occur when *e.g.* the failure of one unit results in a fatal shock for the other

unit. Although it is possible to model internal and external effects separately by connecting an additional component in series with C_3, for convenience both types of effects are lumped together here and modelled by one artificial component.
In the next section a mathematical description of the parallel system is given in terms of a stochastic process.

1.3 The state description process

In order to characterise the performance measures of the parallel system, a stochastic process is defined which describes the state of the system. The size of the state space of this stochastic process is determined by the kind of information wanted about the system and the mathematical techniques used to analyse the system's failure behaviour. As illustrated in this section, the more detailed the information wanted, the more detailed the system's state space will be. For example, when the only object of interest is the reliability or availability of the system, a two-valued state description process suffices: in section 1.4 it is shown that expressions for the system reliability and availability can be translated into equivalent questions about the state description process $\{X(t), t \geq 0\}$, where

$$X(t) = \begin{cases} 1, & \text{if the system is operating at time } t, \\ 0, & \text{if the system is down at time } t. \end{cases}$$

Since transitions occur only from state 0 to 1 and *vice versa*, figure 1.2 shows the one-step transition diagram of the process X(t).

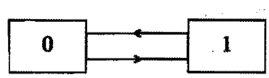


Fig. 1.2: One-step transition diagram of the process X(t)

However, when the parallel system is operating, an interesting question concerns the number of units which is operating. Hence, if one's interest is not only at system level but also at component level, a more detailed state description process is needed.

The three-valued process $\{Y(t), t \geq 0\}$, where

$$Y(t) = \begin{cases} 0, & \text{if both units are operating at time } t, \\ 1, & \text{if one unit is up and one unit is down at time } t, \\ 2, & \text{if both units are down at time } t, \end{cases}$$

can provide information at component level. Figure 1.3 shows the one-step transition diagram. As common cause failures occur, transitions are possible from state 0 to state 2. A stay in state 2 is always followed by a transition to state 1.

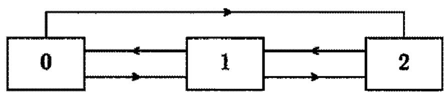


Fig. 1.3: One-step transition diagram of the process Y(t)

When the units are not identical, not only the number but also the identity of the operating units is of interest. Using the symbols 'o' for an operating unit, 'r' for a unit under repair and 'w' for a unit which is waiting for repair (the units are repaired by a single server facility), the system state space S_z is given by the set of ordered pairs {(o,o), (o,r), (r,o), (r,w,), (w,r)}. Hence, in this case a five-valued state description process $\{Z(t), t \geq 0\}$ is defined, where

$$Z(t) = \begin{cases} (o,o), & \text{if both units are up at time } t \\ (o,r), & \text{if unit 1 is operating and unit 2 under repair at time } t \\ (r,o), & \text{if unit 1 is under repair and unit 2 operating at time } t \\ (w,r), & \text{if unit 1 is queued and unit 2 under repair at time } t \\ (r,w), & \text{if unit 1 is under repair and unit 2 queued at time } t. \end{cases}$$

The accompanying transition diagram is shown in figure 1.4. Note that it is supposed that transitions from (o,r) to (r,o) do not occur.

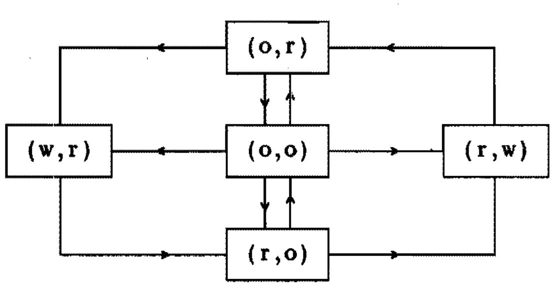


Fig. 1.4: One-step transition diagram of the process Z(t)

The transition diagrams in figure 1.3 and 1.4 are fundamental for this study, since they provide the basic transition mechanism of the dependent parallel system. However, the mathematical techniques in the following chapters necessitate a further extension of the state space S_Y and S_Z, when the life or repair time distributions are assumed to be of phase type (Neuts, 1981). In this case the states in figure 1.3 and 1.4 are extended to include the numbers of the phases of the life respectively repair time distribution.

In section 1.4 it is illustrated how the performance measures are related to the system's state description process. Subsequently, it will be shown in the next chapters that study of an appropriate state description process yields the system's operating characteristics and performance measures explicitly.

1.4 The performance measures

As mentioned in section 1.3, expressions for the system's performance measures can be translated into equivalent questions about the state description process. The characterisation of the performance measures is illustrated by the simple example of a system with two states, *viz.* operating and down. The generalisation to a system with a larger state space is straightforward.

Let $\{X(t),t\geq 0\}$ be the system's state description process, *i.e.*

$$X(t) = \begin{cases} 1, & \text{if the system is operating at time } t \\ 0, & \text{if the system is down at time } t \end{cases}$$

and assume that at time $t=0$ the system is new and put into operation.

The system reliability $R(t)$ is defined as the probability that the system is operating during the interval $(0,t)$:

$$R(t) = \Pr\{X(s)=1,\ 0\leq s\leq t\}.$$

Similarly, the interval reliability $R(t,\tau)$ is defined as the probability the system is operating satisfactorily during the interval $(t,t+\tau)$:

$$R(t,\tau) = \Pr\{X(s)=1,\ t\leq s\leq t+\tau\}.$$

Obviously, $R(t) = R(0,t)$.

Another fundamental quantity of interest is the system point availability A(t), defined as the probability the system is performing satisfactorily at time t:

$$A(t) = \Pr\{X(t)=1\}.$$

The joint availability $A(t,\tau)$ is the probability the system is operating at t and $t+\tau$:

$$A(t,\tau) = \Pr\{X(t)=1 \wedge X(t+\tau)=1\}.$$

Just as reliability and interval reliability are related by $R(t) = R(0,t)$, the availability and joint availability satisfy $A(t) = A(0,t)$.
Further, the time dependent state probabilities $P_i(t)$ are given by the probability the system is in state i at time t:

$$P_i(t) = \Pr\{X(t)=i\}, \; i=0,1.$$

In the simple case with only two states, $P_1(t)$ equals the system point availability at time t and $P_0(t)=1-P_1(t)$.
While the above performance measures are time dependent, their time independent counterparts are of particular interest if the stochastic process $\{X(t),t\geq 0\}$ is stationary or transient, *i.e.* if the process reaches (after some initial effects) a steady state or equilibrium state. Formal definitions are given by Thompson (1988). In the steady state, characteristic quantities are the mean time to system failure (MTSF), the limiting or asymptotic availability A and the stationary state probabilities π_i (i=0,1). Let T_n $(n\geq 1)$ be the length of the n^{th} stay in state 1, then formal definitions are given by

$$MTSF = \lim_{n\to\infty} E\, T_n,$$

$$A = \lim_{t\to\infty} A(t)$$

and

$$\pi_i = \lim_{t\to\infty} P_i(t),$$

assuming that the above limits exist.
If a system fails rarely or is nonrepairable, a special point of interest is formed by two quasi-stationary distributions, *viz.* the quasi-stationary state probabilities and the limiting residual lifetime distribution. The quasi-

stationary state probabilities q_i give the limiting probability as $t \to \infty$ of being in state i, under the condition that no system failures have occurred until time t. Formally,

$$q_i = \lim_{t\to\infty} \Pr\{X(t)=i \mid X(s)\neq 0,\ 0\leq s\leq t\},$$

assumed that the limit exists. Note that, for a pure *two-state* system it is not meaningful to study the quasi-stationary state probabilities, as it is easily seen that $q_0=0$ and $q_1=1$ in this case.

The second quasi-stationary distribution under consideration is the limiting residual life time distribution, given that no system failure has occurred. Let q(.) denote the limiting residual life time distribution, then formally

$$q(x) = 1 - \lim_{t\to\infty} \Pr\{X(s)=1,\ t\leq s\leq t+x \mid X(s)=1,\ 0\leq s\leq t\}.$$

In this study explicit analytic expressions are sought for the above quantities and performance measures and algorithmic forms are developed to allow numerical implementation on a computer. It is assumed that the life and repair time distributions are *known*. In principle the life and repair times may be arbitrarily distributed but, in order to get manageable formulas, the concept of phase type distributions (Neuts, 1981) is used throughout this monograph.

1.5 The techniques

The techniques used to analyse the model depend heavily upon the assumptions made with respect to the components' life and repair time distributions. In principle the aim is to handle a system with generally distributed life and repair times. For a two-unit parallel system with independent units Ohashi *et al.* (1980) obtained expressions for the system reliability R(t) and Subramanian *et al.* (1979) derived expressions for the availability A(t), under arbitrarily distributed life and repair times. In both papers the supplementary variable technique (Cox, 1955[a]) is used to derive formulas for the performance measures. The supplementary variable technique provides a mechanism to convert a non-Markovian to a Markovian process, by including some information in additional variables. These variables are called supplementary variables and usually contain information about the past or the history of the process, *e.g.* the length of time the system or components have been in a particular state. In the present study the supplementary variable technique is used to analyse the system's failure behaviour under phase type distributed life times and generally distributed repair times. The special properties of phase type distributions are exploited to derive expressions not only for the reliability and availability, but also for the interval reliability, joint availability, the system state probabilities, the stationary counterparts of these performance measures and two quasi-stationary distributions. The most important property, used in the analysis, is that the duration of a stay in a particular phase of a phase type distribution is exponentially distributed and hence the lack-of-memory property of the exponential distribution applies locally, *i.e.* per phase. Moreover, phase type distributions are dense in the class of all distribution functions and hence every distribution function can be approximated arbitrarily close by a phase type distribution.

In general, the life times of the components C_i (i=1,2,3) in figure 1.1 are taken to follow a phase type distribution and the repair times of C_i (i=1,2) are arbitrarily distributed. However, two cases are distinguished, dependent upon the distribution of the repair times.

In the first case both life and repair times have a phase type distribution. It is shown that including the phases of the life and repair time distributions in the state space S_Z of the process $\{Z(t),t\geq 0\}$ (as described in section 1.3) renders the system Markovian. Hence, an ordinary Markov process is created and standard Markov theory can be applied to derive expressions for the performance measures. It appears that questions about the performance measures can be translated into eigenvalue problems for the generator of the Markov process.

In the second case, the life times are supposed to follow a phase type distribution, while the repair times are generally distributed. Again, the state space S_Z of the process $\{Z(t),t\geq 0\}$ is extended to include the phases of the life time distributions. Although this extension of S_Z does not render the state description process Markovian, it appears that, as a consequence of the local lack-of-memory property, there exist states $\mathbf{s}$ with the property that the evolution of the state description process *after* an entry into state $\mathbf{s}$ is independent of the history *until* the entry in $\mathbf{s}$. In other words: an entry into state $\mathbf{s}$ is a regeneration point and the underlying stochastic process a regenerative stochastic process. Consequently, the interarrival times between successive entries in state s form an imbedded renewal process. The regenerative nature of the entries, as well as the imbedded renewal process, play an important role in the analysis and are frequently exploited to obtain equations for performance measures. In fact, the phases of the life times are used as a discrete supplementary variable and formulas for the performance measures are derived, conditioned on an entry into a regenerative state. The result is a set of recurrence relations for the reliability functions, the availability functions and the other time dependent measures, conditioned on an entry into a regenerative state. More specifically, the equations are in terms of sets of coupled (convolution) integrals. Hence, taking the Laplace transforms yields a set of linear equations in the Laplace transforms of the reliability functions, the availability functions, etc. Proceeding in vector–matrix notation, the sets of equations is written in the form $Ax = b$, where A is a matrix and b and x vectors. It appears that the matrix A plays the role of the generator in a Markov process and questions about performance measures can be answered by investigating the matrix A.

1.6 The chapters

Commonly used models for the analysis of repairable systems form the subject of chapter 2. A brief review is given of some basic stochastic models and processes. Actually, two categories are considered: probabilistic and regression models. Basic probabilistic models are the homogeneous and nonhomogeneous Poisson process (HPP respectively NHPP), the Markov process (MP) and semi–Markov process (SMP), the renewal and superimposed renewal process (RP respectively SRP), the branching Poisson and branching renewal process (BPP respectively BRP) and the regenerative stochastic process (RSP).
The HPP, MP, RP, SRP and RSP are of special interest here, as elements of these stochastic processes provide the basic techniques throughout this study. Subsequently, three regression models are described, *viz.* the accelerated failure time model (AFTM), the proportional hazards model (PHM) and the additive hazards model (AHM). The ability of regression models to include explanatory variables in the analysis, seems a powerful tool to capture many real world factors. The model, described in section 1.2, appears to be a special case of an AHM, a model about which is very little known in the

literature. There are few references to the AHM and for this reason it is given special attention in chapter 2.
From chapter 3 on, the analysis focuses on a specific system configuration, *viz.* a two-unit dependent parallel system. The joint life time distribution is a bivariate exponential (BVE) distribution and the repair times are arbitrarily distributed. Formulas are derived for the performance measures mentioned in section 1.4, using the theory of regenerative stochastic processes and imbedded renewal processes.
Chapter 4 treats the special case where the stochastic process under investigation becomes Markovian. Starting with BVE life times and exponential repair times, an extension is made to phase type distributed repair times and finally the more general situation is considered with bivariate phase type life times and PH distributed repair times. Standard Markov theory is applied to derive the performance measures.
A further generalisation is made in chapter 5, by studying a system with bivariate phase type life times and arbitrarily distributed repair times.
In chapter 6 an opportunistic replacement policy for the two-unit dependent parallel system is described, in continuous time and at component level. When a unit fails or when its life time exceeds a control limit, it is replaced and this opportunity is used for a possible replacement of the other unit. The stationary joint probability density function of the stochastic process is used to derive a number of operating characteristics. Numerical examples illustrate the techniques.
Finally, chapter 7 concentrates on the gradual deterioration of the parallel system. If the system deteriorates slowly and does not reach the steady state during the period of operation, only the transient, time dependent behaviour is of importance. A method is studied to model gradual deterioration, making the assumption that both units are imbedded in a larger system, called the system body. The system body, in the model represented by one more component, gradually deteriorates. Expressions for the performance measures are derived for arbitrarily distributed life times of the system body component. In this case, introduction of phase type distributions yields a Markov process with a (possibly) very large state space. However, using a special technique from Markov theory, it is still possible to study the transient behaviour.

2. BASIC STOCHASTIC MODELS IN RELIABILITY THEORY

2.1 Introduction

Until the late 1970's, early 1980's relatively little work had been done on repairable systems reliability. The emphasis in the reliability literature had been on nonrepairable systems in spite of the fact that in practice, repairable systems are much more common. Reasons for this unsatisfactory state of affairs are outlined in Ascher (1983) and Ascher *et al.* (1984).

The commonest models of the failure behaviour of both nonrepairable systems and, as a first order approximation, repairable systems are the renewal process (RP) and (as a special case of an RP) the homogeneous Poisson process (HPP). Both models are used because of their mathematical tractability and the fact that other models for repairable systems were ignored or simply overlooked. Apart from the HPP and RP, the nonhomogeneous Poisson process, the superimposed renewal process, the Markov and semi-Markov process, the branching Poisson and branching renewal process are basic stochastic models in reliability theory. These models are not described here at great length. However, as important aspects of the models are used throughout the analysis of the two-unit dependent parallel system, a short description is given in this chapter and a few basic properties are mentioned. For more details the reader is referred to Parzen (1962), Cox *et al.* (1965, 1966, 1984), Çinlar (1975) and Thompson (1981, 1988).

Whereas the above basic models consider the operating time as the only variable of interest, an important role in a system's failure behaviour is often played by factors such as maintenance, overhauls, the effects of repairs, dependence between components, intensity of use, stress situations, etc. Ascher *et al.* (1984) give a list containing 18 'real world factors' which are usually not considered in probabilistic models. To include explanatory and causal factors in a model, the use of regression models is appropriate. Well known regression models are the accelerated failure time model (AFTM), Cox's (1972) proportional hazards model (PHM) and the model introduced by Prentice *et al.* (1981). The latter model (referred to as the PWPM) is an extension of Cox's PHM to the case in which multiple failures of a system can occur and is therefore useful in the repairable systems reliability field. In the sections 2.8 and 2.9, the AFTM, PHM and PWPM are described and connections between basic probabilistic models and the PWPM are shown.

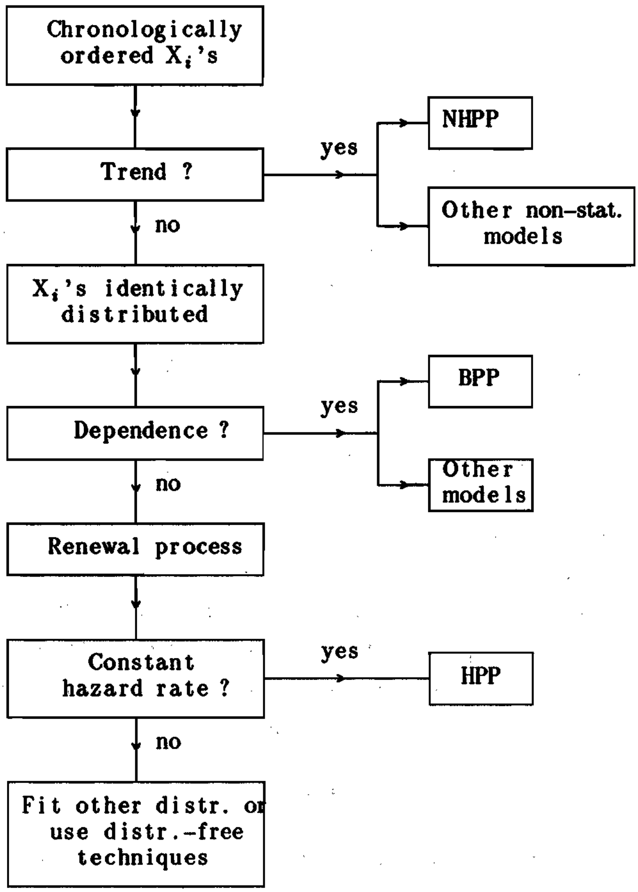
In the PHM the explanatory variables or covariates are assumed to act in a multiplicative way on a baseline hazard function. Nevertheless, it is just as plausible to suppose that the covariates act in an additive way. This kind of model is called an additive hazards model (AHM). Compared with the PHM very little is known about the AHM and only a few texts include references to it. As the model used to analyse the dependent parallel system is a special case of an additive hazards model, the AHM is the subject of section 2.10.

2.2 Repairable vs nonrepairable systems

In section 1.2 a repairable system is defined as a system which can be restored to perform all of its required functions by any method other than replacement of the entire system. Conversely, a nonrepairable system is discarded after its first failure at system level. Hence the fundamental difference is that repairable systems can fail many times, whereas a nonrepairable systems can fail no more than once.

Intuitively, a repair does not renew the system, *i.e.* being repaired a system is not 'as good as new'. However, a repair usually increases the system's reliability, *e.g.* because of the replacement of one or more components. In other words: it is plausible that a repair does not return the system to a 'bad as old' situation. As a result the life times of the system are not necessarily identically distributed. The consequences for a life time data analysis are the following (for the moment the repair times are ignored).

Let the random variable X_i (i=1,2,..) denote the length of the system's i^{th} life time. Focusing on deterioration, the life times X_i are expected to become smaller (alternatively, the repair times can become longer). Hence, given the failure time data, the first thing to do is to check if there is a tendency for successive life times to shorten: a trend test is needed. Several trend tests are described in Ascher *et al.* (1978) and Cox *et al.* (1966). Ascher *et al.* (1984) discuss the difficulties in finding a mathematical satisfying definition of improvement or deterioration in terms of times between failures. If there are no indications of a trend, the random variables X_i are usually supposed to be stationary. If there is also no evidence that the life times X_i are statistically dependent, it is reasonable to accept them to be i.i.d. In the latter case they can be modelled by a renewal process. Several tests for dependence are treated in Cox *et al.* (1966). Figure 2.1 summarises how a statistical analysis of the life times X_i of a repairable system should be performed. A similar diagram, including the analysis of ***repair time data***, is presented by Walls *et al.* (1985). Both diagrams show which model is acceptable under which conditions and are helpful in an exploratory data analysis.



(Source: Ascher et al., 1984)

Fig. 2.1: Analysis of the life times X_i of a repairable system

In the following sections a survey of some important basic probabilistic and regression models is presented.

2.3 The homogeneous and nonhomogeneous Poisson process

Several equivalent definitions of a homogeneous Poisson process (HPP) are found in the literature (Çinlar, 1975, Cox *et al.*, 1965). However, the most straightforward way to define an HPP is as a counting process, generated by a sequence of independent and identically exponentially distributed random variables $\{X_i\}_{i\in N}$. In the context of repairable systems, the random variables X_i are called times between failures or interarrival times. Let the sequence $\{T_i\}_{i\in N}$ represent the epochs at which failures occur and assume that repair times can be neglected. Then $X_i=T_i-T_{i-1}$, i=1,2,.., where by definition $T_0=0$.
Çinlar (1975) proves that the HPP is the unique process with stationary and independent increments and no simultaneous failures. To show how the nonhomogeneous Poisson process (NHPP) is a direct generalisation of the HPP, the following alternative definition is given.
Let $\{N(t),t\geq 0\}$ be a counting process and let $N(t,t+\Delta)$ denote the number of events in $(t,t+\Delta]$. Then the process $N(t)$ is an HPP if it has independent increments and

$$\Pr\{N(t,t+\Delta)=n\} = \frac{(\lambda\Delta)^n}{n!}\, e^{-\lambda\Delta}, \quad n=0,1,2,.. \tag{2.1}$$

where λ is a positive constant.
As the interarrival times X_i are exponentially distributed with parameter λ, they have a constant hazard rate λ. The constant hazard rate or, equivalently, the lack-of-memory property of the exponential distribution, is often unrealistic as it cannot model wear-out or burn-in of a component (Ascher *et al.*, 1984). From (2.1) it follows that $\mathbf{E}\, N(t,t+\Delta) = \lambda\Delta$ and hence

$$\mathbf{E}\, N(0,t) = \lambda t,$$

i.e. the rate of occurrence of failures (ROCOF) of an HPP is also λ. Thus the HPP cannot model the deterioration of or the reliability growth in a system. In spite of these restrictions, the HPP is often used to model the failure behaviour of a repairable system because of its mathematical tractability.
The NHPP (Thompson, 1981) is a direct generalisation of the HPP. The only difference from the HPP is that the ROCOF $\lambda(.)$ is time dependent rather than being a constant.

Hence, the counting process $\{N(t),t\geq 0\}$ is an NHPP if $N(t)$ has independent increments and

$$\Pr\{N(t,t+\Delta)=n\} = \frac{\left[\int_t^{t+\Delta} \lambda(s)\, ds\right]^n}{n\,!} \exp\left[-\int_t^{t+\Delta} \lambda(s)\, ds\right]. \qquad (2.2)$$

From (2.2),

$$\mathbf{E}\, N(t,t+\Delta) = \int_t^{t+\Delta} \lambda(s)\, ds.$$

The minor change in definition leads to a major difference between the HPP and NHPP as under the latter model the interarrival times are *neither* independent *nor* identically distributed.

Two important properties of an NHPP are

$$T_{N(t)+1} - t \text{ is independent of the history up to time } t \qquad (2.3)$$

and

$$\Pr\{X_{n+1}>x \mid T_1=t_1,..,T_n=t_n\} = \Pr\{X_1>t_n+x \mid X_1>t_n\}. \qquad (2.4)$$

Statement (2.3) expresses the lack-of-memory property: the forward recurrence time is independent of the history up to time t. Expression (2.4) states that the reliability is not changed by failure and repair. In fact the left and right hand sides are the system reliability with and without failure. The properties (2.3) and (2.4) make that the NHPP is appropriate to model a bad-as-old situation. The bad-as-old assumption is plausible *e.g.* when a system consists of a large number of components and only a few are replaced at repair.

Thompson (1988) shows that the ROCOF of the process and the hazard rate of X_1, the time to first failure, are numerically equal. Moreover, let $h_n(.)$ be the hazard rate of X_n and suppose that failure n-1 occurs at t_{n-1}. Then $h_n(x)$ is numerically equal to the ROCOF of the process in $t_{n-1}+x$. In this way the NHPP is a natural development from the use of a hazard function for nonrepairable systems. The NHPP is often used to model a trend and is mathematically tractable.

2.4 The Markov and semi-Markov process

Let $\{X(t),t\geq 0\}$ be a stochastic process on a discrete state space $S=\{1,..,n\}$. Further, let $\{p_i\}_{i\in S}$ be a probability distribution on S and $\{P(t),\ t\geq 0\}$ a Markov semi-group on S, *i.e.*

i. P(t) is a Markov matrix on S for all $t>0$, with elements $p_{ij}(t)$, $i\in S$, $j\in S$

ii. $P(0) = I$,

iii. $P(t+s) = P(t)\ P(s)$, for all $s,t>0$.

Then $\{X(t),t\geq 0\}$ is a (time homogeneous) Markov process (Parzen, 1962, Cox *et al.*, 1965) if for all sequences $0=t_0<t_1<..<t_n$, $s_0,s_1,..,s_n\in S$ and for all $n\in\mathbb{N}$

$$\Pr\{X(t_0)=s_0,\ X(t_1)=s_1,\ ..\ ,\ X(t_n)=s_n\} = p_{s_0}\prod_{k=1}^{n} p_{s_{k-1}s_k}(t_k-t_{k-1}).$$

Since P(t) is a Markov semi-group,

$$P(t) = P(s)\ P(t-s). \tag{2.5}$$

Equation (2.5) is the Chapman-Kolmogorov equation for a time homogeneous Markov process (Çinlar, 1975). The behaviour of the process X(t) is completely specified by the functions $p_{ij}(t)$, which are obtained from the knowledge of the (constant) transition rates q_{ij} over short time intervals. Assuming

$$p_{ij}(t) = q_{ij}\ t + o(t),\ t\downarrow 0,\ j\neq i$$

$$p_{ii}(t) = 1 + q_{ii}\ t + o(t),\ t\downarrow 0 \tag{2.6}$$

$$q_{ij}\geq 0,\ j\neq i$$

and

$$q_{ii} = -\sum_{j\neq i} q_{ij},$$

a set of equations can be derived which determine the functions $p_{ij}(t)$. The matrix Q of q_{ij}'s is known as the generator of the Markov process and from (2.5) it can be shown (Çinlar, 1975) that

$$\frac{dP(t)}{dt} = P(t)\ Q \tag{2.7}$$

and

$$\frac{dP(t)}{dt} = Q\ P(t). \tag{2.8}$$

Equations (2.7) and (2.8) are respectively the forward and backward Kolmogorov differential equations. With initial condition $P(0)=I$, they have the solution

$$P(t) = \exp(Qt).$$

A description of a time homogeneous Markov process in terms of the interarrival times X_i is as follows. From (2.6), the sojourn time in a transient state is exponentially distributed with parameter $-q_{ii}$. Given the process leaves state i, a transition to state j occurs with probability $-q_{ij}/q_{ii}$, $i\neq j$. Hence, the interarrival times are exponentially distributed and the transition mechanism is provided by a Markov matrix P with entries

$$P_{ij} = -\frac{q_{ij}}{q_{ii}},\ j\neq i$$

$$P_{ii} = 0.$$

The semi-Markov process (Ross, 1970) links the theory of renewal processes and Markov chains. Again, the state transition mechanism is captured in a Markov matrix P. However, the sojourn times are not necessarily exponentially distributed: if the time spent in state i is followed by a transition to state j, the sojourn time in state i has distribution function $F_{ij}(.)$. Formally, the state description process generated by the sequence $\{X_i\}_{i\in N}$ (of times between successive transitions) forms an SMP if for all $t_0<t_1<..<t_n$, $s_0,s_1,..,s_n\in S$ and $n\in\mathbb{N}$

$$\Pr\{X(t_0)=s_0,\ X_1\leq x_1,\ X(t_1)=s_1,\ X_2\leq x_2,\ ..\ ,\ X(t_{n-1})=s_{n-1},\ X_n\leq x_n,\ X(t_n)=s_n\}$$

$$= p_{s_0}\prod_{k=1}^{n} P_{s_{k-1}s_k}\ F_{s_{k-1}s_k}(x_k),$$

where $x_k=t_k-t_{k-1}$. The sequence $\{s_i\}_{i\in N}$ is called the imbedded Markov chain.

2.5 The renewal and superimposed renewal process

The renewal process (Parzen, 1962, Çinlar, 1975, Thompson, 1988) generalises the HPP by allowing the interarrival times to have an arbitrary, but identical distribution: the sequence $\{X_i\}_{i\in N}$ forms an RP if the X_i's are non-negative, independent and identically distributed with distribution function $F(.)$, which satisfies $F(0)<1$. Hence an RP model for a repairable system postulates that

repairs return the system to a good-as-new situation. This assumption is reasonable *e.g.* when repair consists of replacement of the system.
The ROCOF of an RP is better known as the renewal density. In order to compute the ROCOF, consider $\{N(t), t \geq 0\}$, the counting process of renewals. From

$$N(t) = n \iff T_n \leq t \wedge T_{n+1} > t$$

it follows that

$$H(t) := \mathbb{E}\, N(t) = \sum_{n=1}^{\infty} F^{(n)}(t),$$

where $F^{(n)}(t)$ denotes the n-fold convolution of $F(t)$. The function $H(t)$ is called the renewal function and $h(t) = dH(t)/dt$ the renewal density or ROCOF: $h(t)\, dt$ represents the absolute probability of a renewal in $(t, t+dt)$, $dt \downarrow 0$. Of special interest in this study is the renewal equation:

$$h(t) = f(t) + \int_0^t f(u)\, h(t-u)\, du, \tag{2.9}$$

where $f(t) = dF(t)/dt$.
Expression (2.9) is derived by considering the mutually exclusive and exhaustive events that the renewal at time t is the first or not:

i. With probability $f(t)\, dt$ the first renewal takes place in $(t, t+dt)$.

ii. Given that the first renewal occurred at time u, with probability $h(t-u)\, du$ another renewal (not necessarily the next) occurs approximately $t-u$ units of time later, *i.e.* in $(t, t+du)$. Hence, multiplied by dt the second term in the right hand side of (2.9) represents the probability that the renewal in $(t, t+dt)$ is not the first.

Alternatively, writing

$$h(t-u) = \sum_{n=1}^{\infty} f^{(n)}(t-u)$$

(2.9) is immediate.
Applying Laplace transform techniques to (2.9) gives

$$h^*(s) = \frac{f^*(s)}{1 - f^*(s)}. \tag{2.10}$$

In principle, inversion of (2.10) yields an explicit expression for $h(t)$.

A very useful and famous result in renewal theory is the key renewal theorem: if the distribution function F(.) (with mean μ) is non-lattice and g(x) is directly Riemann integrable, then

$$\lim_{t\to\infty} \int_0^t g(t-x)\, dH(x) = \frac{1}{\mu} \int_0^\infty g(x)\, dx. \tag{2.11}$$

A superimposed renewal process (SRP) can be used to model a system at component level. It is appropriate when a system consists of n components connected in series. When the components are replaced at failure and failures occur independently of one another, the components can be modelled by RP's. The superposition of these RP's is called an SRP. If $\{N_i(t),t\geq 0\}$ is the counting process of failures of component i, then $N(t) = \sum N_i(t)$ gives the total number of system failures. Writing $M(t) = E\ N(t)$ and $M_i(t) = E\ N_i(t)$, it is obvious that the system's ROCOF $dM(t)/dt$ equals $\sum dM_i(t)/dt$. Further, the hazard rate of the time to first failure $\lambda(t)$ equals the sum of the hazard rates $\lambda_i(t)$ of the components: $\lambda(t) = \sum \lambda_i(t)$.

A detailed analysis of SRP's is performed in Drenick (1960), Grigelionis (1964), Blumenthal *et al.* (1971, 1973) and Barlow *et al.* (1975, the discussion is based on the work of Grigelionis). However, most results have been obtained in the limiting case where the number of components or the operating time approaches infinity. Moreover, the fundamental results of Grigelionis are obtained under quite unrealistic assumptions, *viz.*

$$\lim_{n\to\infty} \min_{i\in\{1,..,n\}} \Pr\{N_i(t)=0\} = 1$$

and

$$\lim_{n\to\infty} \sum_{i=1}^{n} \Pr\{N_i(t)\geq 2\} = 0,$$

i.e. a failure of any component is unlikely and two or more failures of any component are unlikely.

2.6 The branching Poisson and branching renewal process

Although the branching Poisson process (BPP) and the branching renewal process (BRP) are considered as basic models in Ascher *et al.* (1984) and Cox *et al.* (1966), they are of minor importance in practical applications. Hence, only a short description is given here. More details are found in the above references.

In the BPP a series of primary events is generated by an HPP and each primary event generates a series of subsidiary events with probability p. The BPP is realistic when a primary failure causes one or more secondary failures, which are not detected until after the system is operating again, *e.g.* because the system does not use all its components all the time. The subsidiary series is a finite RP (*i.e.* an RP which terminates after a finite number of events) and the BPP is the superposition of the primary and subsidiary series. The two types of events are supposed to be indistinguishable.

The BRP is a generalisation of the BPP as the primary series is generated by an RP. In this case only very few results have been obtained.

2.7 Regenerative stochastic processes

Regenerative stochastic processes (Smith, 1955, Ross, 1970) are of fundamental importance in this study since the analysis of the two-unit dependent parallel system is based on the regenerative nature of the underlying stochastic process.

A process is called a regenerative stochastic process (RSP) when it induces regenerative events (with probability one within finite time). A regenerative event E is characterised by the property that if E happens at t_{E_i}, i=1,2,.., the continuation of the process beyond t_{E_i} (i>1) is a probabilistic replica of the process starting from t_{E_1}. The process is said to regenerate or restart itself whenever E occurs. Hence, the process $\{X(t),t\geq 0\}$ with state space S is called an RSP with respect to the regenerative event E, if for all $A\subset S$

$$\Pr\{X(t)\in A \mid N(t)>0,\ X(s),\ 0\leq s\leq t_{N(t)}\} = \Pr\{X(t-t_{N(t)})\in A\},$$

where $\{N(t),t\geq 0\}$ is the counting process of events E. Remark that the above definition implies that the interarrival times between successive regenerative events E form a renewal process, provided that $\mathbf{E}\, N(t)<\infty$ for $t\geq 0$.

An important result for RSP's is obtained applying the key renewal theorem. Defining a cycle as the time between two successive occurrences of the event E, Ross (1970) shows (for cycles having a density function and a finite mean), for all $A \subset S$

$$\lim_{t\to\infty} \Pr\{X(t)\in A\} = \frac{E(\text{amount of time spent in A during one cycle})}{E(\text{length of a cycle})}. \tag{2.12}$$

Simple examples of RSP's are the RP, MP and SMP. In an RP every renewal is a regenerative event, since a renewal induces the start of a stochastic process with identical probabilistic properties. In a transient MP or SMP an entry in state i is a regenerative event and the interarrival times between two successive entries in state i form an imbedded renewal process. In the analysis of the parallel system the occurrence of regenerative events plays a crucial role and an imbedded renewal process is exploited to obtain expressions for a number of performance measures.

The connections between the basic models are illustrated in figure 2.2:

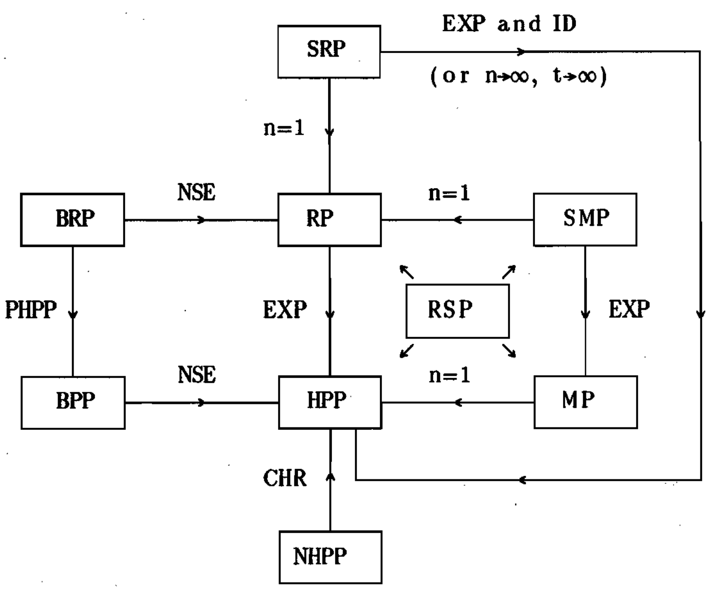


Fig. 2.2: Relations between basic models

The notation used in figure 2.2 is:

CHR : Constant Hazard Rate

EXP : X_i's EXPonentially distributed

ID : X_i's Identically Distributed

NSE : No Subsidiary Events generated

PHPP : Series of Preliminary events generated by an HPP

n : Number of system components (SRP) or number of states (MP, SMP)

t : Time

As mentioned in the introduction, factors such as maintenance, varying environmental stresses, overhauls, repairs, improved quality of new components, changes in design, etc., can effect the system's failure behaviour. A possible way of capturing the effects of real world factors in the basic probabilistic models, described in the previous sections, is to extend the state space of the models. An example in the context of Markov models is given in section 7.4.

Regression models reflect a different kind of view on the modelling of real world factors, by including them as explanatory variables in the analysis of a system's failure behaviour. Two well known families of regression models for life time data are the accelerated failure time model (also known as accelerated life model or location-scale model) and the proportional hazards model (PHM). These models are the subject of the following sections.

2.8 The accelerated failure time model

In the accelerated failure time model a covariate z acts multiplicatively on the system's time scale (Cox *et al.*, 1984). The reliability function, given z, can be written as

$$R(t|z) = R_0\left[\frac{t}{g(z)}\right],$$

where $g(.)$ is a regression function and $R_0(.)$ the reliability function for a system with $g(z)=1$. Obviously, the hazard function is

$$h(t|z) = \frac{1}{g(z)}\, h_0\left[\frac{t}{g(z)}\right],$$

where $h_0(.)$ is the baseline hazard rate:

$$h_0(t) = -\frac{d}{dt} \ln(R_0(t)).$$

As shown in Lawless (1982), an alternative way to formulate the model is in the form

$$Y = \mu(z) + \sigma e,$$

where Y is the log life time, $\mu(.)$ a location parameter, depending upon the value of the covariate z and σ a (constant) scale parameter. The random variable e has a distribution that is independent of z.

2.9 The proportional hazards model

The PHM is a regression model for the system's hazard rate. In general two approaches are distinguished: a parametric and a distribution free approach. In fact the parametric approach extends renewal processes to include regressor variables. In the distribution free approach, first suggested by Cox (1972), the hazard function h(.) is decomposed into a baseline hazard function $h_0(.)$ and a multiplicative term g(.), incorporating the effect of explanatory variables or covariates:

$$h(t|z) = h_0(t)\ g(z), \tag{2.13}$$

where the vector z contains the covariates and t represents the time.
In the non-parametric approach, no particular form is assumed for the baseline hazard function $h_0(.)$, which represents the system's hazard rate if all covariates take the value zero. Note that for different values of the covariates, the hazard functions are proportional to each other over time. This multiplicative effect gives the technique its name. The particular form suggested by Cox is

$$h(t|z) = h_0(t)\ \exp(\beta^T z), \tag{2.14}$$

where β is a vector with regression coefficients, describing the effect of the covariates. An important property of the model (2.14) is that inference about β can be made without knowledge of the baseline hazard function $h_0(.)$ (Cox *et al.*, 1984).

Since the models (2.13) and (2.14) concentrate on the time to first failure, they are not suitable for modelling the failure behaviour of repairable systems. However, Prentice *et al.* (1981) consider an extension of Cox's model to the case where multiple failures of a single system can occur. Two specific cases of the hazard model suggested by Prentice *et al.* are

$$h(t|N(t),z(t)) = h_{0s}(t)\ \exp(\beta_s^T z(t)) \tag{2.15}$$

and

$$h(t|N(t),z(t)) = h_{0s}(t-t_{N(t)})\ \exp(\beta_s^T z(t)), \tag{2.16}$$

where $\{N(t),t\geq 0\}$ is the counting process of failures, $z(.)$ a time dependent covariate and s a stratification variable. A simple example of a stratification variable is $s=N(t)+1$, in which case a system moves to stratum k after its $(k-1)^{st}$ failure, *i.e.* the baseline function $h_{0s}(.)$ depends upon the number of preceding failures. Other examples of stratification variables are given by Prentice *et al.* (1981). Note that applying model (2.16) in fact just results in a repetitive regression analysis for different strata.

Obviously, the time scale in model (2.15) is the time from the beginning of study whereas in (2.16) the time scale is the gap time or interarrival time between failures. The connection with previous models becomes clear in the absence of covariates or when they are taken constant: in the absence of covariates, the choice $h_{0s}(t)=h_0(t)$ reduces model (2.15) to an NHPP. Similarly, the assumption of constant covariates reduces model (2.16) to an SMP while the further restriction that $h_{0s}(.)=h_0(.)$ gives an ordinary RP.

Multiplicative models and particularly Cox's PHM have received a great deal of attention recently, one of the reasons being their analytical tractability. According to Ascher *et al.* (1984) the importance of the Prentice–Williams–Peterson model can hardly be overemphasised.

2.10 The additive hazards model

By analogy with the multiplicative model, for the additive hazards model (AHM) the hazard rate function is defined as

$$h(t|z)=h_0(t)+g(z), \tag{2.17}$$

where $h_0(.)$ is the baseline hazard function and $g(z)$ an additive term, incorporating the effect of the covariate z. The time scale in model (2.17) is the time from the beginning of study. Alternatively, taking the gap time or interarrival time between failures, yields

$$h(t|z)=h_0(t-t_{N(t)})+g(z), \tag{2.18}$$

where $\{N(t),t\geq 0\}$ is the counting process of failures.

Notice that $g(z)$ does not need to be positive throughout the range under consideration in order to have $h(t|z)>0$ in the models (2.17) and (2.18). Good reasons for the use of the AHM are provided by David *et al.* (1978), Aranda–Ordaz (1980) and Elandt–Johnson (1980), who consider a competing risk situation, *i.e.* a system subject to (say) n causes of failure. In this case the system's hazard rate equals the sum of the hazard rates belonging to the n causes of failure. The result is an AHM. Both Aranda–Ordaz and Elandt–Johnson mention the particular form

$$h(t|z)=h_0(t)+\beta^T z,$$

where β is a vector with regression coefficients. According to Elandt–Johnson the baseline hazard rate $h_0(t)$ might be thought of as a 'genetic ageing' effect leading to failure. It can be quite small when the system is started up and be modified by factors expressed in terms of the covariate z. This idea is used as a starting point for the following repairable systems model.

The hazard functions (2.17) and (2.18) can be thought of as the hazard rate of a two–unit series system, consisting of the components C_s and C_c. Component C_s is called the system component and represents the repairable system itself. C_s has hazard function $h_0(t)$. The effects of the covariates on C_s are lumped and modelled by an artificial component C_c, called the covariate component. C_c has hazard rate $g(.)$. The connection with a two–unit dependent parallel system is obvious when C_s is replaced by a two–unit independent parallel system (with components C_1 and C_2). Of course in this case C_c is used to model the common cause effect (figure 2.3).

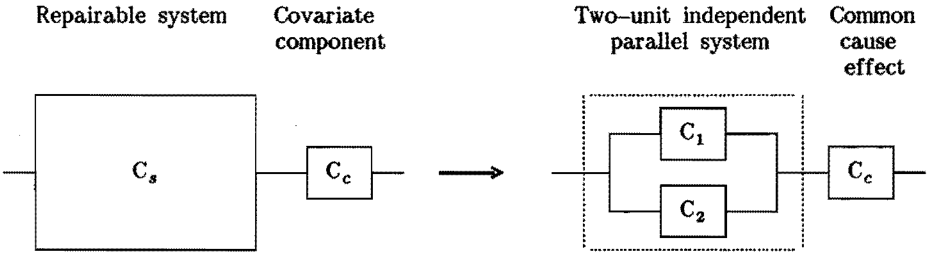


Fig. 2.3: The two-unit dependent parallel system modelled by an AHM

The system's deterioration can be expressed in the failure behaviour of C_c in the following way. Let $\{N(t),t\geq 0\}$ be the counting process of failures at system level, *i.e.* failures of either C_s or C_c. It is supposed that after each failure at system level, C_s is replaced by an identical new component with hazard rate $h_0(.)$ and C_c is replaced by a component with constant hazard rate $g(z_{N(t)})$. Hence, the hazard rate $g(z_{N(t)})$, as well as the covariate z, is assumed to be constant in the intervals between successive system failures. Obvious candidates for z are binary or discrete variables as on/off, high/low, the number of cold starts or repairs up to time t, etc. Continuous variables (temperature, pressure, humidity) may be discretised, the mean can be taken or an extreme value. The term $g(z_{N(t)})$ represents the effect of the covariates on the baseline hazard rate $h_0(.)$. The size of $g(z_{N(t)})$ may be interpreted as a deterioration related to a renewal process. The interpretation of a failure of C_c is that the system has failed under influence of the environment.

Since C_s is renewed after every failure at system level, the time scale of the baseline hazard rate is the gap time. Thus the system's hazard rate can be written as

$$h(t\,|\,N(t),z_{N(t)})=h_0(t-t_{N(t)})+g(z_{N(t)}),$$

an additive variant of the PWPM.

With respect to the functional form of g(.) there are many possibilities. A simple example is the linear form

$$g(z)=\sum_{i=1}^{n}\beta_i z_i,$$

where the β_i's are the regression coefficients. In contrast with the PHM, a parametric form has to be chosen for the baseline hazard function $h_0(.)$, in order to make inference about β. For more details on the AHM, the reader is referred to Pijnenburg (1991).

Summarising this chapter, a brief review is presented of basic stochastic models and processes for the analysis of repairable systems. Starting in chapter 3, the methods and techniques, used to derive the performance measures of the two-unit dependent parallel system, will be based on elements of the stochastic processes described in this chapter.

3. BIVARIATE EXPONENTIALLY DISTRIBUTED LIFE TIMES

3.1 Introduction

It is well known in reliability analysis that provision of redundancy improves the performance of a system. Redundancy reduces the system down time and enhances the reliability, within technological constraints.

Parallel and standby are the two basic redundant configurations. The standby configuration provides a backup for the on-line operating unit. There are two distinct types of standby depending on whether the standby unit deteriorates or not. When the standby deteriorates it is called warm standby, otherwise cold standby.

There is an extensive literature for the parallel redundant system. The original contribution is due to Gaver (1963, 1964), who considered a two unit parallel redundant system with exponentially distributed life times and arbitrarily distributed repair times. Gaver used supplementary variables (Cox, 1955[a]) to derive the mean time to system failure (MTSF) and the limiting or stationary availability. Ever since this reported research there have been attempts to derive the MTSF and the stationary availability under relaxed assumptions on the life and repair time distributions of the units in the system. Some of the notable contributions are Liebowitz (1966), Kodama *et al.* (1974), Linton (1976), Ravichandran (1981) and Osaki (1985). Ravichandran (1990) reviewed the state of the art for a two-unit parallel redundant system.

In order to capture the reality of a practical situation, there are several directions in which the system can be extended. One of them is the situation where the units are statistically *dependent*. In this study dependence arises from the occurrence of common cause failures. Harris (1968) used a bivariate exponential distribution (BVE) to model the life times of the units and derived the MTSF, using the supplementary variable technique. Harris allowed the repair times to be arbitrarily distributed. Osaki (1980) extended the analysis to obtain the availability of the system, using a variant of a semi-Markov process with some non-regeneration points.

In the present study, the analysis of a two-unit dependent parallel system is performed, by constructing an appropriate imbedded renewal process. Important operating characteristics follow from the renewal process in a direct and elegant way: not only the MTSF and availability, but also the (stationary) system state probabilities, the (stationary) interval reliability, the (stationary) joint availability and two quasi-stationary distributions. After a description of the BVE and the system's stochastic behaviour in section 3.2, the imbedded renewal process is investigated in section 3.3. Using the renewal process, explicit expressions are obtained in section 3.4 for the reliability, MTSF, (stationary) availability and (stationary) system state probabilities. The (stationary) interval reliability and joint availability are analysed in 3.5 and the stationary behaviour of the system, under the condition that a system failure has not occurred, is examined in section 3.6. Finally, intensity functions associated with the counting process of various stochastic point events form the subject of 3.7: the expected value of the number of events in (0,t) and its variance are derived.

3.2 The bivariate exponential distribution

Marshall *et al.* (1967) introduced the BVE and generalised it to the multivariate case. The first reported application of the BVE in a reliability model of a two unit parallel redundant system, is in Harris (1968). The BVE used by Harris is the following. Let T_1 and T_2 be random variables representing the life times of the dependent units in the parallel system. The joint survival function of the life times has representation

$$\bar{F}(t_1,t_2) = \Pr\{T_1>t_1,T_2>t_2\} = \exp(-\lambda_1 t_1-\lambda_2 t_2-\lambda_{12}\max(t_1,t_2)), \tag{3.1}$$

where λ_1, λ_2, λ_{12} are nonnegative constants and $t_1,t_2\geq 0$.

The BVE (3.1) is a natural way to model a system with dependent units and has a simple physical interpretation in terms of a shock model. Suppose the system is subjected to three independent sources of shocks. Shocks from source i (i=1,2) are generated by an HPP with rate λ_i and cause a failure of unit i. Shocks from source 3 are generated by an HPP with rate λ_{12} and affect both units simultaneously. Thus, if U_1, U_2 and U_{12} represent the random times until the occurrence of a shock from source 1, 2 and 3 respectively, then

$$T_1 = \min(U_1,U_{12}),$$

$$T_2 = \min(U_2,U_{12})$$

and obviously

$$\Pr\{T_1>t_1,T_2>t_2\} = \exp(-\lambda_1 t_1-\lambda_2 t_2-\lambda_{12}\max(t_1,t_2)).$$

The BVE (3.1) has exponential marginal distribution functions with survival probabilities

$$\Pr\{T_1>t_1\} = \exp(-(\lambda_1+\lambda_{12})t_1)$$

and

$$\Pr\{T_2>t_2\} = \exp(-(\lambda_2+\lambda_{12})t_2).$$

A property of great value in reliability theory is the lack-of-memory property of the (univariate) exponential distribution: if the random variable T is exponentially distributed, then for all $s,t\geq 0$

$$\Pr\{T>s+t \mid T>s\} = \Pr\{T>t\}.$$

Hence the probability of survival for an additional t units of time for a unit of age s is the same as that of a new unit. It is easily seen that the BVE (3.1) enjoys the corresponding bivariate properties

$$\Pr\{T_1>s+t_1,\ T_2>s+t_2 \mid T_1>s,\ T_2>s\} = \Pr\{T_1>t_1,\ T_2>t_2\} \tag{3.2}$$

and

$$\Pr\{T_1>t_1+s,\ T_2>t_2+s \mid T_1>t_1,\ T_2>t_2\} = \Pr\{T_1>s,\ T_2>s\}, \tag{3.3}$$

for all $s,t_1,t_2\geq 0$. Equation (3.2) says that the joint survival function of a pair of units each of age s is the same as that of a pair of new units. On the other hand (3.3) states that the probability of surviving an additional s units of time is independent of the age of the units. From (3.2) and (3.3) it follows that the joint survival function $\bar{F}(.,.)$ satisfies

$$\bar{F}(s+t_1,s+t_2) = \bar{F}(t_1,t_2)\ \bar{F}(s,s). \tag{3.4}$$

In Barlow *et al.* (1975) it is shown that the BVE (3.1) is the unique bivariate distribution with exponential marginals satisfying (3.4).

It is useful to note that:

i. The sojourn time in the state where both units are operating is exponentially distributed with parameter $\lambda_1+\lambda_2+\lambda_{12}$.

ii. The marginal life time distribution of unit i (i=1,2) is exponential with parameter $\lambda_i+\lambda_{12}$.

iii. The time until an occurrence of a common cause failure is exponentially distributed with parameter λ_{12}.

In this chapter, the BVE (3.1) is used to represent the joint life times of two dependent units connected in parallel. The repair time duration of unit i (i=1,2) is an arbitrarily distributed random variable with pdf $g_i(.)$.

When the units are not identical, the state description is in terms of a process $\{X(t, t \geq 0\}$ with state space $S_X=\{(o,o), (o,r), (r,o), (r,w), (w,r)\}$, as in section 1.3. The one-step transition diagram in figure 3.1 shows the essential dynamics of the system.

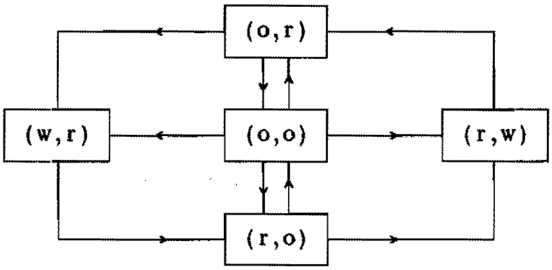


Fig. 3.1: One-step transition diagram of the X(t) process

At time t=0 both units are assumed to be new and to operate. As the joint life time is characterised by the BVE (3.1), after a random duration both the units fail simultaneously (with rate λ_{12}) or either of them fails (unit 1 with rate λ_1 and unit 2 with rate λ_2). When both units fail, with probability p_i (i=1,2) unit i is taken under repair and the system moves to state (r,w) or (w,r). When one of the units fails, the system moves to either state (o,r) or (r,o), depending on which unit has failed. In state (r,o) the following developments are possible: since unit 2 is operating and the other unit is under repair, the life time T_2 of the operating unit is exponentially distributed with parameter $\lambda_2+\lambda_{12}$ and the repair time is a random variable S_1 with pdf $g_1(.)$. A transition occurs at the minimum of (T_2,S_1). If $\min(T_2,S_1)=S_1$, then the system moves to state (o,o), otherwise a transition is made to (r,w). In the latter case the system enters (o,r) after $\max(T_2,S_1)$ units of time, *i.e.* after repair completion of unit 1.

When the units are physically identical but statistically dependent, a state description of the system can be given as the number of units operating (or failed), as in section 1.5. Let X(t) represent the number of failed units in the system at time t. Hence, X(t)=0 means that both units are working, if X(t)=1, one unit is working and the other unit is undergoing a repair and if X(t)=2, no unit is operating. State 2 is the down state and the states 0 and 1 are the up states. The system's one-step transition diagram is given in figure 3.2.

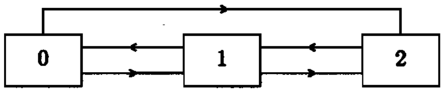


Fig 3.2: One-step transition diagram, identical units

In this chapter attention is focused on a system with identical units. The case with nonidentical units is treated in the chapters 4 and 5. Since they are identical, the units are supposed to have identically distributed repair times, with pdf g(.). Questions about the operating characteristics are translated into questions about the process $\{X(t),t\geq 0\}$. It is supposed that X(0)=0: at time t=0 the system is operating with both units working. The reliability of the system is determined by the distribution of the first passage time of the process to state 2. The system is available if X(t)=0 or X(t)=1. The stationary availability is similarly related to the stationary distribution of the process $\{X(t),t\geq 0\}$. Thus, a study of $\{X(t),t\geq 0\}$ yields the operating characteristics of the system explicitly.
To study the behaviour of the induced stochastic process, it is convenient to introduce the events E_i as

$$E_i \text{ : entrance into state i, i=0,1,2.}$$

Table 3.1 summarises the possibilities for the occurrences of the events E_i and their associated properties.

Event	Possible occurrences	Description	Nature of the event
E_0	Initial occurrence From state 1	Both units start operating Repair completion	Regenerative
E_1	From state 0 From state 2	Failure of one of the units Repair completion	Regenerative
E_2	From state 0 From state 1	Failure of both units Failure of operating unit before repair completion of the other unit	Non - regenerative

Table 3.1: Description of the events E_i

Note that the process X(t) is non–regenerative with respect to the event E_2, since the conditional behaviour of the process X(t) depends upon the elapsed repair time at the moment of the transition to state 2. However, the nature of *common cause failures* is regenerative, since transitions from state 0 to state 2 induce the *start* of a repair. Hence, let E_α denote a transition from state 0 to state 2 and E_β a transition from state 1 to state 2, then X(t) is regenerative with respect to the event E_α, but non–regenerative with respect to E_β.

In the analysis of the system, occurrences of the E_1 event play a crucial role as they are regenerative and thus form a renewal process.

The following notation will be used:

$f(t)$: pdf of a random variable

$F(t) = \int_0^t f(u)du$

$\overline{F}(t) = 1-F(t)$

$*$: convolution symbol: $f(t)*g(t) = \int_0^t f(u)g(t-u)du$

$f^{(n)}(t)$: n–fold convolution of $f(t)$

$f^*(s)$: Laplace transform of $f(t)$: $f^*(s) = \int_0^\infty e^{-st} f(t)\, dt$

$\overline{F}_0(t) = \exp(-(2\lambda+\lambda_{12})t)$

$\overline{F}_1(t) = \exp(-(\lambda+\lambda_{12})t)$

$f_1(t) = (\lambda+\lambda_{12}) \exp(-(\lambda+\lambda_{12})t)$

Remark that, from a mathematical point of view, $(f*g)(t)$ is preferred to denote the convolution of f(.) and g(.) at the point t. However, for convenience the above notation is chosen: it makes the formulas more readable, it is not confusing and is in agreement with the notation used by a number of other authors in the reliability field.

3.3 The imbedded renewal process

In the previous section several events have been identified to describe the stochastic behaviour of the system under consideration. Most of the events are regenerative except for E_2 when it occurs from state 1. It is an important aspect of these systems that the nature of an event is a function of the state from which it was entered. In the present analysis the occurrences of the event E_1 play the most significant role in obtaining the operating characteristics of the system. The random variables relating to E_1 are developed here.

Let X_{11} be the random variable representing the time interval between two successive visits to state 1. At every new visit to state 1 an E_1 event occurs and hence the durations between two successive visits correspond to intervals in a renewal process. The pdf of X_{11} characterises this renewal process. The pdf of X_{11} is composed of three distinct parts, corresponding to the possible paths taken by the process $\{X(t),t\geq 0\}$.

Starting in state 1 these paths are:

i. A transition to state 2 occurs, caused by a failure of the operating unit before repair completion of the other unit. A stay in state 2 is always followed by a transition to state 1, *i.e.* the occurrence of an E_1 event.

On the other hand a transition to state 0 occurs when the repair is completed before a failure of the operating unit. Subsequently there are two possibilities:

ii. A single unit fails. Hence the process enters state 1, inducing an event E_1.

iii. Both units fail simultaneously, caused by the occurrence of a common cause failure. Hence state 2 is entered, followed by an E_1 event when the unit under repair restarts operating.

The above paths are denoted by respectively 1–2–1, 1–0–1 and 1–0–2–1. Before writing a formal expression for the pdf of the random variable X_{11}, the pdf of the random variables characterising transitions from state 0 to state 1 and 2 are obtained. Let $f_{0i}(t)$ denote the pdf of the length of a stay in state 0, followed by a transition to state i. Formally,

$$f_{0i}(t)\ \Delta t = \Pr\{X(u)=0,\ 0\le u\le t,\ X(t+\Delta t)=i\},\ \Delta t\downarrow 0.$$

To obtain $f_{0i}(t)$, note that the sojourn time in state 0 is exponentially distributed with parameter $2\lambda+\lambda_{12}$. Further, shocks which destroy a single unit occur with rate 2λ and shocks which destroy both units simultaneously occur with rate λ_{12}, when the system is in state 0. Hence,

$$f_{01}(t) = 2\lambda \exp(-(2\lambda+\lambda_{12})t) \tag{3.5}$$

and

$$f_{02}(t) = \lambda_{12} \exp(-(2\lambda+\lambda_{12})t). \tag{3.6}$$

Remark that $f_{01}(.)$ and $f_{02}(.)$ are defective pdf's (since a stay in state 0 is not necessarily followed by a transition to state 1, respectively 2), but $f_{01}(t)+f_{02}(t)$ is a proper density function (namely of the length of the sojourn time in state 0).

From (3.5), (3.6) and the description of X_{11} it follows that the pdf of X_{11}, denoted by $f_{11}(.)$, satisfies

$$f_{11}(t) = g(t)F_1(t) + [g(t)\bar{F}_1(t)]*f_{01}(t) + [g(t)\bar{F}_1(t)]*f_{02}(t)*g(t). \tag{3.7}$$

The three terms in (3.7) correspond to the paths 1–2–1, 1–0–1 and 1–0–2–1. Expression (3.7) completely characterises the renewal process induced by E_1 events. It is evident that the random variable X_{11} is proper. Its Laplace transform is

$$f^*_{11}(s) = g^*(s) + g^*(\lambda+\lambda_{12}+s)\left[\frac{2\lambda+\lambda_{12}g^*(s)}{2\lambda+\lambda_{12}+s} - 1\right]. \tag{3.8}$$

From (3.8),

$$\mathsf{E}\,X_{11} = \mu + \frac{g^*(\lambda+\lambda_{12})}{2\lambda+\lambda_{12}}(1+\lambda_{12}\mu), \tag{3.9}$$

where μ is the mean repair time.
Further, the pdf of a modified version of the random variable X_{11} is needed in the next section, where expressions are obtained for the system reliability and availability. Let $\tilde{X}_{11}$ denote the time interval between two successive occurrences of E_1 events, while the process X(t) does *not* visit state 2 in between the visits to state 1. Let $\tilde{f}_{11}(.)$ denote the pdf of $\tilde{X}_{11}$, *i.e.*

$$\tilde{f}_{11}(t)\,\Delta t = \Pr\{t\le X_{11}<t+\Delta t,\ X(u)\ne 2,\ 0\le u\le t\},\ \Delta t\downarrow 0.$$

The function $\tilde{f}_{11}(.)$ is obtained by dropping the first and the third term in (3.7). Retaining the second term, corresponding to transitions 1–0–1, it follows

$$\tilde{f}_{11}(t) = [g(t)\bar{F}_1(t)]*f_{01}(t). \tag{3.10}$$

The density given by (3.10) for obvious reasons is defective. The Laplace transform of (3.10) is

$$\tilde{f}^*_{11}(s) = \frac{2\lambda g^*(\lambda+\lambda_{12}+s)}{2\lambda+\lambda_{12}+s}. \tag{3.11}$$

In section 3.4 it is shown how the system reliability and availability can be obtained in terms of X_{11} and $\tilde{X}_{11}$.

3.4 Reliability and availability

Let $R_1(t)$ and $A_1(t)$ be the reliability and the availability of the system, conditioned on an E_1 event at the time origin. Further, define the function $\alpha(t)$ as

$$\alpha(t) = \bar{G}(t)\bar{F}_1(t) + [g(t)\bar{F}_1(t)]*\bar{F}_0(t). \tag{3.12}$$

Then

$$R_1(t) = \alpha(t) + \sum_{n=1}^{\infty} \tilde{f}_{11}^{(n)}(t)*\alpha(t) \tag{3.13}$$

and

$$A_1(t) = \alpha(t) + \sum_{n=1}^{\infty} f_{11}^{(n)}(t)*\alpha(t). \tag{3.14}$$

The derivation of the expressions (3.12)–(3.14) is achieved by observing the stochastic behaviour of the process $\{X(u),u\geq 0\}$ in the interval $(0,t)$. The function $\alpha(t)$ is used both in the reliability and availability function and represents the probability that in an interval of length t initiated by an E_1 event, the system *neither* fails *nor* induces the occurrence of an E_1 event:

$$\alpha(t) = \Pr\{X_{11}>t,\ X(u)\neq 2,\ 0\leq u\leq t\}.$$

Expression (3.12) is obtained by considering the following mutually exclusive cases:

i. The repair which started at the time origin is not completed at time t. Hence the operating unit is required to work without failure until time t, which occurs with probability $\bar{G}(t)\bar{F}_1(t)$.

ii. The repair commenced at the time origin is completed in $(u,u+du)$, $u<t$. In this case the operating unit is required to work without failure up to the repair completion of the other unit, resulting in a transition into state 0. An entry into state 0 occurs with probability $g(u)\bar{F}_1(u)du$ in $(u,u+du)$. Further, *none* of the units is allowed to fail in (u,t) as any failure (individual or common cause) would either induce an E_1 event or a system failure (which is automatically followed by an E_1 event). The latter probability being $\bar{F}_0(t-u)$, the computation of $\alpha(t)$ is complete.

Given (3.12), the expressions (3.13) and (3.14) are obtained by classifying the time interval (0,t) as:

i. No E_1 events occur in (0,t).

ii. Several occurrences of E_1 events take place in (0,t) and the last one occurred in (u,u+du), u<t.

Notice that in (3.14) the system can visit state 2 between two successive E_1 events, whereas in (3.13) such visits are not possible. These requirements are met using the functions $f_{11}(.)$ and $\tilde{f}_{11}(.)$. The expressions (3.12)–(3.14) are fundamental for the analysis of the system and the reasoning used in their derivation is typical and standard for the subsequent results in this study.

The reliability and availability are two key operating characteristics of a system. Subsequently, two important measures associated with the performance of the system are derived, *viz.* the mean time to system failure (MTSF) and the system's stationary availability (A). They are well known in the reliability literature (Birolini, 1985) in terms of their direct physical interpretation. Starting point for the computation of the MTSF and A are the Laplace transforms of (3.12)–(3.14). From (3.12),

$$\alpha^*(s) = \frac{1}{\lambda+\lambda_{12}+s} + g^*(\lambda+\lambda_{12}+s)\left[\frac{1}{2\lambda+\lambda_{12}+s} - \frac{1}{\lambda+\lambda_{12}+s}\right]. \tag{3.15}$$

The reliability $R_1(t)$ of the system is

$$R_1(t) = \Pr\{X(u)\neq 2,\ 0\leq u\leq t \mid E_1 \text{ at } t=0\}$$

and

$$\text{MTSF} = \int_0^{\infty} R_1(t)dt$$

Equivalently,

$$\text{MTSF} = R_1^*(s)\Big|_{s=0}.$$

From (3.13), the Laplace transform of the system reliability, conditioned on an E_1 event at t=0, is

$$R_1^*(s) = \frac{\alpha^*(s)}{1-\tilde{f}^*_{11}(s)} \tag{3.16}$$

and after substitution of (3.11) and (3.15),

$$R_1^*(s) = \frac{2\lambda+\lambda_{12}+s-\lambda g^*(\lambda+\lambda_{12}+s)}{(2\lambda+\lambda_{12}+s-2\lambda g^*(\lambda+\lambda_{12}+s))\ (\lambda+\lambda_{12}+s)}. \tag{3.17}$$

Hence,

$$\text{MTSF} = \frac{2\lambda+\lambda_{12}-\lambda g^*(\lambda+\lambda_{12})}{(2\lambda+\lambda_{12}-2\lambda g^*(\lambda+\lambda_{12}))\ (\lambda+\lambda_{12})}. \tag{3.18}$$

Using (3.14), the Laplace transform of the system's availability, conditioned on an E_1 event at $t=0$ is

$$A_1^*(s) = \frac{\alpha^*(s)}{1-f_{11}^*(s)}.$$

Substitution of (3.8) and (3.15) yields

$$A_1^*(s) = \frac{2\lambda+\lambda_{12}+s-\lambda g^*(\lambda+\lambda_{12}+s)}{\left[(2\lambda+\lambda_{12}+s+\lambda_{12}g^*(\lambda+\lambda_{12}+s))\ (1-g^*(s)) + sg^*(\lambda+\lambda_{12}+s)\right]\ (\lambda+\lambda_{12}+s)}. \tag{3.19}$$

The stationary availability A of the system, conditioned on an E_1 event at the time origin, is obtained as the limiting value of $A_1(t)$, as $t\to\infty$. Note that the stationary availability is independent of the initial condition by the regenerative nature of the E_1 events. The key renewal theorem (Smith, 1958) is used to derive an expression for A. To apply the key renewal theorem, suppose that the pdf g(.) is non-lattice (Feller, 1966) and Riemann integrable. Then it follows from (3.12) that $\alpha(t)$ is non-lattice and Riemann integrable. Moreover, $\alpha(t)$ is also direct Riemann integrable, since

i. $\alpha(t)$ is non-negative and non-increasing,

ii. from (3.15), $\int_0^\infty \alpha(t)\,dt < \infty$.

Now, the key renewal theorem yields

$$A = \frac{1}{E\,X_{11}} \int_0^\infty \alpha(u)du$$

or, equivalently,

$$A = \frac{2\lambda+\lambda_{12}-\lambda g^*(\lambda+\lambda_{12})}{(\lambda+\lambda_{12})\ \left[\mu(2\lambda+\lambda_{12})+g^*(\lambda+\lambda_{12})(1+\lambda_{12}\mu)\right]}. \tag{3.20}$$

An alternative way to compute the stationary availability is by using the Abelian theorem for Laplace transforms (Cohen, 1982, p.651), which state that A may be computed as $A = \lim_{s \downarrow 0} s\, A_1^*(s)$.

The stationary availability can also be obtained from the stationary state probabilities of the process X(t). The system state probabilities are characterised using the imbedded renewal process corresponding to the occurrences of E_1 events and arguments similar to those used in the derivation of (3.13) and (3.14). The state probabilities are defined as

$$P_{1,i}(t) = \Pr\{X(t)=i \mid E_1 \text{ at } t=0\},\ i=0,1,2.$$

From the stated behaviour of the process, it follows that

$$P_{1,i}(t) = \gamma_i(t) + \sum_{n=1}^{\infty} f_{11}^{(n)}(t) * \gamma_i(t), \tag{3.21}$$

where

$$\gamma_0(t) = [g(t)\bar{F}_1(t)] * \bar{F}_0(t),$$

$$\gamma_1(t) = \bar{G}(t)\bar{F}_1(t)\cdot$$

and

$$\gamma_2(t) = \bar{G}(t)F_1(t) + [g(t)\bar{F}_1(t)] * f_{02}(t) * \bar{G}(t).$$

The functions $\gamma_i(t)$ (i=0,1) represent the probability that the system is found in state i at time t and it neither has failed nor has induced an E_1 event in (0,t). Similarly, $\gamma_2(t)$ represents the probability that the system is down at time t and has not induced an E_1 event during the interval (0,t). Note that $\alpha(t)=\gamma_0(t)+\gamma_1(t)$. The key renewal theorem yields the stationary distribution $\{\pi_i\}_{i=0,1,2}$ of the process X(t) as

$$\pi_i = \frac{1}{\mathsf{E}\, X_{11}} \int_0^{\infty} \gamma_i(t)dt.$$

Using Laplace transforms and substitution of $\mathsf{E}\, X_{11}$ yields explicit expressions for the stationary state probabilities. It follows that

$$\pi_0 = \frac{g^*(\lambda+\lambda_{12})}{\mu(2\lambda+\lambda_{12})+g^*(\lambda+\lambda_{12})(1+\lambda_{12}\mu)}, \tag{3.22}$$

$$\pi_1 = \frac{(2\lambda+\lambda_{12})(1-g^*(\lambda+\lambda_{12}))}{(\lambda+\lambda_{12})\ \left[\mu(2\lambda+\lambda_{12})+g^*(\lambda+\lambda_{12})(1+\lambda_{12}\mu)\right]} \tag{3.23}$$

and

$$\pi_2 = \frac{\mu(\lambda+\lambda_{12})(2\lambda+\lambda_{12}) - (2\lambda+\lambda_{12})(1-g^*(\lambda+\lambda_{12}))}{(\lambda+\lambda_{12})\ \left[\mu(2\lambda+\lambda_{12})+g^*(\lambda+\lambda_{12})(1+\lambda_{12}\mu)\right]} \tag{3.24}$$

$$+ \frac{\lambda_{12}\mu(\lambda+\lambda_{12})g^*(\lambda+\lambda_{12})}{(\lambda+\lambda_{12})\ \left[\mu(2\lambda+\lambda_{12})+g^*(\lambda+\lambda_{12})(1+\lambda_{12}\mu)\right]}.$$

Alternatively, the stationary distribution π_i can be computed as

$$\pi_i = \lim_{s\downarrow 0} s\ P^*_{1,i}(s).$$

The advantage of studying the stochastic process itself rather than transforms is the ability to deal with additional performance measures such as the cost of operating or maintaining the system, since the π_i's represent the fraction of time the system spends in state i (i=0,1,2) in the stationary case. It is easily verified that $\pi_0+\pi_1$, determined by (3.22)–(3.23), agrees with the stationary availability A obtained in (3.20).

Expressions (3.13) and (3.14) give the reliability and the availability of the system, conditioned by an E_1 event at the time origin. To conclude this section, it is shown that the above measures can be extended easily to the case where a (regenerative) E_0 event occurs at the time origin. Using $R_0(t)$ and $A_0(t)$ to represent the reliability and availability under the changed initial condition, it is easily found that

$$R_0(t) = \bar{F}_0(t) + f_{01}(t)*R_1(t) \tag{3.25}$$

and

$$A_0(t) = \bar{F}_0(t) + [f_{01}(t)+f_{02}(t)*g(t)]*A_1(t), \tag{3.26}$$

where $R_1(t)$ and $A_1(t)$ are determined by (3.13) and (3.14).

The derivation of (3.25) and (3.26) is based on considering whether there is a failure or not in the time interval under consideration. When there is no failure in (0,t), the system is reliable in (0,t) with probability $\bar{F}_0(t)$ (and hence available at time t). When there is a failure, an E_1 event is induced

and the required probability is related to $R_1(t)$ and $A_1(t)$ as in the second term of the equations (3.25) and (3.26).
Obviously,

$$A_0^*(s) = \frac{2\lambda}{2\lambda+\lambda_{12}+s} + \frac{2\lambda+\lambda_{12}g^*(s)}{2\lambda+\lambda_{12}+s} A_1^*(s)$$

and hence

$$\lim_{s\downarrow 0} s\, A_1^*(s) = \lim_{s\downarrow 0} s\, A_0^*(s),$$

which confirms that the limiting availability is independent of the initial condition. Note however that the MTSF *does* depend on the initial condition.
With respect to the system state probabilities $P_{0,i}(t)$, which are conditioned on an E_1 event at the time origin, it is easily verified that

$$P_{0,0}(t) = F_0(t) + [f_{01}(t)+f_{02}(t)*g(t)]*P_{1,0}(t),$$

$$P_{0,1}(t) = [f_{01}(t)+f_{02}(t)*g(t)]*P_{1,1}(t)$$

and

$$P_{0,2}(t) = f_{01}(t)*P_{1,2}(t) + f_{02}(t)*[\overline{G}(t)+g(t)*P_{1,2}(t)].$$

Taking the Laplace transforms it is immediate that

$$\lim_{s\downarrow 0} s\, P_{0,i}^*(s) = \lim_{s\downarrow 0} s\, P_{1,i}^*(s) = \pi_i,$$

where the stationary state probabilities π_i are given by (3.22)–(3.24).

3.5 Interval reliability and joint availability

The analysis is now extended to some of the more general operating characteristics, *viz.* the interval reliability and joint availability. Again, the imbedded renewal process described by (3.7) plays a dominant role in the derivations. The interval reliability $R_1(t,\tau)$ represents the probability that the system is available for a duration τ, beginning at time t, conditioned by an E_1 event at the time origin. Formally,

$$R_1(t,\tau) = \Pr\{X(u)\neq 2,\ t\leq u\leq t+\tau \mid E_1 \text{ at } t=0\}.$$

The interval reliability function is a combined measure of availability and reliability introduced earlier. The reliability and availability conditioned on an E_1 event at $t=0$ are recovered from $R_1(t,\tau)$ by setting t and τ respectively to zero.

The joint availability $A_1(t,\tau)$ represents the joint probability that the system is available at the time epochs t and $t+\tau$, given an E_1 event at t=0:

$$A_1(t,\tau) = \Pr\{X(u)=0 \vee X(u)=1,\ u=t,t+\tau \mid E_1 \text{ at } t=0\}.$$

The interval reliability of the system, conditioned by an E_1 event at t=0, is given by

$$R_1(t,\tau) = \varphi(t,\tau) + \sum_{n=1}^{\infty} \int_0^t f_{11}^{(n)}(u)\ \varphi(t-u,\tau)\ du, \tag{3.27}$$

where

$$\varphi(t,\tau) = \bar{G}(t+\tau)\bar{F}_1(t+\tau) + \int_0^{t+\tau} g(u)\bar{F}_1(u)\bar{F}_0(t+\tau-u)du + \int_t^{t+\tau} \tilde{f}_{11}(u)R_1(t+\tau-u)du. \tag{3.28}$$

Expression (3.27) is obtained by classifying the events according to the number of E_1 events in the intervals (0,t) and $(t,t+\tau)$. The function $\varphi(t,\tau)$ gives the probability that no E_1 events occur in (0,t) and no system failures in $(0,t+\tau)$, given an E_1 event at t=0:

$$\varphi(t,\tau) = \Pr\{X_{11}>t,\ X(u)\neq 2,\ 0\leq u\leq t+\tau\} \tag{3.29}$$

In $\varphi(t,\tau)$ the first two terms follow from the non-occurrence of an E_1 event in both (0,t) and $(t,t+\tau)$ and the third term follows from a non-occurrence in (0,t) and one or more occurrences in $(t,t+\tau)$. Notice that

$$\varphi(t,\tau) = \alpha(t+\tau) + \int_t^{t+\tau} \tilde{f}_{11}(u)\ R_1(t+\tau-u)du.$$

The integral in (3.27) represents the probability of exactly n occurrences, $n\geq 1$, in (0,t) and any number in $(t,t+\tau)$.

Subsequently, the joint availability of the system, conditioned by an E_1 event at t=0, is given by

$$A_1(t,\tau) = \psi(t,\tau) + \sum_{n=1}^{\infty} \int_0^t f_{11}^{(n)}(u)\ \psi(t-u,\tau)du, \tag{3.30}$$

where

$$\psi(t,\tau) = \bar{G}(t+\tau)\ \bar{F}_1(t+\tau) \tag{3.31}$$

$$+ \int_t^{t+\tau} g(u)\ \bar{F}_1(u)\ A_0(t+\tau-u)du$$

$$+ \int_t^{t+\tau} f_1(u) \int_0^{t+\tau-u} g(u+v)\ A_1(t+\tau-u-v)dv\ du$$

$$+ \int_0^{t} g(u)\ \bar{F}_1(u)\ \bar{F}_0(t+\tau-u)du$$

$$+ \int_0^{t} g(u)\ \bar{F}_1(u) \int_{t-u}^{t+\tau-u} f_{01}(v)\ A_1(t+\tau-u-v)dv\ du$$

$$+ \int_0^{t} g(u)\ \bar{F}_1(u) \int_{t-u}^{t+\tau-u} f_{02}(v) \int_0^{t+\tau-u-v} g(w)\ A_1(t+\tau-u-v-w)dw\ dv\ du.$$

The derivation of (3.30) involves a careful consideration of the process until time $t+\tau$. The major classification is the number of occurrences of E_1 events in $(0,t)$. The further subclassification is based on whether the process is in state 0 or state 1 at time t and whether the process remains in this state or not during $(t,t+\tau)$. The function $\psi(t,\tau)$ represents the probability of no E_1 events and no system failures in $(0,t)$ and the system being up at time $t+\tau$:

$$\psi(t,\tau) = \Pr\{X_{11}>t,\ X(u)\neq 2,\ 0\leq u\leq t,\ X(t+\tau)\neq 2\}. \tag{3.32}$$

More specific, the first three terms of $\psi(t,\tau)$ represent the probability of no E_1 events in $(0,t)$, the system being in state 1 at time t and:

i. No transitions occur in $(t,t+\tau)$.

ii. The first transition after t is to state 0.

iii. The first transition after t is to state 2.

The remaining three terms represent the probability of no E_1 events in $(0,t)$, the system being in state 0 at time t and:

iv. No transitions occur in $(t,t+\tau)$.

v. The first transition after t is to state 1.

vi. The first transition after t is to state 2.

The limiting behaviour of the interval reliability and the joint availability is obtained by Laplace transform techniques. Regarding τ as a constant, the Laplace transform $R_1^*(s,\tau)$ is defined as

$$R_1^*(s,\tau) = \int_0^\infty e^{-st} R_1(t,\tau)\, dt.$$

Let $\Re(\tau)$ denote the stationary interval reliability for an interval of length τ, *i.e.*

$$\Re(\tau) = \lim_{t\to\infty} R_1(t,\tau),$$

then

$$\Re(\tau) = \lim_{s\downarrow 0} s\, R_1^*(s,\tau).$$

From expression (3.27),

$$R_1^*(s,\tau) = \frac{\varphi^*(s,\tau)}{1 - f_{11}^*(s)}$$

and applying De l'Hôpital's rule

$$\Re(\tau) = \lim_{s\downarrow 0} \frac{\varphi^*(s,\tau)+s\dot{\varphi}^*(s,\tau)}{-\dot{f}_{11}^*(s)},$$

where the dot is used to denote the derivative of $\varphi^*(s,\tau)$ with respect to s:

$$\dot{\varphi}^*(s,\tau) = \frac{d}{ds}\, \varphi^*(s,\tau).$$

Remark that the notation $\dot{\varphi}^*(.)$ is used to denote the derivative of the Laplace transform of $\varphi(.)$ and *not* the Laplace transform of the derivative.

If $\dot{\varphi}^*(0,\tau)$ and $\varphi^*(0,\tau)$ are finite, it follows that

$$\Re(\tau) = \frac{\varphi^*(0,\tau)}{E\, X_{11}}. \tag{3.33}$$

In order to prove that (3.33) holds, note that $\varphi(t,0)=\alpha(t)$ and as a result $\varphi(t,\tau)\le\alpha(t)$. Thus

$$0 \le -\dot{\varphi}^*(0,\tau) = \int_0^\infty t\, \varphi(t,\tau)dt \le \int_0^\infty t\, \alpha(t)dt = -\dot{\alpha}^*(0)$$

and from (3.15)

$$\dot{\alpha}^*(0) = \frac{-1}{(\lambda+\lambda_{12})^2} + \frac{-\lambda\, \dot{g}^*(\lambda+\lambda_{12})}{(\lambda+\lambda_{12})(2\lambda+\lambda_{12})} + g^*(\lambda+\lambda_{12}) \left[\frac{1}{(\lambda+\lambda_{12})^2} - \frac{1}{(2\lambda+\lambda_{12})^2}\right].$$

Further, the assumption that $g^*(s)$ exists for $s\geq 0$ guarantees that the repair time has finite moments, since the Laplace transform is analytic in the halfplane right of its abscissa of convergence. Hence $\dot{\alpha}^*(0)$ is finite.
On the other hand, it follows from (3.29) that

$$\varphi^*(0,\tau) = \int_0^\infty \varphi(t,\tau)dt \leq \text{E } X_{11}.$$

Thus both $\dot{\varphi}^*(0,\tau)$ and $\varphi^*(0,\tau)$ are finite. ∎

Analogously, let $\mathfrak{A}(\tau)$ denote the stationary joint availability for an interval of length τ, *i.e.*

$$\mathfrak{A}(\tau) = \lim_{t\to\infty} A_1(t,\tau),$$

then

$$\mathfrak{A}(\tau) = \lim_{s\downarrow 0} s\, A_1^*(s,\tau),$$

where

$$A_1^*(s,\tau) = \int_0^\infty e^{-st} A_1(t,\tau)dt.$$

It is easily seen that

$$\mathfrak{A}(\tau) = \frac{\psi^*(0,\tau)}{\text{E } X_{11}}, \tag{3.34}$$

if both $\dot{\psi}^*(0,\tau)$ and $\psi^*(0,\tau)$ are finite.
To prove that (3.34) holds, notice that from (3.32)

$$\psi(t,\tau) \leq \Pr\{X_{11}>t\}$$

and hence

$$\psi^*(0,\tau) = \int_0^\infty \psi(t,\tau)\, dt \leq \text{E } X_{11}.$$

Further, let $F_{11}(.)$ denote the distribution function of X_{11}, then

$$0 \le -\dot{\psi}^{*}(0,\tau) = \int_0^{\infty} t\ \psi(t,\tau)dt \le \int_0^{\infty} t\ \bar{F}_{11}(t)dt = -\dot{\bar{F}}_{11}^{*}(0).$$

Using $\bar{F}_{11}^{*}(s) = (1-f_{11}^{*}(s))/s$ and applying De l'Hôpital's rule twice yields

$$-\dot{\bar{F}}_{11}^{*}(0) = \frac{\ddot{f}_{11}^{*}(0)}{2}.$$

(Once more, note that $\dot{\bar{F}}_{11}^{*}(.)$ denotes the derivative of the Laplace transform of $\bar{F}_{11}(t)$ and *not* the Laplace transform of the derivative of $\bar{F}_{11}(t)$.)
Since the existence of $g^{*}(s)$ for $s \ge 0$ guarantees that X_{11} has finite moments, it follows from (3.8) that $\dot{\psi}^{*}(0,\tau)$ is finite. ■

3.6 Quasi-stationary distributions

The expressions (3.22)–(3.24) capture the stationary distribution of the process $\{X(t),t\ge0\}$. Here the stationary distribution of the process $\{X(t),t\ge0\}$, under the additional condition that the process has not visited state 2, is investigated. Such limiting distributions conditioned on an event whose probability tends to zero in the long run, are known as quasi-stationary distributions in the stochastic process literature (Cavender, 1978). Some of the earlier contributions to quasi-stationarity in the context of Markov chains, are found in Darroch *et al.* (1965, 1967) and subsequent developments are discussed in Seneta *et al.* (1967, 1985), Tweedie (1974), Pollak *et al.* (1986) and Ziedins (1987). The two major questions discussed in the literature in the context of quasi-stationarity concern the system's residual life time distribution and the stationary distribution of the process $\{X(t),t\ge0\}$, both conditioned on the event that absorption into the system's down state has not occurred.
The use of quasi-stationary distributions *in the context of reliability modelling* is of recent origin. One of the few attempts that have been made in this direction is that of Kalpakam *et al.* (1983). In the cited reference, the tail of the reliability function of a two-unit warm standby system is studied, *i.e.* the distribution of the residual lifetime at time t is investigated for $t\to\infty$, given that there was no system failure up to time t. Kalpakam *et al.* establish that this distribution is exponential for a two-unit warm standby system, under the assumption that the Laplace transforms of the life and repair time distribution are rational.

Since quasi-stationary distributions are conditioned on the fact that there was no earlier failure, they are useful for systems that rarely enter the failed state, but may experience repairs. In this section the conditional state limiting behaviour and the conditional residual life time distribution are investigated for the two-unit parallel redundant system with dependent units. The assumption made by Kalpakam *et al.* is relaxed, *i.e.* the results are derived for arbitrarily distributed repair times.
Let q(.) denote the limiting residual life time distribution and q_i (i=0,1) the quasi-stationary state probabilities, then formal definitions of q(.) and q_i are

$$q(x) = \lim_{t\to\infty} \Pr\{X(u)\neq 2,\ t\leq u\leq t+x \mid X(u)\neq 2,\ 0\leq u\leq t,\ E_1 \text{ at } t=0\},\ x\geq 0$$

and

$$q_i = \lim_{t\to\infty} \Pr\{X(t)=i \mid X(u)\neq 2,\ 0\leq u\leq t,\ E_1 \text{ at } t=0\},\ i=0,1.$$

Obviously

$$q(x) = 1 - \lim_{t\to\infty} \frac{R_1(t+x)}{R_1(t)} \tag{3.35}$$

and

$$q_i = \lim_{t\to\infty} \frac{\mathfrak{P}_{1,i}(t)}{R_1(t)},\ i=0,1 \tag{3.36}$$

where

$$\mathfrak{P}_{1,i}(t) = \Pr\{X(t)=i,\ X(u)\neq 2,\ 0\leq u\leq t \mid E_1 \text{ at } t=0\},\ i=0,1.$$

The primary objective is to study the limiting distributions in (3.35) and (3.36). The limiting distribution (3.35) is shown to be negative exponential. The results are obtained by using Laplace transform techniques to identify the dominating terms in the asymptotic expansions for (3.35) and (3.36).
The probabilities functions $\mathfrak{P}_{1,i}(t)$ are obtained by replacing $f_{11}^{(n)}(t)$ by $\tilde{f}_{11}^{(n)}(t)$ in (3.21):

$$\mathfrak{P}_{1,i}(t) = \gamma_i(t) + \sum_{n=1}^{\infty} \tilde{f}_{11}^{(n)}(t) * \gamma_i(t),$$

where

$$\gamma_0(t) = [g(t)\bar{F}_1(t)] * \bar{F}_0(t)$$

and

$$\gamma_1(t) = \bar{G}(t)\bar{F}_1(t).$$

The Laplace transforms are

$$\mathfrak{P}^*_{1,i}(s) = \frac{\gamma^*_i(s)}{1-\tilde{f}^*_{11}(s)}, \tag{3.37}$$

$$\gamma^*_0(s) = \frac{g^*(\lambda+\lambda_{12}+s)}{2\lambda+\lambda_{12}+s} \tag{3.38}$$

and

$$\gamma^*_1(s) = \frac{1-g^*(\lambda+\lambda_{12}+s)}{\lambda+\lambda_{12}+s}, \tag{3.39}$$

where $\tilde{f}^*_{11}(s)$ is given by (3.11). After substitution of (3.11), (3.38) and (3.39) in (3.37), explicit expressions for $\mathfrak{P}^*_{1,i}(s)$ are

$$\mathfrak{P}^*_{1,0}(s) = \frac{g^*(\lambda+\lambda_{12}+s)}{\lambda_{12}+s+2\lambda(1-g^*(\lambda+\lambda_{12}+s))}, \tag{3.40}$$

$$\mathfrak{P}^*_{1,1}(s) = \frac{(2\lambda+\lambda_{12}+s)\ (1-g^*(\lambda+\lambda_{12}+s))}{(\lambda+\lambda_{12}+s)\ (\lambda_{12}+s+2\lambda(1-g^*(\lambda+\lambda_{12}+s)))}. \tag{3.41}$$

Further, substitution of (3.11) and (3.15) in (3.16) yields

$$R^*_1(s) = \frac{2\lambda+\lambda_{12}+s-\lambda g^*(\lambda+\lambda_{12}+s))}{(\lambda+\lambda_{12}+s)\ (\lambda_{12}+s+2\lambda(1-g^*(\lambda+\lambda_{12}+s)))}. \tag{3.42}$$

For the moment attention will be concentrated on the asymptotic behaviour of $\mathfrak{P}_{1,0}(t)$, as $t\to\infty$. The asymptotic expansions of $\mathfrak{P}_{1,1}(t)$ and $R_1(t)$ are derived in exactly the same way.

With respect to (3.40), notice that

i. the pole δ with maximum real part of $\mathfrak{P}^*_{1,0}(s)$ is real (Widder, 1946, theorem 5b, p.58),

ii. for real $s\geq-(\lambda+\lambda_{12})$, the denominator $\mathfrak{D}^*_{1,0}(s)$ of $\mathfrak{P}^*_{1,0}(s)$ is strictly increasing in s,

iii. $\mathfrak{D}^*_{1,0}(-(\lambda+\lambda_{12})) = -\lambda$ and $\mathfrak{D}^*_{1,0}(-\lambda_{12}) = 2\lambda(1-g^*(\lambda)) > 0$.

Hence, $\delta\in(-(\lambda+\lambda_{12}),-\lambda_{12})$ and δ is the unique real valued root of $\mathfrak{D}^*_{1,0}(s)$ in the interval $(-(\lambda+\lambda_{12}),\infty)$. Obviously, δ is also pole with maximum real part of $\mathfrak{P}^*_{1,1}(s)$ and $R^*_1(s)$.

Moreover, since

$$\lim_{s\to\delta} (s-\delta)\, \mathfrak{P}^*_{1,0}(s) = \frac{g^*(\lambda+\lambda_{12}+\delta)}{1-2\lambda \dot{g}^*(\lambda+\lambda_{12}+\delta)},$$

$$\lim_{s\to\delta} (s-\delta)\, \mathfrak{P}^*_{1,1}(s) = \frac{(2\lambda+\lambda_{12}+\delta)\ (1-g^*(\lambda+\lambda_{12}+\delta))}{(\lambda+\lambda_{12}+\delta)\ (1-2\lambda \dot{g}^*(\lambda+\lambda_{12}+\delta))},$$

$$\lim_{s\to\delta} (s-\delta)\, R^*_1(s) = \frac{\lambda+\lambda_{12}+\delta\ +\ \lambda(1-g^*(\lambda+\lambda_{12}+\delta))}{(\lambda+\lambda_{12}+\delta)\ (1-2\lambda \dot{g}^*(\lambda+\lambda_{12}+\delta))}$$

and $g^*(s)$ is strictly decreasing in s, it follows that δ is a *simple* pole of $\mathfrak{P}^*_{1,0}(s)$, $\mathfrak{P}^*_{1,1}(s)$ and $R^*_1(s)$. Hence,

$$\mathfrak{P}_{1,i}(t) = c_i \exp(\delta t) + o(\exp(\delta t)),\ t\to\infty,\ i=0,1$$

and

$$R_1(t) = c_2 \exp(\delta t) + o(\exp(\delta t)),\ t\to\infty.$$

The coefficients c_i (i=0,1,2) are found from (3.37), respectively (3.16), as

$$c_i = L\ \gamma^*_i(\delta),\ i=0,1$$

and

$$c_2 = L\ \alpha^*(\delta),$$

where

$$L = \lim_{s\to\delta} \frac{s-\delta}{1-\tilde{f}^*_{11}(s)}.$$

Using De l'Hôpital's rule

$$L = \frac{-1}{\mathfrak{F}^*(\delta)},$$

where

$$\mathfrak{F}^*(\delta) = \frac{d}{ds}\tilde{f}^*_{11}(s)\ \Big|_{s=\delta}.$$

Thus

$$\mathfrak{P}_{1,i}(t) = -\frac{\gamma^*_i(\delta)}{\mathfrak{F}^*(\delta)} \exp(\delta t) + o(\exp(\delta t)),\ t\to\infty,\ i=0,1 \tag{3.43}$$

and

$$R_1(t) = -\frac{\alpha^*(\delta)}{\mathfrak{F}^*(\delta)}\exp(\delta t) + o(\exp(\delta t)),\ t\to\infty. \tag{3.44}$$

As a result

$$q_i = \frac{\gamma_i^*(\delta)}{\alpha^*(\delta)},\ i=0,1.$$

Substitution of (3.15), (3.38), (3.39) and noting that the pole δ satisfies

$$\lambda g^*(\lambda+\lambda_{12}+\delta)) = \lambda+\lambda_{12}+\delta+\lambda(1-g^*(\lambda+\lambda_{12}+\delta)) \tag{3.45}$$

finally yields

$$q_0 = \frac{\lambda+\lambda_{12}+\delta}{\lambda} \tag{3.46}$$

and

$$q_1 = -\frac{\lambda_{12}+\delta}{\lambda}. \tag{3.47}$$

Subsequently it is established that the quasi-stationary distribution $\mathbf{q}=(q_0,q_1)$ of the process $\{X(t),t\geq0\}$ is independent of the parameter λ_{12}. As $\mathbf{q}$ is the limiting distribution of the process conditioned on the information that the system has *not* visited the down state, this means in particular that *no* common cause failures have occurred. Hence, intuitively the rate at which common cause failures occur should not influence $\mathbf{q}$.

For a formal proof, consider the case where $\lambda_{12}=0$, *i.e.* the situation with statistically independent units. Let δ_{ind} be the root with maximum real part of the denominator of (3.40), then δ satisfies

$$\frac{2\lambda}{2\lambda+\delta_{ind}}\, g^*(\lambda+\delta_{ind}) = 1.$$

Thus, defining δ_{dep} as $\delta_{dep}=\delta_{ind}-\lambda_{12}$, it follows that

$$\frac{2\lambda}{2\lambda+\lambda_{12}+\delta_{dep}}\, g^*(\lambda+\lambda_{12}+\delta_{dep}) = 1.$$

Obviously, δ_{dep} is the root with maximum real part in the case with statistically dependent units. From $\delta_{dep}=\delta_{ind}-\lambda_{12}$ it is immediate that (3.46) and (3.47) are independent of λ_{12}. ∎

Further, it is clear from (3.44) that the tail of the life time distribution is negative exponentially distributed with parameter $-\delta$:

$$q(x) = 1 - \exp(\delta x).$$

Note that the parameter δ is *not* independent of λ_{12} as $\delta_{dep}=\delta_{ind}-\lambda_{12}$. Hence, the limiting residual life time distribution *does* depend upon the rate at which common cause failures occur.

On the other hand, the distributions $q(x)$ and $\mathbf{q}$ do not depend upon the initial condition: since $q(x)$ and $\mathbf{q}$ are conditioned on the information that the system does not visit state 2 before time t, only transitions from state 0 to 1 and *vice versa* occur before t and the first visit to state 1 will occur (with probability one) within finite time. However, an entry into state 1 is a regenerative event and as $t\to\infty$, the conclusion is that the quasi-stationary distributions $q(x)$ and $\mathbf{q}$ do not depend on the initial condition.

3.7 Intensity of events

In this section the number of point events generated by the process $\{X(t),t\geq 0\}$ is studied. The objective is to obtain expressions for the expected value, the variance and the covariance of the counting measures associated with the events E_i (i=0,1,2). Since an entrance into state 2 is regenerative if it occurs from state 0 and non-regenerative if it occurs from state 1, both types of events are distinguished hereafter and called respectively E_α and E_β events.

Let $N_i(t)$ be the counting measure associated with event E_i, $i=0,1,\alpha,\beta$, then the objective is to obtain $\mathbb{E}\, N_i(t)$, Var $N_i(t)$ and $\mathrm{Cov}(N_\alpha(t),N_\beta(t))$. The basic approach is to use the product densities associated with the events. The product densities are the counter parts of the renewal density in the context of non-renewal stochastic point processes and systematically discussed and illustrated in Srinivasan (1974) and Cox *et al.* (1980).

The analysis starts with E_1 events. The renewal density of E_1 events is used to obtain the expected value and the second order characteristics. The renewal density $h_1(t)$ is defined as

$$h_1(t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \Pr\{N_1(t+\Delta t)-N_1(t)=1 \mid E_1 \text{ at } t=0\}$$

and, for $t_1 \neq t_2$, the product density $h_1(t_1,t_2)$ is defined as

$$h_1(t_1,t_2) = \lim_{\Delta t_1, \Delta t_2 \to 0} \frac{1}{\Delta t_1 \Delta t_2} \Pr\{N_1(t_i+\Delta t_i)-N_1(t_i)=1,\ i=1,2 \mid E_1 \text{ at } t=0\}.$$

The functions $h_1(t)$ and $h_1(t_1,t_2)$ have a simple interpretation: for small Δt the quantity $h_1(t)\Delta t$ represents the probability of an occurrence of an E_1 event in $(t,t+\Delta t)$ and for small Δt_1 and Δt_2 $h_1(t_1,t_2)\Delta t_1 \Delta t_2$ represents the joint occurrence of E_1 events in $(t_1,t_1+\Delta t_1)$ and $(t_2,t_2+\Delta t_2)$. From the renewal nature of E_1 events and the possibility that the required event in $(t,t+\Delta t)$ may be the first or any subsequent event,

$$h_1(t) = \sum_{n=1}^{\infty} f_{11}^{(n)}(t), \quad t>0 \tag{3.48}$$

and

$$h_1(t_1,t_2) = h_1(t_1)\ h_1(t_2-t_1), \quad t_1<t_2. \tag{3.49}$$

For obvious reasons the roles of t_1 and t_2 are interchanged if $t_2<t_1$. The expected number of E_1 events in $(0,t)$ is obtained by integrating the function $h_1(u)$ over $(0,t)$:

$$\mathbf{E}\ N_1(t) = \int_0^t h_1(u)\ du. \tag{3.50}$$

Cox *et al.* (1980) give the second factorial moment as

$$\mathbf{E}\ [N_1(t)\ (N_1(t)-1)] = \int_{\substack{0<t_1,\, t_2 \le t \\ t_1 \neq t_2}} h_1(t_1,t_2)\ dt_1\ dt_2$$

and hence

$$\mathrm{Var}\ N_1(t) = \int_{\substack{0<t_1,\, t_2 \le t \\ t_1 \neq t_2}} h_1(t_1,t_2)\ dt_1\ dt_2 + \mathbf{E}\ N_1(t) - \mathbf{E}^2 N_1(t). \tag{3.51}$$

The stationary frequency Λ_1 of E_1 events is obtained from the limiting behaviour of $h_1(t)$:

$$\Lambda_1 = \lim_{t\to\infty} h_1(t). \tag{3.52}$$

Applying the key renewal theorem or, alternatively, Laplace transform techniques, yields $\Lambda_1 = 1/\mathbb{E}\, X_{11}$. Substitution of (3.9) finally gives

$$\Lambda_1 = \frac{2\lambda+\lambda_{12}}{\mu(2\lambda+\lambda_{12}) + g^*(\lambda+\lambda_{12})(1+\lambda_{12}\mu)}.$$

Further, let $N_1(t_1,t_2)$ denote the number of E_1 events in the interval (t_1,t_2). Then, in the stationary case the expected number of E_1 events in an interval of length τ is, from (3.52),

$$\lim_{t\to\infty} \mathbb{E}\, N_1(t,t+\tau) = \Lambda_1\tau.$$

To find the variance of the number of E_1 events in an interval of length τ in the stationary case, note that from (3.49) the asymptotic behaviour of $h_1(t_1,t_2)$ is

$$\lim_{t\to\infty} h_1(t,t+\tau) = \Lambda_1\, h_1(\tau).$$

Thus, using (3.51) and (3.49)

$$\lim_{t\to\infty} \mathrm{Var}\, N_1(t,t+\tau) = \Lambda_1\tau(1-\Lambda_1\tau) + 2\Lambda_1 \int_0^\tau m(u)\, du,$$

where $m(.)$ is the renewal function for X_{11}:

$$m(u) = \sum_{n=1}^{\infty} F_{11}^{(n)}(u). \tag{3.53}$$

Since $N_0(t)$ can be analysed similarly to $N_1(t)$, the counting process of the E_0 events is not examined here. A more interesting sequence of point events is generated by E_2 events. In the rest of this section the intensity and product density functions associated with E_2 events are derived and their limiting behaviour is considered. Subsequently, expectations, variances and their limiting behaviour can be obtained as shown above for the E_1 events.

Define $h_2(t)$, the intensity associated with E_2 events, as

$$h_2(t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \Pr\{N_2(t+\Delta t)-N_2(t)=1 \mid E_1 \text{ at } t=0\}.$$

Using arguments similar to those in the derivation of (3.12)–(3.14), it is immediate that

$$h_2(t) = a(t) + \sum_{n=1}^{\infty} f_{11}^{(n)}(t)*a(t), \tag{3.54}$$

with $a(t)$ defined as

$$a(t) = [g(t)\bar{F}_1(t)]*f_{02}(t) + \bar{G}(t)f_1(t). \tag{3.55}$$

Obviously, $a(t)\,dt$ represents the probability that the system enters state 2 in $(t,t+dt)$ whereas in the interval $(0,t)$ that began with an E_1 event, the system neither fails nor induces the further occurrence of an E_1 event. Actually, (3.55) is a defective pdf, as in an interval between two successive E_1 events not necessarily an E_2 event occurs. The terms in the right hand side of (3.55) correspond with the paths 1–0–2 and 1–2, *i.e.* with the occurrence of an E_α and E_β event. To distinguish between E_α and E_β events, the function $h_2(t)$ is decomposed and written as the sum of the intensity functions of E_α and E_β events, denoted by $h_\alpha(t)$, respectively $h_\beta(t)$. Clearly

$$h_2(t) = h_\alpha(t) + h_\beta(t),$$

where

$$h_\alpha(t) = a_\alpha(t) + \sum_{n=1}^{\infty} f_{11}^{(n)}(t)*a_\alpha(t),$$

$$h_\beta(t) = a_\beta(t) + \sum_{n=1}^{\infty} f_{11}^{(n)}(t)*a_\beta(t)$$

and

$$a_\alpha(t) = [g(t)\bar{F}_1(t)]*f_{02}(t),$$

$$a_\beta(t) = \bar{G}(t)f_1(t).$$

The key renewal theorem yields expressions for the stationary intensities Λ_2, Λ_α and Λ_β.

It is clear that

$$\Lambda_2 = \frac{1}{E\,X_{11}} \int_0^\infty a(u)\,du. \tag{3.56}$$

Laplace transform techniques, applied to (3.56), yield

$$\Lambda_2 = \frac{2\lambda+\lambda_{12}-2\lambda g^*(\lambda+\lambda_{12})}{\mu(2\lambda+\lambda_{12}) + g^*(\lambda+\lambda_{12})(1+\lambda_{12}\mu)}.$$

Analogously,

$$\Lambda_\alpha = \frac{\lambda_{12} g^*(\lambda+\lambda_{12})}{\mu(2\lambda+\lambda_{12}) + g^*(\lambda+\lambda_{12})(1+\lambda_{12}\mu)}$$

and

$$\Lambda_\beta = \frac{(2\lambda+\lambda_{12})(1-g^*(\lambda+\lambda_{12}))}{\mu(2\lambda+\lambda_{12}) + g^*(\lambda+\lambda_{12})(1+\lambda_{12}\mu)}.$$

It is immediate that $\Lambda_2=\Lambda_\alpha+\Lambda_\beta$.

The E_α events, generated as a special sequence of E_2 events are regenerative and the second order properties are derived in a similar way to those of the events E_1 and E_0. Details are not repeated here. However, the E_β events are non-regenerative and it is useful to obtain the second order properties of the counting process $N_\beta(t)$ from the appropriate product densities. To distinguish the second order product densities of E_α and E_β events from the second order *cross* product densities, a double index is used: for $t_1 \neq t_2$ $h_{\beta\beta}(t_1,t_2)$ is defined as

$$h_{\beta\beta}(t_1,t_2) = \lim_{\Delta t_1,\Delta t_2 \to 0} \frac{1}{\Delta t_1 \Delta t_2} \Pr\{N_\beta(t_i+\Delta t_i)-N_\beta(t_i)=1,\ i=1,2 \mid E_1 \text{ at } t=0\}.$$

Then

$$h_{\beta\beta}(t_1,t_2) = \bar{G}(t_1)\, f_1(t_1) \int_0^{t_2-t_1} \frac{g(t_1+u)}{\bar{G}(t_1)}\, h_\beta(t_2-t_1-u)du \tag{3.57}$$

$$+ \sum_{n=1}^{\infty} \int_0^{t_1} f_{11}^{(n)}(t_1-v)\, f_1(v)\, \bar{G}(v) \int_0^{t_2-t_1} \frac{g(v+u)}{\bar{G}(v)}\, h_\beta(t_2-t_1-u)du\,dv.$$

Expression (3.57) is derived by considering the following specifications. It is required to obtain an E_β event (a transition of state 1 to state 2) at t_1 and t_2 ($t_1 \neq t_2$) conditioned by an E_1 event at the origin. Calling the time between two successive E_1 events a cycle, the first term in the right hand side of (3.57) corresponds to an occurrence of an E_β event at t_1 in the first cycle and the second term to an occurrence at t_1 in a later cycle. The first term is derived as follows:

i. The E_β event realised at t_1 in the first cycle is caused by a repair which is not completed until t_1 and a failure of the operating unit at t_1. This explains the contribution $\bar{G}(t_1)\ f_1(t_1)$.

ii. The E_β event realised at t_1 is non-regenerative, but the elapsed repair time at t_1 is precisely the time until the last E_1 event. Thus, u units of time after the E_β event at t_1, the repair time has hazard rate $g(t_1+u)/\bar{G}(t_1)$. Hence the repair is completed in (t_1+u, t_1+u+du) with probability $g(t_1+u)/\bar{G}(t_1)\ du$.

iii. When the repair which began at the last E_1 event before t_1 is completed, a fresh E_1 event is generated and using this, the required E_β event at t_2 is identified as the first order product density in the appropriate interval.

In the second term in the right hand side of (3.57), the E_β event at t_1 is preceded by several E_1 events before t_1 while the last one occurred at t_1-v, so that v represents the elapsed repair duration of the unit under repair at time t_1. The rest of the analysis follows the above reasoning under *i–iii.*

Let the function $h_{\beta\beta}(\tau)$ describe the stationary behaviour of $h_{\beta\beta}(t_1,t_2)$, *i.e.*

$$h_{\beta\beta}(\tau) = \lim_{t\to\infty} h_{\beta\beta}(t,t+\tau).$$

Then

$$h_{\beta\beta}(\tau) = \lim_{t\to\infty} \sum_{n=1}^{\infty} \int_0^t f_{11}^{(n)}(t-v)\ f_1(v) \int_0^\tau g(u+v)\ h_\beta(\tau-u)\ du\ dv,$$

or, equivalently

$$h_{\beta\beta}(\tau) = \lim_{t\to\infty} \int_0^t \tilde{f}_\beta(\tau,t-v)\ dm(v),$$

where

$$\Gamma_\beta(\tau,v) = f_1(v) \int_0^\tau g(u+v)\ h_\beta(\tau-u)du.$$

As before, m(.) is the renewal function for X_{11}.
From the key renewal theorem,

$$h_{\beta\beta}(\tau) = \frac{1}{E\ X_{11}} \int_0^\infty \Gamma_\beta(\tau,v)dv.$$

To complete the analysis of the stochastic point events associated with E_2 events, expressions for the second order cross product densities are obtained. For $0<t_1<t_2$, the product density $h_{\alpha\beta}(t_1,t_2)$ is defined as

$$h_{\alpha\beta}(t_1,t_2)=\lim_{\Delta_1,\Delta_2\to 0} \frac{1}{\Delta_1\Delta_2}\Pr\{N_\alpha(t_1+\Delta_1)-N_\alpha(t_1)=1,\ N_\beta(t_2+\Delta_2)-N_\beta(t_2)=1 \mid E_1 \text{ at } t=0\}.$$

The function $h_{\beta\alpha}(t_1,t_2)$ is defined similarly by interchanging the role of α and β. From the renewal nature of the E_α events it immediately follows that

$$h_{\alpha\beta}(t_1,t_2) = h_\alpha(t_1) \int_0^{t_2-t_1} g(u)\ h_\beta(t_2-t_1-u)du \tag{3.58}$$

and defining

$$h_{\alpha\beta}(\tau) = \lim_{t\to\infty} h_{\alpha\beta}(t,t+\tau)$$

it is clear that

$$h_{\alpha\beta}(\tau) = \Lambda_\alpha \int_0^\tau g(u)\ h_\beta(\tau-u)\ du.$$

The observation that the E_β events are non-regenerative and every E_β event is followed by an E_1 event, yields, for $t_1<t_2$

$$h_{\beta\alpha}(t_1,t_2) = f_1(t_1)\ \bar{G}(t_1) \int_0^{t_2-t_1} \frac{g(t_1+v)}{\bar{G}(t_1)}\ h_\alpha(t_2-t_1-v)dv$$

$$+\sum_{n=1}^{\infty} \int_0^{t_1} f_{11}^{(n)}(t_1-v)\ f_1(v)\ \bar{G}(v) \int_0^{t_2-t_1} \frac{g(u+v)}{\bar{G}(v)}\ h_\alpha(t_2-t_1-u)\ du\ dv. \tag{3.59}$$

Further, let $h_{\beta\alpha}(\tau)$ describe the limiting behaviour of $h_{\beta\alpha}(t,t+\tau)$ for $t\to\infty$, then

$$h_{\beta\alpha}(\tau) = \frac{1}{\mathsf{E}\, X_{11}} \int_0^{\infty} \tilde{f}_{\alpha}(\tau,v)\; dv,$$

where

$$\tilde{f}_{\alpha}(\tau,v) = f_1(v) \int_0^{\tau} g(u+v)\; h_{\alpha}(\tau-u)\; du.$$

The above relations summarise the properties of the counting measures associated with the E_i events. The covariance function of the counting processes associated with E_α and E_β events connects several counting measures:

$$\mathrm{Cov}[N_\alpha(t),N_\beta(t)] = \int_0^t \int_0^v \{h_{\alpha\beta}(u,v)+h_{\beta\alpha}(u,v)\}\; du\; dv\; - \mathsf{E}\, N_\alpha(t)\; \mathsf{E}\, N_\beta(t).$$

Finally, the covariance function in the stationary case, over an interval of length τ, is given by

$$\lim_{t\to\infty} \mathrm{Cov}[N_\alpha(t,t+\tau),N_\beta(t,t+\tau)] = \int_0^{\tau} \int_0^v \{h_{\alpha\beta}(v-u)+h_{\beta\alpha}(v-u)\}\; du\; dv\; - \Lambda_\alpha \Lambda_\beta \tau^2.$$

The above analysis of the intensity functions concludes the stochastic analysis of the two-unit dependent parallel system with *identical* units. In this chapter, the joint life time distribution of the units was a BVE and the repair times were generally distributed. Expressions are obtained for a number of performance measures, two quasi-stationary distributions are studied and the intensity functions of E_0, E_1 and E_2 events are analysed. A key role in the analysis is played by E_1 events. The regenerative nature of the E_1 events is used to construct an imbedded renewal process, which is exploited to compute the system's performance measures. Expressions for the *transient* behaviour are given by convolution integrals and the *stationary* measures of the system with identical units are obtained by applying Laplace transform techniques. The methods and techniques in this chapter are illustrative for the study and will be applied in the chapters 4 and 5 to investigate the situation with *non-identical* units. In chapter 4, the life and repair times are restricted to the class of phase type distributions (Neuts, 1981), which render the state description process Markovian, whereas in chapter 5 this assumption is relaxed.

4. PHASE TYPE DISTRIBUTED LIFE AND REPAIR TIMES

4.1 Introduction

In the previous chapter a parallel system was analysed with bivariate exponentially distributed life times and general repair times. In this case the time dependent behaviour of the system is captured by a set of convolution integrals and the Laplace transform technique is used to describe the stationary behaviour. The computation of time dependent measures as the reliability function (3.13), the availability function (3.14) or the system state probabilities (3.21) can be a tough and time consuming job. However, for some particular choices of the repair time distribution, the stochastic behaviour of the state description process becomes Markovian and then the time dependent performance measures can be computed relatively easy. In the case of a Markov process (MP), the stochastic behaviour of the process is completely determined by the generator of the MP.

In this chapter the concept of phase type (PH) distributions (Neuts, 1981) is introduced and it is shown how they can be used to approximate the general situation where the life and repair times of the units are arbitrarily distributed. Since the use of phase type distributions results in an MP, the method is useful in particular when special interest is in the time dependent behaviour of a system.

In section 4.2 the life times of the units follow a BVE and the repair times are assumed to be exponentially distributed. The generator of the resulting MP is given and the performance measures described in section 1.3 are derived. The case of exponentially distributed repair times is simple but the techniques illustrate the approach in the remainder of this chapter. In fact the only modification to be made in the other sections is to replace the generator of the MP in 4.2. In section 4.3 PH distributions are introduced and in section 4.4 the repair times are phase type, while the life times follow a BVE. A generalisation is made towards the system's life time distribution in section 4.5, where the bivariate phase type distribution (BVPH) is constructed. In section 4.6, the life times have a BVPH and the repair times a PH distribution. Finally, in section 4.7 the randomisation technique is described. The technique is used to handle Markov processes with a large state space and is suitable for computing time dependent measures, even when the number of system states becomes very large.

4.2 BVE life times and exponential repair times

Consider the model described in chapter 3, where the units are identical and follow the BVE life time distribution (3.1) with $\lambda_1=\lambda_2=\lambda$. Let the state description process $\{X(t),t\geq 0\}$ denote the number of failed units at time t and assume that the repair times are exponentially distributed with parameter μ. Then the observation that the BVE has exponential marginals implies that the process $\{X(t),t\geq 0\}$ is an MP and obviously the generator Q of this MP is

$$Q = \begin{array}{c} \\ 0 \\ 1 \\ 2 \end{array} \begin{array}{c} \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ \begin{bmatrix} -(2\lambda+\lambda_{12}) & 2\lambda & \lambda_{12} \\ \mu & -(\lambda+\lambda_{12}+\mu) & \lambda+\lambda_{12} \\ 0 & \mu & -\mu \end{bmatrix} \end{array}.$$

The above case can be extended straightforwardly to the situation where the units are not identical. For nonidentical units, the repair time of unit i is assumed to be exponentially distributed with parameter μ_i, while the life times follow the BVE (3.1). It is easily seen that the state description process $\{X(t),t\geq 0\}$ with state space $S_X=\{(o,o),(r,o),(o,r),(r,w),(w,r)\}$ is an MP with generator

$$Q = \begin{array}{c} \\ (o,o) \\ (r,o) \\ (o,r) \\ (r,w) \\ (w,r) \end{array} \begin{array}{c} \begin{array}{ccccc} (o,o) & (r,o) & (o,r) & (r,w) & (w,r) \end{array} \\ \left[\begin{array}{ccc|cc} -(\lambda_1+\lambda_2+\lambda_{12}) & \lambda_1 & \lambda_2 & p\lambda_{12} & (1-p)\lambda_{12} \\ \mu_1 & -(\lambda_2+\lambda_{12}+\mu_1) & 0 & \lambda_2+\lambda_{12} & 0 \\ \mu_2 & 0 & -(\lambda_1+\lambda_{12}+\mu_2) & 0 & \lambda_1+\lambda_{12} \\ \hline 0 & 0 & \mu_1 & -\mu_1 & 0 \\ 0 & \mu_2 & 0 & 0 & -\mu_2 \end{array}\right] \end{array}.$$

As before, p is the probability that unit 1 is repaired first when a common cause failure occurs.

In this section it is shown how the performance measures of the system can be derived, applying standard Markov theory. The analysis is performed for the simple Markov process of the system with nonidentical units. The only modification that has to be made in the more complicated situations considered in the sections 4.4 and 4.5, is the substitution of an appropriate generator.

The main advantage of using MP's is the computational ease of obtaining the performance measures, caused by the lack-of-memory property of the exponential distribution and the fact that an MP is nothing but a sequence of exponentially distributed phases, with a transition mechanism which is governed by the underlying Markov chain. In the subsequent sections the concept of exponential phases is used to approximate arbitrary distribution functions. Following this approach, a system with generally distributed life and repair times can be approximated by an MP, with all its computational and analytical advantages.

To start with, the system state probabilities are considered. Let $P(t)$ denote the row vector with the time dependent state probabilities $P_i(t)$ ($i \in S_X$) and p_0 the initial probability vector, *i.e.* $p_0(i)$ (the i^{th} component of p_0) represents the probability of starting in state i at $t=0$. Then the time dependent behaviour of the system is governed by the Chapman-Kolmogorov differential equations and solving them yields (Çinlar, 1975)

$$P(t) = p_0 \exp(Qt), \; t \geq 0.$$

The stationary distribution of the process $\{X(t), t \geq 0\}$ is obtained by setting the time derivative in the Chapman-Kolmogorov equations to zero. As a result, the stationary state probabilities π_i satisfy (Çinlar, 1975)

$$\pi \, Q = 0.$$

Together with the normalisation equation $\sum \pi_i = 1$, the distribution π is uniquely determined.

The reliability function is obtained by lumping the system down states (r,w) and (w,r) into one absorbing state, denoted as 'abs'. The result is an MP with generator Q_a, where

$$Q_a = \begin{array}{c} \\ (o,o) \\ (r,o) \\ (o,r) \\ abs \end{array} \begin{array}{c} \begin{array}{cccc} (o,o) & (r,o) & (o,r) & abs \end{array} \\ \left[\begin{array}{ccc|c} -(\lambda_1+\lambda_2+\lambda_{12}) & \lambda_1 & \lambda_2 & \lambda_{12} \\ \mu_1 & -(\lambda_2+\lambda_{12}+\mu_1) & 0 & \lambda_2+\lambda_{12} \\ \mu_2 & 0 & -(\lambda_1+\lambda_{12}+\mu_2) & \lambda_1+\lambda_{12} \\ \hline 0 & 0 & 0 & 0 \end{array}\right] \end{array} .$$

It is clear that the system reliability equals the distribution function of the entrance time into the absorbing state. Hence,

$$R(t) = p_0^r \exp(Q_r t)\ e_r,$$

where p_0^r is the row vector containing the initial probabilities of starting in a system up state, e_r is a unit column vector of appropriate dimension and Q_r the submatrix of Q_a describing the transient behaviour until absorption, *i.e.*

$$Q_r = \begin{array}{c} \\ (o,o) \\ (r,o) \\ (o,r) \end{array} \begin{array}{c} \begin{array}{ccc} (o,o) & (r,o) & (o,r) \end{array} \\ \begin{bmatrix} -(\lambda_1+\lambda_2+\lambda_{12}) & \lambda_1 & \lambda_2 \\ \mu_1 & -(\lambda_2+\lambda_{12}+\mu_1) & 0 \\ \mu_2 & 0 & -(\lambda_1+\lambda_{12}+\mu_2) \end{bmatrix} \end{array}.$$

To prevent misunderstandings: a unit vector is defined here (and in chapter 5) as a vector which elements are all ones.
The MTSF is found after integration of the reliability function as

$$\text{MTSF} = -\ p_0^r\, Q_r^{-1}\ e_r.$$

Obviously, the availability function is obtained as the summation of the system state probabilities over the set of system up states:

$$A(t) = \sum_{i \in S_{up}} P_i(t),$$

where $S_{up}=\{(o,o),(o,r),(r,o)\}$.
Similarly, the stationary availability A is

$$A = \sum_{i \in S_{up}} \pi_i.$$

From the lack-of-memory property of the exponential distribution, the interval reliability $R(t,\tau)$ can be decomposed into the probability of being up at time t and surviving an additional τ units of time:

$$R(t,\tau) = P_r(t)\ \exp(Q_r\tau)\ e_r, \tag{4.1}$$

where $P_r(t)$ is the row vector with the system state probabilities $P_i(t), i \in S_{up}$.

Clearly, the stationary interval reliability $\mathfrak{R}(\tau)$ is found from (4.1) as

$$\mathfrak{R}(\tau) = \pi_r \exp(Q_r\tau)\, e_r,$$

where π_r is defined analogously to $P_r(t)$.

In the same way as the interval reliability, the joint availability $A(t,\tau)$ can be decomposed into the probability of being available at time t and being available τ units of time later:

$$A(t,\tau) = P_a(t) \exp(Q\tau)\, e_a, \tag{4.2}$$

where the vector $P_a(t)$ is obtained from $P(t)$, replacing the system down probabilities by zeros. Similarly, e_a is a unit column vector with zeros in the positions representing down states. From (4.2), the stationary joint availability $\mathfrak{A}(\tau)$ is

$$\mathfrak{A}(\tau) = \pi_a \exp(Q\tau)\, e_a.$$

As in section 3.6, the quasi-stationary state probability vector $\mathbf{q}=\{q_i\}_{i\in S_{up}}$ and the limiting residual life time distribution $q(.)$ are

$$\mathbf{q} = \lim_{t\to\infty} \frac{\mathfrak{P}(t)}{R(t)} \tag{4.3}$$

and

$$q(x) = 1 - \lim_{t\to\infty} \frac{R(t+x)}{R(t)}, \tag{4.4}$$

where

$$\mathfrak{P}(t) = p_0^r \exp(Q_r t).$$

In order to get an asymptotic expansion for $\mathfrak{P}(t)$ and $R(t)$, the Perron–Frobenius theorem (Seneta, 1973) is applied to express $\exp(Q_r t)$ in terms of the dominant eigenvalue of Q_r. Write

$$Q_r = S^{-1}\, D\, S,$$

where D is a diagonal matrix with the eigenvalues of the matrix Q_r if Q_r is diagonalisable and a Jordan normal matrix in the other case. Then

$$\exp(Q_r t) = S^{-1} \exp(Dt)\, S$$

and $\exp(Q_r t)$ is nonnegative and substochastic.

Applying the Perron–Frobenius theorem to $\exp(Q_r t)$, it follows that the eigenvalue ρ with maximum real part of Q_r is real, simple and negative. The eigenvalue ρ being simple, yields

$$\exp(Q_r t) = wv^T \exp(\rho t) + o(\exp(\tilde{\rho} t)), \quad t\to\infty, \tag{4.5}$$

where w and v are the right, respectively left eigenvector corresponding to ρ and $\tilde{\rho}$ is the eigenvalue with second largest maximum real part of Q_r. Substitution of (4.5) in (4.3) gives

$$\mathbf{q} = \frac{v}{v^T e_r}. \tag{4.6}$$

It is clear that the distribution determined by (4.6) is proper, since the vector v is strictly positive (or strictly negative) by the Perron–Frobenius theorem.

Besides, the quasi-stationary distribution $\mathbf{q}$ of the process $\{X(t),t\geq 0\}$ is independent of λ_{12}. To provide a formal proof, define

$$Q_{ind} = \begin{bmatrix} -(\lambda_1+\lambda_2) & \lambda_1 & \lambda_2 \\ \mu_1 & -(\lambda_2+\mu_1) & 0 \\ \mu_2 & 0 & -(\lambda_1+\mu_2) \end{bmatrix}$$

and

$$Q_{dep} = \begin{bmatrix} -(\lambda_1+\lambda_2+\lambda_{12}) & \lambda_1 & \lambda_2 \\ \mu_1 & -(\lambda_2+\lambda_{12}+\mu_1) & 0 \\ \mu_2 & 0 & -(\lambda_1+\lambda_{12}+\mu_2) \end{bmatrix}.$$

Since $Q_{dep} = Q_{ind} - \lambda_{12} I$ (where I is an identity matrix of appropriate dimension), it is immediate that

i. ρ is eigenvalue of $Q_{ind} \Leftrightarrow \rho-\lambda_{12}$ is eigenvalue of Q_{dep}

ii. v is eigenvector corresponding to ρ in the independent case
$\Leftrightarrow$ v is eigenvector corresponding to $\rho-\lambda_{12}$ in the dependent case.

Hence, Q_{dep} and Q_{ind} have identical eigenvectors. Consequently $\mathbf{q}$ is independent of λ_{12}. ■

On the other hand, from substitution of (4.5) in (4.4), it follows that the quasi-stationary distribution q(x) is negative exponential with parameter $-\rho$:

$$q(x) = 1-\exp(\rho x), \quad x \geq 0. \tag{4.7}$$

Note that the eigenvalue ρ with maximum real part of the matrix Q is *not* independent of λ_{12}, as $\rho_{dep}=\rho_{ind}-\lambda_{12}$, where ρ_{ind} (ρ_{dep}) is the eigenvalue with maximum real part of the system with independent (dependent) units.

4.3 Phase type distributions

Section 4.2 clearly shows the advantages of the use of exponential distributions. Obviously, the derivation of the performance measures is relatively easy, compared to the situation in chapter 3, where the repair times are arbitrarily distributed. However, the use of exponential distributions is restrictive and unsatisfactory. Hence, the dilemma is to use exponential distributions, which render the state description process Markovian and make the analysis relatively simple or to assume generally distributed repair times, which results in a more complicated analysis. The compromise is in terms of phase type (PH) distributions. The definition of a PH distribution used here is the definition of Neuts (1981), *i.e.* a probability distribution is of phase type if it is the distribution of the time until absorption in a Markov process with a finite number of transient states and one absorbing state. The notation used by Neuts is the following. Let Q represent the generator of the associated absorbing Markov process on the states 1,..,n+1 (state n+1 is the absorbing state), *i.e.*

$$Q = \begin{bmatrix} T & T^0 \\ \mathbf{0} & 0 \end{bmatrix},$$

where T is an nxn matrix, T^0 an nx1 vector and $\mathbf{0}$ a 1xn vector. The vector T^0 contains the transition rates into the absorbing state and hence $Te+T^0=0$, where e is a unit column vector of dimension n. The n-dimensional vector α is used to denote the initial state probabilities assigned to the transient states: the Markov process starts with probability $\alpha(i)$ in state i, where $\alpha(i)$ denotes the i^{th} component of α. A PH distribution F(.) is represented by the pair (α,T) and n is called its dimension. It will be supposed in the following chapters that the representation (α,T) is irreducible (Neuts, 1981, p.49).

Neuts shows that without loss of generality, attention may be restricted to irreducible representations. He proves that from any reducible representation of a phase type distribution F(.), a smaller, irreducible representation can be obtained by deleting appropriate rows and columns of the generator Q.

A formal expression for the probability distribution F(.) of the time until absorption is

$$F(t) = 1 - \alpha^T \exp(Tt)\ e, \quad t \geq 0$$

and its pdf f(.) is

$$f(t) = -\ \alpha^T \exp(Tt)\ T^0.$$

Hence, the Laplace transform $f^*(s)$ of f(t) is

$$f^*(s) = \alpha^T (sI-T)^{-1} T^0$$

and $f^*(s)$ is a rational function.

The central moments μ_k of F(.) are all finite and given by

$$\mu_k = (-1)^k\ k!\ (\alpha^T T^{-k} e).$$

Since a PH distribution represents the distribution of the sum of a random number of exponentially (but in general not identically) distributed phases, the lack-of-memory property applies locally, *i.e.* per phase. As a result, PH repair times render the state description process Markovian, when it is extended to include the phases of the repair times and when the life times follow the BVE (3.1). In that case, the system can be analysed applying standard Markov theory.

Another important property of PH distributions is that they are dense in the class of distribution functions. Consequently, *any* repair time distribution can be approximated with arbitrary precision by a PH distribution. Methods to fit PH distributions are described in Bux *et al.* (1977), Kao (1988) and Johnson *et al.* (1989, 1990a, 1990b).

The class of PH distributions contains, among others, the hyper exponential distribution, the generalised Erlang and the phase type distributions considered by Cox (1955^b) and Tijms (1986). Tijms uses mixtures of Erlang-k and Erlang-(k-1) distributions, which are not dense in the class of probability distributions, but which are useful in many practical applications. However, Neuts' PH distributions are preferred in this study, since they are more general than those of Tijms and secondly, Tijms' PH distributions have a representation with a generator which has multiple eigenvalues. In chapter 5 it will be shown how the latter property complicates the numerical computation of the system's stationary performance measures.

4.4 BVE life times and PH repair times

Suppose that unit i has a PH repair time distribution with representation (β_i,S_i) and dimension m_i and assume that the life times follow the BVE (3.1). Then the process under consideration is an MP, if the phases of the repair time distributions are included in the state description process.
To incorporate the repair time phases in the state description process $\{X(t),t\geq 0\}$, define the state space S_X as

$$S_X = (o,o)\cup(r,o)\cup(o,r)\cup(r,w)\cup(w,r)$$

where the elements of S_X are the following sets of states:

$$\begin{aligned}(r,o) &= \{(r_i,o)\}_{1\leq i\leq m_1},\\ (o,r) &= \{(o,r_i)\}_{1\leq i\leq m_2},\\ (r,w) &= \{(r_i,w)\}_{1\leq i\leq m_1},\\ (w,r) &= \{(w,r_i)\}_{1\leq i\leq m_2}.\end{aligned}$$

The notation r_i is used to denote that a component is in phase i of its repair time distribution. Further, let I_i be the identity matrix of dimension m_i, let '$\otimes$' denote the Kronecker product of matrices (Marcus *et al.*, 1964) and assume that S_i^0 satisfies $S_ie_i+S_i^0=0$, where e_i is a unit column vector of dimension m_i. Then the generator Q of the resulting MP with state space S_X is

$$Q = \begin{array}{c} \\ (o,o) \\ (r,o) \\ (o,r) \\ (r,w) \\ (w,r) \end{array} \begin{array}{c} \begin{array}{ccccc} (o,o) & (r,o) & (o,r) & (r,w) & (w,r) \end{array} \\ \left[\begin{array}{ccc|cc} -(\lambda_1+\lambda_2+\lambda_{12}) & \lambda_1\beta_1^T & \lambda_2\beta_2^T & p\lambda_{12}\beta_1^T & (1-p)\lambda_{12}\beta_2^T \\ S_1^0 & S_1-(\lambda_2+\lambda_{12})I_1 & 0 & (\lambda_2+\lambda_{12})I_1 & 0 \\ S_2^0 & 0 & S_2-(\lambda_1+\lambda_{12})I_2 & 0 & (\lambda_1+\lambda_{12})I_2 \\ \hline 0 & 0 & S_1^0\otimes\beta_2^T & S_1 & 0 \\ 0 & \beta_1^T\otimes S_2^0 & 0 & 0 & S_2 \end{array}\right] \end{array}.$$

The entries of the generator Q are obtained as follows. Transitions from state (o,o) to (r,o) occur with rate λ_1 and the repair of unit 1 starts with probability $\beta_1(i)$ in phase i. Secondly, the vector S_1^0 contains the transition rates from state (r,o) to (o,o). Thirdly, when the system is in state (w,r), the repair of unit 2 is completed with rate $S_2^0(i)$ if the repair is in phase i and at repair completion unit 1 starts in phase j of its repair with probability $\beta_1(j)$.

With the above generator, the performance measures of the system are computed as in section 4.2.

4.5 The bivariate phase type distribution

As mentioned in the previous sections, the advantages of phase type distributed repair times are:

i. PH type distributions are dense in the class of distribution functions and

ii. they render the state description process Markovian when the life times follow a BVE.

Hence, in principle arbitrarily distributed repair times can be handled, approximating them by PH distributions. However, an unsatisfactory state of affairs is that the life times of the components of the model in section 1.4 are exponentially distributed when a BVE is used. Given the interpretation of the BVE in section 3.2 in terms of a shock model, a logical extension is to consider shocks which are generated by PH distributions. When the life time of component C_i has a PH distribution with representation (α_i,T_i) and dimension n_i (i=1,2,3), the joint survival function $\bar{F}(t_1,t_2)$ of two new units is easily derived as

$$\bar{F}(t_1,t_2) = \alpha^T \exp(T_1 \otimes I_2 \otimes I_3 t_1 + I_1 \otimes T_2 \otimes I_3 t_2 + I_1 \otimes I_2 \otimes T_3 \max(t_1,t_2))\ e,$$

where $\alpha=\alpha_1\otimes\alpha_2\otimes\alpha_3$, $e=e_1\otimes e_2\otimes e_3$ and e_i is a unit column vector of dimension n_i.

The probability distribution $F(t_1,t_2)$ is called a bivariate phase type distribution (BVPH).

The BVPH has PH marginals: denoted by $\bar{F}_1(.)$ and $\bar{F}_2(.)$, the marginal survival functions are

$$\bar{F}_1(t) = \alpha_1^T \otimes \alpha_3^T \exp((T_1 \otimes I_3 + I_1 \otimes T_3)t)\ e_1 \otimes e_3$$

and

$$\bar{F}_2(t) = \alpha_2^T \otimes \alpha_3^T \exp((T_2 \otimes I_3 + I_2 \otimes T_3)t)\ e_2 \otimes e_3.$$

Since the lack-of-memory property applies per phase, property (3.2) holds, conditioned on the phase number of the life time distributions. Formally, let X_i denote the life time of unit i (i=1,2) and $X_i(t)$ the phase of the life time distribution of component C_i (i=1,2,3) at time t, then

$$\Pr\{X_1>t_1+s,\ X_2>t_2+s \mid X_1(s)=k_1,\ X_2(s)=k_2,\ X_3(s)=k_3\} =$$
$$\Pr\{X_1>t_1,\ X_2>t_2 \mid X_1(0)=k_1,\ X_2(0)=k_2,\ X_3(0)=k_3\}. \tag{4.8}$$

The construction of the BVPH implies that, in combination with PH repair times, the state description process $\{X(t),t\geq 0\}$ can be modelled as an MP, as will be shown in the next section.

4.6 BVPH life times and PH repair times

Suppose the life times of the parallel system follow a BVPH and the repair times a PH distribution. Let the life time distribution of component C_i have representation (α_i,T_i) and dimension n_i (i=1,2,3) and the repair time representation (β_i,S_i) and dimension m_i (i=1,2).

In order to model the system by an MP, the state space is extended to include the phases of the life and repair time distributions. Ordered triples are used to denote the phases of the components C_1, C_2 and C_3 respectively. Phase i of a component's life time is denoted by ℓ_i, phase j of a repair time by r_j and w is used to denote that a component is waiting for repair. Hence the state space S_X is

$$S_X = (o,o) \cup (r,o) \cup (o,r) \cup (r,w) \cup (w,r),$$

where

$$
\begin{aligned}
(o,o) &= \{(\ell_i,\ell_j,\ell_k)\}_{1\le i\le n_1,1\le j\le n_2,1\le k\le n_3},\\
(r,o) &= \{(r_i,\ell_j,\ell_k)\}_{1\le i\le m_1,1\le j\le n_2,1\le k\le n_3},\\
(o,r) &= \{(\ell_i,r_j,\ell_k)\}_{1\le i\le n_1,1\le j\le m_2,1\le k\le n_3},\\
(r,w) &= \{(r_i,w,w)\}_{1\le i\le m_1},\\
(w,r) &= \{(w,r_j,w)\}_{1\le j\le m_2}.
\end{aligned}
$$

Let the above sets of triples be ordered lexicographically on the indices of the phases of the life and repair times, then the generator Q of the Markov process under consideration is

$$
Q = \begin{array}{c|ccc|cc|}
 & (o,o) & (r,o) & (o,r) & (r,w) & (w,r) \\
\hline
(o,o) & \begin{matrix} T_1\otimes I_2\otimes I_3+ \\ I_1\otimes T_2\otimes I_3+ \\ I_1\otimes I_2\otimes T_3 \end{matrix} & T_1^0\beta_1^T\otimes I_2\otimes I_3 & I_1\otimes T_2^0\beta_2^T\otimes I_3 & p_1 e_1\beta_1^T\otimes e_2\otimes T_3^0 & p_2 e_1\otimes e_2\beta_2^T\otimes T_3^0 \\
(r,o) & S_1^0\alpha_1^T\otimes I_2\otimes I_3 & \begin{matrix} S_1\otimes I_2\otimes I_3+ \\ I_4\otimes T_2\otimes I_3+ \\ I_4\otimes I_2\otimes T_3 \end{matrix} & 0 & \begin{matrix} I_4\otimes T_2^0\otimes e_3+ \\ I_4\otimes e_2\otimes T_3^0 \end{matrix} & 0 \\
(o,r) & I_1\otimes S_2^0\alpha_2^T\otimes I_3 & 0 & \begin{matrix} T_1\otimes I_5\otimes I_3+ \\ I_1\otimes S_2\otimes I_3+ \\ I_1\otimes I_5\otimes T_3 \end{matrix} & 0 & \begin{matrix} T_1^0\otimes I_5\otimes e_3+ \\ e_1\otimes I_5\otimes T_3^0 \end{matrix} \\
\hline
(r,w) & 0 & 0 & S_1^0\alpha_1^T\otimes\beta_2^T\otimes\alpha_3^T & S_1 & 0 \\
(w,r) & 0 & \beta_1^T\otimes S_2^0\alpha_2^T\otimes\alpha_3^T & 0 & 0 & S_2 \\
\hline
\end{array}
$$

where:

i. $I_1,..,I_5$ are identity matrices of respective orders n_1,n_2,n_3,m_1,m_2,

ii. $e_1,..,e_5$ are unit column vectors of respective orders n_1,n_2,n_3,m_1,m_2,

iii. T_i^0 satisfies $T_ie_i+T_i^0=0$ $(i=1,2,3)$ and S_i^0 satisfies $S_ie_{i+3}+S_i^0=0$ $(i=1,2)$.

The entries of Q are obtained as follows. The elements of $T_1\otimes I_2\otimes I_3+I_1\otimes T_2\otimes I_3+I_1\otimes I_2\otimes T_3$ correspond to the case where only a change in the phase of the life time of one of the components C_i $(i=1,2,3)$ occurs. The elements of $I_1\otimes T_2^0\beta_2^T\otimes I_3$ correspond to a failure of component C_2 while the other components do not change of phase. Being in phase i of its life time, C_2 fails with rate $T_2^0(i)$ and subsequently starts with probability $\beta_2(j)$ in phase j of its repair time distribution. The entry $p_1e_1\beta_1^T\otimes e_2\otimes T_3^0$ follows from a failure of the common cause component C_3, while unit 1 is the first to be repaired. Finally, the elements of $S_1^0\alpha_1^T\otimes\beta_2\otimes\alpha_3$ are obtained from a repair completion of C_1. After repair completion, C_1 is restarted in phase i of its life time distribution with probability $\alpha_1(i)$, the repair of C_2 is started in phase j with probability $\beta_2(j)$ and C_3 is restarted in phase k with probability $\alpha_3(k)$. The other elements of Q are interpreted similarly.

In view of the high order of the generator Q, it is essential to use its special structure to evaluate the above quantities, as in Neuts *et al.* (1981).

In a few situations an important reduction of the state space can be obtained. Four of them are mentioned here.

i. If C_3 has an exponential life time distribution with parameter λ, then $T_3=-\lambda$, $T_3^0=\lambda$, $I_3=e_3=\alpha_3=1$. Further, as the life time distribution of C_3 has only one phase, the state sets are reduced to pairs:

$$\begin{aligned}
(o,o) &= \{(\ell_i,\ell_j)\}_{1\le i\le n_1,1\le j\le n_2}\\
(o,r) &= \{(\ell_i,r_j)\}_{1\le i\le n_1,1\le j\le m_2}\\
(r,o) &= \{(r_i,\ell_j)\}_{1\le i\le m_1,1\le j\le n_2}\\
(r,w) &= \{(r_i,w)\}_{1\le i\le m_1}\\
(w,r) &= \{(w,r_j)\}_{1\le j\le m_2}
\end{aligned}$$

ii. If the units of the parallel system are identical, then the components C_1 and C_2 are identical. The state space of the process can be reduced, since the states (ℓ_i,ℓ_j,ℓ_k) and (ℓ_j,ℓ_i,ℓ_k) are identical, as are the states (ℓ_i,r_j,ℓ_k) and (r_j,ℓ_i,ℓ_k) and the states (r_i,w,w) and (w,r_i,w). The generator of the Markov process is obtained from the generator Q by
- lumping identical states, *i.e.* by adding the columns of these states,
- removing one of the rows of every pair of identical states.

iii. When the units are identical and C_3 has an exponential distribution, then a combination of *i* and *ii* occurs.

iv. When the lifetimes of the units are identically *exponentially* distributed and C_3 has an exponential distribution, then the system has a BVE as lifetime distribution. The generator of the MP is given in 4.2.

Obviously, the number of system states N of the MP is

$$N = n_1n_2n_3 + m_1n_2n_3 + n_1m_2n_3 + m_1 + m_2.$$

Hence, the number of states can become large, even if the life and repair time distributions have a small number of phases. In this case classical techniques for the numerical computation of *e.g.* the time dependent state probabilities fail. However, a special technique, called the randomisation technique (Grassmann, 1977, Gross *et al.* 1984), can be used to compute the time dependent behaviour of MP's with huge generators. The technique is useful, in particular when the generator of the MP is sparse.

4.7 The randomisation technique

The Kolmogorov forward differential equations for the time dependent state probabilities in an MP are (section 2.4)

$$\frac{dP(t)}{dt} = P(t)\ Q, \tag{4.9}$$

with solution

$$P(t) = p_0 \exp(Qt). \tag{4.10}$$

The Runge–Kutta method (Gerald, 1978) is a classical numerical integration technique for (4.9) and a common method of computing (4.10) is based on the spectral representation of Q (Çinlar, 1975). Both methods are suitable when the number of system states is relatively small, but severe round off errors arise when the state space is large. In the latter case the Runge–Kutta method has the disadvantage that it is a formal numerical analysis technique, which ignores any exploitable probabilistic structure of the problem. On the other hand, evaluation of the matrix exponential in (4.10), using the spectral decomposition, involves the computation of the eigenvalues and eigenvectors of Q, which is extremely prone to round off errors when Q is large.

As Grassmann (1977) points out, the existence of negative diagonal elements in Q leads to round off errors. The randomisation method (Gross *et al.*, 1984, Grassmann, 1977) is a technique which uses a discrete time Markov chain to derive the transient measures of a continuous time Markov process. Since the transition matrix of a Markov chain contains probabilities, the algorithms work with positive numbers only, which minimises the round off error.

The technique is based on the following property (Gross *et al.*, 1984). Let $\{X(t),t\geq 0\}$ be an MP on a finite state space S_X with generator Q. Define $\Lambda = \max_i Q_{ii}$, then there exists a discrete time Markov chain $\{Y_n,n\in\mathbb{N}\}$ on S_X with transition matrix $P=Q/\Lambda+I$ and a Poisson process $\{N(t),t\geq 0\}$ with rate Λ, such that the processes $\{Y_{N(t)},t\geq 0\}$ and $\{X(t),t\geq 0\}$ are probabilistically identical. Gross *et al.* show that

$$P(t) = \exp(-\Lambda t)\ p_0 \sum_{n=0}^{\infty} \frac{(\Lambda t)^n}{n!}\ P^n \tag{4.11}$$

and truncating the infinite series (4.11), a recursive procedure is used to compute P(t). Arsham *et al.* (1983) and Grassmann (1977) consider the choice of truncation points. In fact, computation of (4.11) is a straightforward matrix multiplication. However, efficient computation should exploit the sparseness of the generators Q in the previous sections. Gross *et al.* show how to compute transient measures and formulate an algorithm to obtain the time dependent state probabilities, exploiting the sparseness of the generator of the Markov process to the fullest degree. De Souza e Silva *et al.* (1986) apply the technique to a practical situation, to calculate a number of transient quantities of a repairable computer system.

Concluding this chapter, it is clear that:

i. approximating the life and repair time distributions by PH distributions is a powerful tool to analyse the failure bahaviour of a system, since PH distributions render the state description process Markovian,

ii. the randomisation technique is a powerful tool in computing the time dependent performance measures of an MP. The technique is useful when special interest is in the transient behaviour of a system or when the system reaches its steady state very slowly.

Nevertheless, in chapter 5 attention is concentrated on the situation with BVPH life times and *generally* distributed repair times. It appears that computation of the transient performance measures is relatively complicated in this case. However, the *stationary* measures can be obtained relatively easily and quickly, since the number of system states is reduced considerably under generally distributed repair times. Consequently, attention is focused primarily on the stationary behaviour of the system and the Laplace transform technique is used to develop algorithmic forms for computing the stationary operating characteristics.

5. BIVARIATE PHASE TYPE DISTRIBUTED LIFE TIMES

5.1 Introduction

In the previous chapters two situations have been examined. In chapter 3 the system's life time distribution was a BVE and the repair time distribution general and in chapter 4 the concept of phase type distributions was introduced, making the process under consideration Markovian. As mentioned, phase type distributions are dense in the class of distribution functions and hence any distribution function can be approximated by a phase type distribution. However, the number of states in the resulting Markov process can become very large and in this case standard Markov techniques fail in practical applications. Nevertheless, computation of performance measures is still possible when the randomisation technique (Grassmann, 1977, Gross *et al.*, 1984) is applied. With this technique both the transient and stationary behaviour of the process can be analysed in a relatively easy, but time consuming way. The technique seems especially useful for the transient behaviour.

Compared to chapter 3, a further generalisation is made here towards the system's life time distribution and, compared to chapter 4, towards the repair time distribution. The life time distribution is assumed to be a BVPH while the repair times are generally distributed. In this case the time dependent behaviour can be obtained by solving sets of coupled integral equations. Further, the stationary behaviour is governed by a matrix which is closely connected with the generator of the Markov processes in chapter 4. This matrix is the Laplace transform of the description of the system's transition mechanism and its dimension is small compared to that of the generator of the Markov processes under consideration. As a consequence, stationary measures such as the mean time to system failure, the limiting availability and the stationary state probabilities can be obtained easily and quickly. Thus, the randomisation technique seems preferable when the system's time dependent behaviour is of interest and the methods described in this chapter seem preferable with respect to the stationary behaviour.

The joint distribution function of the life times of the units is a BVPH with an exponential common cause. In principle the results can be extended without difficulties to the case where the common cause component C_3 has a phase type distribution, as shown in appendix A. However, an exponential distribution is chosen for C_3 because it simplifies the formulas and contributes to a clearer insight into the mechanisms and techniques used. The life time of component C_k has, as in chapter 4, a phase type distribution with representation (α_k,T_k) and dimension n_k, k=1,2, but now the repair time of C_k is general (with pdf $g_k(.)$). Thus, at any time C_k can be found in

i. phase i of its life time,

ii. in the repair state 'r', or

iii. in the waiting-for-repair state 'w'.

Since the state of C_3 (working or not working) follows from the state description of C_1 and C_2, the state of the system can be denoted by the pair (x_1,x_2), where x_k denotes the state of C_k and $x_k \in \{1,..,n_k,r,w\}$, k=1,2.
Subsequently, an instantaneous state 'o' is introduced. A component is said to be in state 'o' at time t, if it (re)starts operating at t, *i.e.* at time t a phase is selected and with probability $\alpha_k(i)$ component C_k starts operating in phase i of its life time ($i=1,..,n_k$, k=1,2). As will become clear later, the main reason for introducing this instantaneous state is to simplify the formulas for the system reliability, availability, etc.
The system state description is in terms of a vector valued process $\{X(t),t\geq0\}$, where $X(t)=(X_1(t),X_2(t))$ and $X_k(t)$ denotes the state of component k. Obviously, $X_k(t)\in\{o,1,..,n_k,r,w\}$ and in fact the state description process $\{Z(t),t\geq0\}$ in section 1.5 is a special case with only one operating state.
To derive expressions for the system's performance measures, such as (interval) reliability, (joint) availability, state probabilities, their stationary behaviour and the quasi-stationary distributions, the events $E_{(x_1,x_2)}$, $x_k\in\{o,1,..,n_k,r,w\}$, k=1,2, are defined as

$$E_{(x_1,x_2)} : \text{the system enters state } (x_1,x_2).$$

As expected, the regenerative nature of events $E_{(x_1,x_2)}$ will be used to determine the above performance measures. Since the sojourn time in a particular phase of a phase type distribution is exponentially distributed, the lack-of-memory property applies locally, *i.e.* per phase. As a consequence, the state description process $\{X(t),t\geq0\}$ is regenerative with respect to the events $E_{(x_1,x_2)}$ for $x_k\in\{o,1,..,n_k\}$, k=1,2. However, the process $\{X(t),t\geq0\}$ is non-regenerative with respect to the events $E_{(i,r)}$, $E_{(r,j)}$, $E_{(r,w)}$ and $E_{(w,r)}$, since an entry into the corresponding states can be generated by a transition of an operating component *during* the repair of the other component. The nature of the events $E_{(x_1,x_2)}$ is shown in table 5.1.

The notation is:

i. R : the process X(t) is regenerative with respect to the event $E_{(x_1,x_2)}$

ii. N : the process X(t) is non-regenerative with respect to event $E_{(x_1,x_2)}$

iii. – : an entrance into state (x_1,x_2) is not possible.

x_1 \ x_2	o	phase j	r	w
o	–	R	R	–
phase i	R	R	N	–
r	R	N	–	N
w	–	–	N	–

Table 5.1: Nature of the events $E_{(x_1,x_2)}$

As mentioned, the process $\{X(t),t\geq 0\}$ is non-regenerative with respect to the events $E_{(i,r)}$, $E_{(r,j)}$, $E_{(r,w)}$ and $E_{(w,r)}$. However, in fact there are two possibilities to enter the states (i,r), (r,j) (r,w) and (w,r). According to these possibilities, define

$\tilde{E}_{(i,r)}$: The system enters state (i,r) by a transition of component C_1 to phase i of its life time, during the repair of C_2,

$\tilde{E}_{(r,w)}$: The system enters state (r,w) by a failure of C_2 during the repair of C_1.

and redefine $E_{(i,r)}$ and $E_{(r,w)}$ as:

$E_{(i,r)}$: The system enters state (i,r), caused by a failure of C_2, while C_1 is in phase i of its life time,

$E_{(r,w)}$: The system enters state (r,w) by a common cause failure.

The events $\tilde{E}_{(r,j)}$, $\tilde{E}_{(w,r)}$, $E_{(r,j)}$ and $E_{(w,r)}$ being defined similarly, it is clear that the process $\{X(t),t\geq 0\}$ is non-regenerative with respect to $\tilde{E}_{(i,r)}$, $\tilde{E}_{(r,j)}$, $\tilde{E}_{(r,w)}$ and $\tilde{E}_{(w,r)}$, but regenerative with respect to $E_{(i,r)}$, $E_{(r,j)}$, $E_{(r,w)}$ and $E_{(w,r)}$, since the events $E_{(i,r)}$, $E_{(r,j)}$, $E_{(r,w)}$ and $E_{(w,r)}$ are initiated by the *start* of a repair.

In the remainder of this chapter, the latter, more detailed definition of the (*regenerative*) events $E_{(i,r)}$, $E_{(r,j)}$, $E_{(r,w)}$ and $E_{(w,r)}$ is used.

The section concludes with some specific notation:

e_k	unit column vector of dimension n_k.
$\alpha_{k,i}$	column vector of dimension n_k, filled with zeros, except for position i, which has a one.
$\alpha_{k,o}$	the initial probability vector of component C_k: $\alpha_{k,o}=\alpha_k$.
I_k	identity matrix of dimension k.
T_k^0	satisfies $T_k e_k + T_k^0 = 0$.
$T_{.n}$	n thcolumn of a matrix T.
$F_{k,i}(t)$	$= 1 - \alpha_{k,i}^T \exp(T_k t)\ e_k$, $k=1,2$, $i \in \{o,1,..,n_k\}$.
$f_{k,i}(t)$	$= \alpha_{k,i}^T \exp(T_k t)\ T^0$, $k=1,2$, $i \in \{o,1,..,n_k\}$.
rhs	right hand side
mrp	maximum real part

Note that $F_{k,i}(.)$ represents the life time distribution of component C_k, given a start in $i \in \{o,1,..,n_k\}$ at t=0. Similarly, $f_{k,i}(.)$ represents the pdf of the life time of C_k, given a start in state i at t=0.

5.2 System reliability

The reliability of the two-unit dependent parallel system, given an $E_{(i,j)}$ event at t=0, is denoted by $R_{ij}^{dep}(t)$:

$$R_{ij}^{dep}(t) = \Pr\{X(u) \neq (r,w) \wedge X(u) \neq (w,r),\ 0 \leq u \leq t \mid E_{(i,j)} \text{ at } t=0\},$$

where $(i,j) \in \{o,1,..,n_1,r\} x \{o,1,..,n_2,r\} \backslash \{(r,r)\}$.

Hence, the reliability of a new system is $R_{oo}^{dep}(.)$ and when the system restarts operating after it has been down, the reliability is given by $R_{or}^{dep}(.)$ or $R_{ro}^{dep}(.)$, depending on which unit has been repaired first.

As described in section 1.4, the two-unit dependent parallel system is modelled as a series system consisting of a common cause component and a two-component independent parallel system (figure 1.1).

Hence the reliability can be written as

$$R_{ij}^{dep}(t) = R_{ij}^{ind}(t)\ R_{cc}(t), \tag{5.1}$$

where $R_{ij}^{ind}(.)$ is the reliability of the two–component independent parallel system, consisting of C_1 and C_2 (figure 1.1), and $R_{cc}(.)$ is the reliability of the common cause component C_3. Thus, in fact the problem of finding $R_{ij}^{dep}(.)$ reduces to the derivation of an expression for $R_{ij}^{ind}(.)$. It is useful to note that, whenever dependence is modelled by connecting an artificial component in series with a system, the decomposition (5.1) can be used to compute the system's reliability.

In order to find the reliability $R_{ij}^{ind}(.)$, the discrete supplementary variable technique is applied, where, obviously, the supplementary variables contain the phases of the life time distributions. Ravichandran (1981) followed a similar approach to obtain the MTSF and limiting availability of a two–unit *independent* parallel system.

For simplicity, the reliability $R_{ij}^{ind}(.)$ is abbreviated to $R_{ij}(.)$. Considering the mutually exclusive and exhaustive events that C_1 fails in (0,t) or does not fail in (0,t), gives, for $(i,j)\in\{o,1,..,n_1\}x\{o,1,..,n_2\}$,

$$R_{ij}(t) = \bar{F}_{1,i}(t) + \sum_{k=1}^{n_2} \int_0^t f_{1,i}(u)\ P_{2,jk}(u)\ R_{rk}(t-u)\ du, \tag{5.2}$$

where

$$P_{k,ij}(t) = \Pr\{X_k(t)=j \mid X_k(0)=i,\ X_{3-k}(u)\neq r,\ X_{3-k}(u)\neq w,\ 0\leq u\leq t\},\ k=1,2. \tag{5.3}$$

Expression (5.3) represents the probability that component C_k is found in phase j at time t, given that C_k starts in phase i at t=0 and component C_{3-k} does not fail until time t.

Formally, for k=1,2, $i,j=1,..,n_k$

$$P_{k,ij}(t) = \exp(T_k t)_{ij} + f_{k,i}(t) * g_k(t) * P_{k,oj}(t), \tag{5.4}$$

and

$$P_{k,oj}(t) = \left[\sum_{m=0}^{\infty} \left[f_{k,0}(t)*g_k(t)\right]^{(m)}\right] * \alpha_k^T \exp(T_k t)_{.j}. \tag{5.5}$$

$P_{k,oj}(t)$ represents the probability that C_k fails a finite number of times, is repaired every time and finally reaches state j at time t. The interpretation of (5.4) is analogous: the first term gives the probability that C_k does not fail until time t and the second term the probability that C_k fails one or more times.

The roles of the components C_1 and C_2 may also be interchanged to derive a dual expression for $R_{ij}(t)$:

$$R_{ij}(t) = \bar{F}_{2,j}(t) + \sum_{k=1}^{n_1} \int_0^t f_{2,j}(u)\ P_{1,ik}(u)\ R_{kr}(t-u)\ du. \tag{5.6}$$

Further, for $i \in \{o,1,..,n_1\}$ and $j \in \{o,1,..,n_2\}$,

$$R_{ir}(t) = \bar{F}_{1,i}(t)\ \bar{G}_2(t) + \sum_{k=1}^{n_1} \int_0^t g_2(u)\ \alpha_{1,i}^T\ \exp(T_1u)_{.k}\ R_{ko}(t-u)\ du, \tag{5.7}$$

and

$$R_{rj}(t) = \bar{F}_{2,j}(t)\ \bar{G}_1(t) + \sum_{k=1}^{n_2} \int_0^t g_1(u)\ \alpha_{2,j}^T\ \exp(T_2u)_{.k}\ R_{ok}(t-u)\ du. \tag{5.8}$$

The expressions (5.7) and (5.8) are obtained by considering the events that a repair is completed in (0,t) or not. Alternative expressions for $R_{rj}(t)$ and $R_{ir}(t)$ can be found by considering the events that the operating component fails or does not fail in (0,t). The equations (5.2) and (5.6)–(5.8) describe the reliability of the system with independent units and hence from (5.1) the reliability of the system with dependent units is obtained.

In principle, the reliability at time t, given an $E_{(i,j)}$ event at the time origin, can be computed solving the above system of integral equations. However, here attention is concentrated on the MTSF conditioned on an $E_{(i,j)}$ event at t=0 rather than on the time dependent behaviour. It is shown that this conditional MTSF can be computed relatively easily.

In order to find the MTSF of the two-unit dependent parallel system, note that $(R_{ij}^{ind})^*(0)$ is the mean time to system failure of the parallel system with two independent components, given $E_{(i,j)}$ at t=0. Now, for the moment suppose that C_3 has an exponential life time distribution with parameter λ, *i.e.*

$$R_{cc}(t) = \exp(-\lambda t).$$

Then, from (5.1)

$$(R_{ij}^{dep})^*(s) = (R_{ij}^{ind})^*(\lambda+s), \tag{5.9}$$

and obviously, the MTSF of the parallel system with dependent components is

$$(R_{ij}^{dep})^*(0) = (R_{ij}^{ind})^*(\lambda).$$

To obtain $(R_{ij}^{ind})^*(\lambda)$, the Laplace transforms of (5.2) and (5.6)–(5.8) are computed:

$$R_{ij}^*(s) = F_{1,i}^*(s) + \sum_{k=1}^{n_2} \mathcal{L}\left(f_{1,i}(t)\ P_{2,jk}(t)\right)\ R_{rk}^*(s), \tag{5.10}$$

$$R_{ij}^*(s) = F_{2,j}^*(s) + \sum_{k=1}^{n_1} \mathcal{L}\left(f_{2,j}(t)\ P_{1,ik}(t)\right)\ R_{kr}^*(s), \tag{5.11}$$

$$R_{rj}^*(s) = \mathcal{L}\left(\bar{F}_{2,j}(t)\ \bar{G}_1(t)\right) + \sum_{k=1}^{n_2} \mathcal{L}\left(g_1(t)\ \alpha_{2,j}^T\ \exp(T_2t)_{.k}\right)\ R_{ok}^*(s), \tag{5.12}$$

and

$$R_{ir}^*(s) = \mathcal{L}\left(\bar{F}_{1,i}(t)\ \bar{G}_2(t)\right) + \sum_{k=1}^{n_1} \mathcal{L}\left(g_2(t)\ \alpha_{1,i}^T\ \exp(T_1t)_{.k}\right)\ R_{ko}^*(s). \tag{5.13}$$

It is clear that the expressions (5.10)–(5.13) form a set of linear equations in the Laplace transforms of the reliability functions $R_{ij}(t)$. The coefficients in the equations can be determined easily if the matrices T_1 and T_2 are diagonalisable. In this case, let $S_k^{-1}D_kS_k$ be the diagonalisation of T_k, *i.e.* D_k is the diagonal matrix with the eigenvalues of T_k, S_k^{-1} the matrix with right eigenvectors and S_k the matrix with left eigenvectors. Further, define a function $f^*(.)$ on an n–dimensional diagonal matrix $D=\mathrm{diag}(d_1,..,d_n)$ as $f^*(D)=\mathrm{diag}(f^*(d_1),..,f^*(d_n))$, then explicit expressions for the coefficients in (5.10)–(5.13) are

$$F_{k,i}^*(s) = \alpha_{k,i}^T\ (sI_k-T_k)^{-1}e_k,$$

$$f_{k,i}^*(s) = \alpha_{k,i}^T\ (sI_k-T_k)^{-1}T_k^0,$$

$$\mathfrak{L}\big(f_{1,k}(t)\ P_{2,ij}(t)\big) = \alpha_{1,k}^{T}\ S_1^{-1}\ P_{2,ij}^{*}(sI_1-D_1)\ S_1\ T_1^{0},$$

where

$$P_{k,ij}^{*}(s) = ((sI_k-T_k)^{-1})_{ij} + f_{k,i}^{*}(s)\ g_k^{*}(s)\ P_{k,oj}^{*}(s),$$

$$P_{k,oj}^{*}(s) = \frac{\alpha_k^{T}((sI_k-T_k)^{-1})_{.j}}{1-f_{k,o}^{*}(s)\ g_k^{*}(s)},$$

$$\mathfrak{L}\big(\bar{F}_{k,i}(t)\ \bar{G}_j(t)\big) = \alpha_{k,i}^{T}\ S_k^{-1}\ \bar{G}_j^{*}(sI_k-D_k)\ S_k\ e_k,$$

and finally

$$\mathfrak{L}\big(\exp(T_k t)_{ij}\ g_m(t)\big) = \Big[S_k^{-1}\ g_m^{*}(sI_k-D_k)\ S_k\Big]_{ij}.$$

To solve the set of equations (5.10)–(5.13), write the equations in vector–matrix notation as

$$r^{*}(s) = \mathfrak{M}^{*}(s)\ r^{*}(s) + b_r^{*}(s), \tag{5.14}$$

where the column vector $r^{*}(s)$ contains quantities $R_{ij}^{*}(s)$, the matrix $\mathfrak{M}^{*}(s)$ the coefficients and the column vector $b_r^{*}(s)$ the first term on the right hand side of the equations (5.10)–(5.13). From a computational point of view, it is of course important to construct a matrix $\mathfrak{M}^{*}(s)$ with dimensions as small as possible. Suppose that $n_2 \le n_1$, then observation of (5.10) and (5.12) shows that an appropriate choice for $r^{*}(s)$ is

$$r^{*}(s) = [R_{o1}^{*}(s),..,R_{on_2}^{*}(s),R_{r1}^{*}(s),..,R_{rn_2}^{*}(s)]^{T}.$$

As a result

$$b_r^{*}(s) = [\mathfrak{L}(\bar{F}_{1,0}(t)),..,\mathfrak{L}(\bar{F}_{1,0}(t)),\mathfrak{L}(\bar{F}_{2,1}(t)\bar{G}_1(t)),..,\mathfrak{L}(\bar{F}_{2,n_2}(t)\bar{G}_1(t))]^{T}$$

and

$$\mathfrak{M}^{*}(s) = \left[\begin{array}{c|c} O & \mathfrak{M}_{12}^{*}(s) \\ \hline \mathfrak{M}_{21}^{*}(s) & O \end{array}\right],$$

where

$$(\mathfrak{M}_{12}^{*}(s))_{ij} = \mathfrak{L}(f_{1,0}(t)\ P_{2,ij}(t)),\ \ i,j\in\{1,..,n_2\}$$

and

$$(\mathfrak{M}^*_{21}(s))_{ij} = \mathfrak{L}(g_1(t)\ \alpha^T_{2,i}\ (\exp(T_2 t))_{.j}),\ i,j\in\{1,..,n_2\}.$$

From (5.14)

$$r^*(s) = (I_{2n_2} - \mathfrak{M}^*(s))^{-1}\ b^*_r(s). \tag{5.15}$$

Since

$$\forall_{i\in\{1,..,n_2\}} \sum_{j=1}^{n_2} (\mathfrak{M}^*_{12}(s))_{ij} \le f^*_{1,o}(s)$$

and

$$\forall_{i\in\{1,..,n_2\}} \sum_{j=1}^{n_2} (\mathfrak{M}^*_{21}(s))_{ij} \le g^*_1(s),$$

it follows that $\mathfrak{M}^*(s)$ is substochastic for s>0, *i.e.* the matrix $(I_{2n_2} - \mathfrak{M}^*(s))^{-1}$ exists for s>0. Hence, setting $s=\lambda$ in (5.15) gives $R^*_{oj}(\lambda)$ and $R^*_{rj}(\lambda)$, $j=1,..,n_2$. Substitution of $R^*_{rj}(\lambda)$ in (5.10) yields $R^*_{ij}(\lambda)$, $(i,j)\in\{o,1,..,n_1\}$x $\{o,1,..,n_2\}$ and substitution of $R^*_{ko}(\lambda)$ in (5.13) gives $R^*_{ir}(\lambda)$, $i\in\{o,1,..,n_1\}$. Quantities of special interest are $R^*_{oo}(\lambda)$, $R^*_{ro}(\lambda)$ and $R^*_{or}(\lambda)$: $R^*_{oo}(\lambda)$ is the mean time to first system failure of a new system. On the other hand $R^*_{or}(\lambda)$ represents the mean time to system failure given $E_{(o,r)}$ at t=0, *i.e.* at t=0 a transition from state (r,w) to (o,r) occurs.

In the more general situation where the common cause component has a phase type distribution with representation (α_3,T_3) and dimension n_3, the state description of the components is in terms of triples (x_1,x_2,x_3). The events $E_{(x_1,x_2,x_3)}$ are defined analogously to the events $E_{(x_1,x_2)}$. The reliability of the dependent system, given $E_{(i,j,k)}$ at t=0, is

$$R^{dep}_{ijk}(t) = R^{ind}_{ij}(t)\ R_{k,cc}(t),$$

where $R_{k,cc}(t)$, the reliability of C_3 given a start in state k, is

$$R_{k,cc}(t) = \alpha^T_{3,k}\ \exp(T_3 t)\ e_3.$$

The mean time to system failure, given $E_{(i,j,k)}$ at t=0, is found by setting s=0 in $(R^{dep}_{ijk})^*(s)$, where

$$(R^{dep}_{ijk})^*(s) = \alpha^T_{3,k}\ S_3^{-1} (R^{ind}_{ij})^* (sI_3 - D_3)\ S_3 e_3. \tag{5.16}$$

Note that (5.9) is the special case of (5.16) with $I_3=S_3=1$, $D_3=-\lambda$ and $\alpha_{3,k}=e_3=1$.

5.3 The diagonalisation problem

In the previous section the diagonalisation of the matrices T_k was used to obtain explicit expressions for various Laplace transforms. The assumption that T_k is diagonalisable is not an essential restriction but, from a practical point of view, the complexity of the computation of the matrix $\mathfrak{M}^*(s)$ is reduced considerably by this assumption.

To illustrate the computational consequences of a non-diagonalisable matrix T, $\mathfrak{L}(f(t)H(t))$ is constructed, where H(t) is an arbitrary function with Laplace transform $H^*(s)$ and f(t) is the pdf of the phase type distribution with representation (α, T).

Suppose $\lambda_1,..,\lambda_M$ are the eigenvalues of the nxn non-diagonalisable matrix T and let m_i be the multiplicity of λ_i (thus $\sum m_i = n$). Then (Finkbeiner, 1978, p.198) it follows that at least one eigenvalue λ_i has multiplicity $m_i>1$.

Alternatively, the pdf f(t) has a rational Laplace transform (Neuts, 1981) and hence

$$f^*(s) = \frac{N(s)}{D(s)},$$

where D(s) and N(s) are polynomials in s of degree n and (at most) n–1 respectively. As the eigenvalues λ_i of the matrix T are the roots of D(s), it follows $f^*(s)$ can be written in partial fractions as

$$f^*(s) = \sum_{i=1}^{M} \sum_{j=1}^{m_i} c_{ij} \left[\frac{\lambda_i}{\lambda_i+s}\right]^j.$$

Taking the inverse Laplace transform yields

$$f(t) = \sum_{i=1}^{M} \sum_{j=1}^{m_i} c_{ij} \frac{\lambda_i \, (\lambda_i t)^{j-1}}{(j-1)!} \exp(-\lambda_i t) \tag{5.17}$$

and hence

$$F(t) = \sum_{i=1}^{M} \sum_{j=1}^{m_i} \sum_{k=0}^{j-1} c_{ij} \frac{(\lambda_i t)^k}{k!} \exp(-\lambda_i t). \tag{5.18}$$

Note that from (5.17) and (5.18) it follows that every phase type distribution can be written as a finite mixture of Erlang distributions.

Further, since

$$\mathfrak{L}(t^j\, H(t)) = (-1)^j \frac{d^j}{ds^j} H^*(s)$$

it is easily seen

$$\mathfrak{L}(f(t)H(t)) = \sum_{i=1}^{M} \sum_{j=0}^{m_i-1} c_{ij}\, \lambda_i \frac{(-\lambda_i)^j}{j!} \left[\frac{d^j}{ds^j} H^*(s) \bigg|_{\lambda_i+s} \right]$$

and thus, in order to compute $\mathfrak{L}(f(t)H(t))$, the first m derivatives of $H^*(s)$ are needed, where $m = \max_i (m_i-1)$.

Hence, computation of *e.g.* $\mathfrak{L}(\mathrm{F}_{k,i}(t)\bar{G}_{3-k}(t))$ in (5.12) and (5.13) is still relatively easy for non-diagonalisable T_1 or T_2, but computing the Laplace transforms of the required multiple integrals in the next section can be a very tedious job. Therefore, the conclusion is that, although the techniques described in this chapter also apply to the case where the matrices T_k are non-diagonalisable, implementation will be tedious.

5.4 System availability

In the following analysis the common cause component is supposed to have an exponentially distributed life time, which makes the formulas more easily readable. Secondly, no additional insight is gained in the situation where the common cause component has a phase type life time distribution. However, the *main* reason is that in practical applications the common cause effect is often taken exponentially, caused by a lack of data on common cause failures. For the interested reader, the modifications which have to be made under a phase type distributed common cause effect are outlined in appendix A.

For the system with dependent units the availability function $A_{ij}^{dep}(t)$, $(i,j)\in \{o,1,..,n_1,r,w\}x\{o,1,..,n_2,r,w\}\backslash\{(r,r),(w,w)\}$, is defined as

$$A_{ij}^{dep}(t) = \Pr\{X(t)\neq(w,r) \wedge X(t)\neq(r,w) \mid E_{(i,j)} \text{ at } t=0\}.$$

Since the availability $A_{ij}^{ind}(t)$ of the independent parallel system consisting of the components C_1 and C_2 is influenced by the availability of the common cause component C_3 and *vice versa*,

$$A_{ij}^{dep}(t) \neq A_{ij}^{ind}(t)\; A_{cc}(t).$$

However, observation of the process X(t) yields the following expressions for $A_{ij}^{dep}(t)$. For simplicity $A_{ij}^{dep}(t)$ is abbreviated to $A_{ij}(t)$. Further, let p denote the probability that C_1 is the first component to be repaired when a common cause failure occurs.

Then, for $(i,j)\in\{o,1,..,n_1\}x\{o,1,..,n_2\}$,

$$A_{ij}(t) = \bar{F}_{1,i}(t)\ \exp(-\lambda t) \tag{5.19}$$

$$+\sum_{k=1}^{n_2} \int_0^t f_{1,i}(u)\ P_{2,jk}(u)\ \exp(-\lambda u)\ A_{rk}(t-u)\ du$$

$$+\int_0^t \int_0^u \int_0^v f_{1,i}(v)\ P_{2,jr}(v,w)\ \exp(-\lambda v)\ \frac{g_2(u-v+w)}{\bar{G}_2(w)}\ A_{ro}(t-u)\ dw\ dv\ du$$

$$+\sum_{k=1}^{n_2} \int_0^t \bar{F}_{1,i}(u)\ P_{2,jk}(u)\ \lambda\exp(-\lambda u)\ \left[p\ A_{rw}(t-u) + (1-p)\ A_{wr}(t-u)\right]\ du$$

$$+\int_0^t \int_0^u \int_0^v \bar{F}_{1,i}(v)\ P_{2,jr}(v,w)\ \lambda\exp(-\lambda v)\ \frac{g_2(u-v+w)}{\bar{G}_2(w)}\ A_{ro}(t-u)\ dw\ dv\ du,$$

where

$$A_{wr}(t) = \int_0^t g_2(u)\ A_{ro}(t-u)\ du,$$

and

$$A_{rw}(t) = \int_0^t g_1(u)\ A_{or}(t-u)\ du.$$

The interpretation of the function $P_{2,jr}(v,w)$ is that $P_{2,jr}(v,w)\ dw$ represents the probability that C_2 fails in (v−w,v−w+dw) and is still in repair at time v, given that C_2 starts in phase j at time 0 and C_1 and C_3 do not fail until time v. Formally,

$$P_{2,jr}(v,w)\ dw = \left[f_{2,j}(v-w) * \left[\sum_{m=0}^{\infty} \left[f_{2,0}(v-w) * g_2(v-w) \right]^{(m)} \right] \right] \bar{G}_2(w)\ dw.$$

Expression (5.19) is obtained by considering the following mutually exclusive and exhaustive cases:

i. The first term represents the probability that both C_1 and C_3 survive t.

Subsequently, if C_1 or C_3 does not survive to time t, then C_1 fails before C_3 or C_3 fails before C_1:

ii. The second term represents the probability that C_1 fails before a failure of C_3 has occurred and at the moment of failure C_2 is in state k,

iii. The third term gives the probability that C_1 fails before a failure of C_3 has occurred, at the moment of failure C_2 has been in repair for w units of time and the repair of C_2 is completed at time u.

Analogously to the second and third term, the fourth and fifth term give the probability that C_3 fails before a failure of C_1 has occurred.

The interpretation of $P_{2,jr}(v,w)\,dw$ is quite easy as the product of two probabilities, *viz.* the probability that a failure of C_2 (not necessarily the first) occurs in (v–w,v–w+dw) and the probability that the repair time of C_2 exceeds w units of time. Of course, a dual expression for $A_{ij}(t)$ is found by interchanging the roles of C_1 and C_2.
An expression for the availability function, conditioned on an $E_{(i,r)}$ event at the time origin is, for $i\in\{o,1,..,n_1\}$

$$A_{ir}(t) = \bar{F}_{1,i}(t)\exp(-\lambda t)\,\bar{G}_2(t) \tag{5.20}$$

$$+\sum_{k=1}^{n_1}\int_0^t g_2(u)\,\alpha_{1,i}^T\,(\exp(T_1u))_{.k}\,\exp(-\lambda u)\,A_{ko}(t-u)\,du$$

$$+\int_0^t\int_0^u f_{1,i}(v)\,\exp(-\lambda v)\,dv\;g_2(u)\,A_{ro}(t-u)\,du$$

$$+\int_0^t\int_0^u \lambda\exp(-\lambda v)\,\bar{F}_{1,i}(v)\,dv\;g_2(u)\,A_{ro}(t-u)\,du.$$

Equation (5.20) is derived considering the events that the repair of C_2 is finished before time t or not. If the repair is completed before time t, the first event which happens is

i. the repair completion of C_2 (second term),

ii. a failure of C_1 (third term) or

iii. a failure of C_3 (fourth term).

Obviously, by interchanging the roles of C_1 and C_2 in (5.20), for $j \in \{o,1,..,n_2\}$

$$A_{rj}(t) = \bar{F}_{2,j}(t)\ \exp(-\lambda t)\ \bar{G}_1(t) \tag{5.21}$$

$$+\sum_{k=1}^{n_2} \int_0^t g_1(u)\ \alpha_{2,j}^T\ (\exp(T_2 u))_{.k}\ \exp(-\lambda u)\ A_{ok}(t-u)\ du$$

$$+\int_0^t \int_0^u f_{2,j}(v)\ \exp(-\lambda v)\ dv\ g_1(u)\ A_{or}(t-u)\ du$$

$$+\int_0^t \int_0^u \lambda\exp(-\lambda v)\ \bar{F}_{2,j}\ (v)\ dv\ g_1(u)\ A_{or}(t-u)\ du.$$

To obtain the limiting availability, the Laplace transforms of (5.19)–(5.21) are computed. The result is a set of linear equations in the Laplace transforms of the availability functions $A_{ij}(t)$:

$$A_{ij}^*(s) = \bar{F}_{1,i}^*(\lambda+s) \tag{5.22}$$

$$+\sum_{k=1}^{n_2} \mathfrak{L}\Big[f_{1,i}(t)\ P_{2,jk}(t)\ \exp(-\lambda t)\Big]\ A_{rk}^*(s)$$

$$+\mathfrak{L}\Big[\int_0^t \int_0^v f_{1,i}(v)\ P_{2,jr}(v,w)\ \exp(-\lambda v)\ \frac{g_2(t-v+w)}{\bar{G}_2(w)}\ dw\ dv\Big]\ A_{ro}^*(s)$$

$$+\sum_{k=1}^{n_2} \mathfrak{L}\Big[\bar{F}_{1,i}(t)\ P_{2,jk}(t)\ \lambda\exp(-\lambda t)\Big]\ \big[pA_{rw}^*(s) + (1-p)A_{wr}^*(s)\big]$$

$$+\mathfrak{L}\Big[\int_0^t \int_0^v \bar{F}_{1,i}(v)\ P_{2,jr}(v,w)\ \lambda\exp(-\lambda v)\ \frac{g_2(t-v+w)}{\bar{G}_2(w)}\ dw\ dv\Big]\ A_{ro}^*(s),$$

where

$$A^*_{wr}(s) = g^*_2(s)\ A^*_{ro}(s)$$

and

$$A^*_{rw}(s) = g^*_1(s)\ A^*_{or}(s).$$

Subsequently,

$$\begin{aligned} A^*_{ir}(s) = {} & \mathfrak{L}\big(\bar{F}_{1,i}(t)\ \exp(-\lambda t)\ \bar{G}_2(t)\big) \\ & + \sum_{k=1}^{n_1} \mathfrak{L}\Big[g_2(t)\ \alpha^T_{1,i}\ (\exp(T_1 t))_{.k}\ \exp(-\lambda t)\Big]\ A^*_{ko}(s) \\ & + \mathfrak{L}\Big[g_2(t) \int_0^t f_{1,i}(u)\ \exp(-\lambda u)\ du\Big]\ A^*_{ro}(s) \\ & + \mathfrak{L}\Big[g_2(t) \int_0^t \lambda \exp(-\lambda u)\ \bar{F}_{1,i}(u)\ du\Big]\ A^*_{ro}(s) \end{aligned} \tag{5.23}$$

and

$$\begin{aligned} A^*_{rj}(s) = {} & \mathfrak{L}(\bar{F}_{2,j}(t)\ \exp(-\lambda t)\ \bar{G}_1(t)) \\ & + \sum_{k=1}^{n_2} \mathfrak{L}\Big[g_1(t)\ \alpha^T_{2,j}\ (\exp(T_2 t))_{.k}\ \exp(-\lambda t)\Big]\ A^*_{ok}(s) \\ & + \mathfrak{L}\Big[g_1(t) \int_0^t f_{2,j}(u)\ \exp(-\lambda u)\ du\Big]\ A^*_{or}(s) \\ & + \mathfrak{L}\Big[g_1(t) \int_0^t \lambda \exp(-\lambda u)\ \bar{F}_{2,j}(u)\ du\Big]\ A^*_{or}(s). \end{aligned} \tag{5.24}$$

Again, as in section 5.2, the coefficients in (5.22)–(5.24) can be determined explicitly by means of the diagonalisation $S_k^{-1} D_k S_k$ of T_k:

$$\mathfrak{L}\Big[g_{3-k}(t) \int_0^t f_{k,i}(u)\ \exp(-\lambda u)\ du\Big]$$

$$= \alpha^T_{k,i}\ (\lambda I_k - T_k)^{-1}\ S_k^{-1}\ \Big[g^*_{3-k}(sI_k) - g^*_{3-k}((\lambda+s)I_k - D_k)\Big]\ S_k\ T^0_k,$$

$$\mathfrak{L}\left[g_{3-k}(t)\int_0^t \lambda\exp(-\lambda u)\ F_{k,j}(u)\ du\right]$$

$$= \lambda\ \alpha_{k,i}^T\ (\lambda I_k - T_k)^{-1}\ S_k^{-1}\ \left[g^*_{3-k}(sI_k) - g^*_{3-k}((\lambda+s)I_k - D_k)\right]\ S_k\ e_k$$

and

$$\mathfrak{L}\left(F_{1,i}(t)\ P_{2,jk}(t)\ \lambda\exp(-\lambda t)\right) = \lambda\ \alpha_{1,i}^T\ S_1^{-1}\ P^*_{2,jk}((\lambda+s)I_1 - D_1)\ S_1\ e_1.$$

In order to find the Laplace transforms of the multiple integrals in (5.22), define

$$\mathfrak{F}(t) := \int_0^t \int_0^v F_{1,i}(v)\ P_{2,jr}(v,w)\ \lambda\exp(-\lambda v)\ \frac{g_2(t-v+w)}{\bar{G}_2(w)}\ dw\ dv$$

and, for fixed v,

$$\mathfrak{P}_{2,jr}(v-w) := \frac{P_{2,jr}(v,w)}{\bar{G}_2(w)}.$$

Then

$$\mathfrak{F}(t) = \int_0^t F_{1,i}(v)\ \lambda\exp(-\lambda v) \int_0^v \mathfrak{P}_{2,jr}(v-w)\ g_2(t-(v-w))\ dw\ dv$$

and the inner integrand is a function of the difference v–w only. Substitution of u=v–w and changing the order of integration results in

$$\mathfrak{F}(t) = \int_0^t \mathfrak{P}_{2,jr}(u)\ g_2(t-u) \int_u^t F_{1,i}(v)\ \lambda\exp(-\lambda v)\ dv\ du.$$

Using the diagonalisation of T_1 to obtain a closed expression for the inner integral, finally gives

$$\mathfrak{F}^*(s) = \lambda\ \alpha_{1,i}^T\ (\lambda I_1 - T_1)^{-1} S_1^{-1}\ \mathfrak{P}^*_{2,jr}((\lambda+s)I_1 - D_1)\ \left[g_2^*(sI_1) - g_2^*((\lambda+s)I_1 - D_1)\right]\ S_1\ e_1,$$

where

$$\mathfrak{P}^*_{2,jr}(s) = \frac{f^*_{2,j}(s)}{1 - f^*_{2,o}(s)\ g^*_2(s)}.$$

Similarly,

$$\mathfrak{L}\left[\int_0^t \int_0^v f_{1,i}(v)\; P_{2,jr}(v,w)\; \exp(-\lambda v)\; \frac{g_2(t-v+w)}{\bar{G}_2(w)}\; dw\; dv\right]$$

$$= \alpha_{1,i}^T\; (\lambda I_1-T_1)^{-1}\; S_1^{-1}\; \mathfrak{P}_{2,jr}^*((\lambda+s)I_1-D_1)\; \left[g_2^*(sI_1)-g_2^*((\lambda+s)I_1-D_1)\right]\; S_1\; T_1^0.$$

Let A_{ij} denote the limiting availability given $E_{(i,j)}$ at t=0,

$$A_{ij} = \lim_{t\to\infty} A_{ij}(t)$$

then, applying the Abelian theorem for Laplace transforms (Cohen, 1982, p.651)

$$A_{ij} = \lim_{s\downarrow 0}\; s\; A_{ij}^*(s) \tag{5.25}$$

and the rhs of (5.25) is obtained from (5.22)–(5.24).
However, from the regenerative nature of the events $E_{(i,j)}$, the limiting availability is independent of the initial condition, thus for the meaningful pairs $(i,j)\in\{o,1,..,n_1,r,w\}\times\{o,1,..,n_2,r,w\}$,

$$A_{ij} = A. \tag{5.26}$$

In order to find the limiting availability A, write the set of equations (5.22)–(5.24) in the form

$$a^*(s) = M^*(s)\; a^*(s) + b_a^*(s), \tag{5.27}$$

where the column vector $a^*(s)$ contains quantities $A_{ij}^*(s)$ and the vector $b_a^*(s)$ and the matrix $M^*(s)$ contain the first term, respectively the coefficients on the rhs of the equations (5.22)–(5.24). For practical applications it is important to construct a matrix $M^*(s)$ with a dimension as small as possible. In the case $n_2\le n_1$, an appropriate choice for $a^*(s)$ is

$$a^*(s)=\left[A_{1o}^*(s),..,A_{n_1}^*(s),A_{ro}^*(s),A_{o1}^*(s),..,A_{on_2}^*(s),A_{or}^*(s),A_{r1}^*(s),..,A_{rn_2}^*(s)\right]^T.$$

To obtain the limiting availability A from (5.27), note that

$$A = \lim_{s\downarrow 0}\; s\; a^*(s) = \lim_{s\downarrow 0}\; s\; (I-M^*(s))^{-1} b_a^*(s). \tag{5.28}$$

However, it follows from probabilistic arguments that

$$\sum_{j=1}^{n} M^*_{ij}(0) = 1, \; i=1,..,n \tag{5.29}$$

where n denotes the dimension of $M^*(s)$. Thus $M^*(0)$ is a stochastic matrix (since $\forall_{i,j}\ \forall_{s\geq 0}\ M^*_{ij}(s)\geq 0$) and consequently the inverse in (5.28) does not exist for s=0.

To illustrate (5.29), consider (5.19) and remark that the five terms in (5.19) correspond to mutually exclusive and exhaustive events. In other words, the probabilities corresponding to the events sum to one. However, for $t\to\infty$ the first tends to zero, *i.e.* the probability that the first event occurs vanishes for $t\to\infty$ and hence the probabilities of the remaining four events sum to one. Now (5.29) follows from (5.22) and (5.25).

Although the inverse in (5.28) does not exist for s=0, the limit can be computed, applying the Perron–Frobenius theorem (Seneta, 1973, p.20) to $M^*(s)$.

To illustrate this, let $\mu_1(s),..,\mu_n(s)$ be the eigenvalues of $M^*(s)$ and suppose that $\mu_1(s)$ is the eigenvalue with mrp of $M^*(s)$. In appendix B it is proved that $M^*(s)$ is irreducible for $s\geq 0$ and that $\mu_1(s)$ is differentiable for $s\geq 0$. $M^*(s)$ being irreducible, the Perron–Frobenius theorem states that $\mu_1(s)$ is real, simple, positive and strictly decreasing in s.

Next, write $M^*(s)=S^{-1}(s)\ D(s)\ S(s)$, where

$$D(s) = \left[\begin{array}{c|c} \mu_1(s) & 0 \\ \hline 0 & \mathfrak{D}(s) \end{array}\right]$$

and

i. $\mathfrak{D}(s)=\mathrm{diag}(\mu_2(s),..,\mu_n(s))$ if $M^*(s)$ is diagonalisable

ii. $\mathfrak{D}(s)$ is a Jordan normal matrix if $M^*(s)$ is not diagonalisable.

Setting

$$(I-M^*(s))^{-1} = S^{-1}(s)\ (I-D(s))^{-1}\ S(s),$$

it follows from (5.28) that

$$\lim_{s\downarrow 0} s\ a^*(s) = \frac{S^{-1}(0)_{.1}\ S(0)_{1.}\ b^*_a(0)}{-\dot{\mu}_1(0)}, \tag{5.30}$$

since, irrespective of $M^*(s)$ being diagonalisable, $\mu_1(s)$ has multiplicity one.

As before the notation $\dot{\mu}_1(0)$ is used to denote the derivative of the function $\mu_1(s)$ at s=0. Assumed that the limiting availability exists, (5.30) yields $0<-\dot{\mu}_1(0)\leq S^{-1}(0)_{.1}\ S(0)_{1.}\ b_a^*(0)$.

Expression (5.30) confirms that the limiting availability is independent of the initial condition, as stated in (5.26). To show this, note that $S(0)_{1.}$ and $S^{-1}(0)_{.1}$ are the left and right eigenvector respectively, corresponding to the eigenvalue $\mu_1(0)=1$. Hence $S(0)_{1.}$ and $S^{-1}(0)_{.1}$ satisfy

$$S(0)_{1.}M^*(0) = S(0)_{1.}$$

and

$$M^*(0)\ S^{-1}(0)_{.1} = S^{-1}(0)_{.1}.$$

Further, since $M^*(0)$ is a stochastic matrix, $S(0)_{1.}$ is (apart from a multiplicative constant) the stationary probability vector corresponding to $M^*(0)$ and $S^{-1}(0)_{.1}$ is the n-dimensional unit column vector e_n. Clearly $S^{-1}(0)_{.1}\ S(0)_{1.}$ is a matrix with identical rows and (5.26) follows.

Note that practical computation of (5.30) demands the use of numerical software in order to compute $\mu_1(s)$ in a neigbourhood of s=0 and to get an approximation of $\dot{\mu}_1(0)$.

5.5 System state probabilities

Define the state probabilities $P_{ij,kl}(t)$, for the allowed pairs (i,j) and $(k,l)\in\{o,1,..,n_1,r,w\}x\{o,1,..,n_2,r,w\}$ as

$$P_{ij,kl}(t) = Pr\{X(t)=(k,l)\ |\ E_{(i,j)} \text{ at } t=0\}.$$

Considering the mutually exclusive and exhaustive cases used to derive the expressions for the availability functions (5.19)–(5.21) immediately yields the system state probabilities.

For $(i,j)\in\{o,1,..,n_1\}x\{o,1,..,n_2\}$ and $(k,l)\in\{1,2,..,n_1,r\}$ x $\{1,2,..,n_2,r,w\}$,

$$P_{ij,kl}(t) = P_{2,jl}(t)\ \exp(T_1t)_{ik}\ \exp(-\lambda t)\ (1-\delta_{kr}) \tag{5.31}$$

$$+\sum_{h=1}^{n_2}\int_0^t f_{1,i}(u)\ P_{2,jh}(u)\ \exp(-\lambda u)\ P_{rh,kl}(t-u)\ du$$

$$+\int_0^t\int_0^u\int_0^v f_{1,i}(v)\ P_{2,jr}(v,w)\ \exp(-\lambda v)\ \frac{g_2(u-v+w)}{\bar{G}_2(w)}\ P_{ro,kl}(t-u)\ dw\ dv\ du$$

$$+\sum_{h=1}^{n_2}\int_0^t \bar{F}_{1,i}(u)\ P_{2,jh}(u)\ \lambda\exp(-\lambda u)\ \left(p\ P_{rw,kl}(t-u)\ +(1-p)\ P_{wr,kl}(t-u)\right)\ du$$

$$+\int_0^t\int_0^u\int_0^v \bar{F}_{1,i}(v)\ P_{2,jr}(v,w)\ \lambda\exp(-\lambda v)\ \frac{g_2(u-v+w)}{\bar{G}_2(w)}\ P_{ro,kl}(t-u)\ dw\ dv\ du$$

where for all k,l

$$P_{rw,kl}(t) = \int_0^t g_1(u)\ P_{or,kl}(t-u)\ du\ +\ \bar{G}_1(t)\ \delta_{lw}$$

and

$$P_{wr,kl}(t) = \int_0^t g_2(u)\ P_{ro,kl}(t-u)\ du\ +\ \bar{G}_2(t)\ \delta_{kw}.$$

In (5.31) δ is Kronecker's δ and the probability p and the function $P_{2,jr}(v,w)$ are defined as in 5.4. Finally, $P_{2,jr}(t)$ represents the probability that component C_2 is found in the repair state at time t, given a start in phase i at t=0 and given that C_1 does not fail until time t. Hence (*cf.* (5.4) and (5.5)), for $j\in\{o,1,..,n_2\}$,

$$P_{2,jr}(t) = \sum_{n=0}^{\infty}\left[f_{2,o}(t)*g_2(t)\right]^{(n)}*\ f_{2,j}(t)*\bar{G}_2(t).$$

Note that for $(i,j)\in\{o,1,..,n_1\}x\{o,1,..,n_2\}$ and $(k,l)=(w,r)$ the first term in the rhs of (5.31) is replaced by

$$\int_0^t \int_0^v f_{1,i}(v)\ P_{2,jr}(v,w)\ \exp(-\lambda v)\ \frac{\bar{G}_2(t-v+w)}{\bar{G}_2(w)}\ dw\ dv$$

$$+\int_0^t \int_0^v \bar{F}_{1,i}(v)\ P_{2,jr}(v,w)\ \lambda\exp(-\lambda v)\ \frac{\bar{G}_2(t-v+w)}{\bar{G}_2(w)}\ dw\ dv, \tag{5.32}$$

i.e. the probability that component C_1 *or* the common cause component fails before time t, while component C_2 is in repair at the moment of failure and the repair of C_2 exceeds time t. Obviously, (5.32) is obtained from the third and fifth term in the rhs of (5.31).

The expressions for $P_{ir,kl}(t)$ and $P_{rj,kl}(t)$ are derived by considering the instant of repair completion of C_2 and C_2 respectively, as in (5.20) and (5.21).

For $i\in\{o,1,..,n_1\}$ and $(k,l)\in\{1,..,n_1,r\}x\{1,..,n_2,r,w\}$,

$$P_{ir,kl}(t) = \bar{G}_2(t)\ \exp(-\lambda t)\ \exp(T_1 t)_{ik}\ \delta_{lr} \tag{5.33}$$

$$+\sum_{h=1}^{n_1} \int_0^t g_2(u)\ \alpha_{1,i}^T\ (\exp(T_1 u))_{.h}\ \exp(-\lambda u)\ P_{ho,kl}(t-u)\ du$$

$$+\int_0^t \int_0^u f_{1,i}(v)\ \exp(-\lambda v)\ dv\ g_2(u)\ P_{ro,kl}(t-u)\ du$$

$$+\int_0^t \int_0^u \lambda\exp(-\lambda v)\ \bar{F}_{1,i}(v)\ dv\ g_2(u)\ P_{ro,kl}(t-u)\ du.$$

The function $P_{ir,wr}(t)$ is obtained by replacing the first term in the rhs of (5.33) by

$$\bar{G}_2(t) \left\{\int_0^t f_{1,i}(u)\ \exp(-\lambda u)\ du + \int_0^t \lambda\ \exp(-\lambda u)\ \bar{F}_{1,i}(u)\ du \right\}, \tag{5.34}$$

the probability that the repair time of C_2 exceeds time t and a failure of component C_1 *or* the common cause component occurs before t. Interchanging the roles of the components C_1 and C_2 gives, for $j\in\{o,1,..,n_2\}$ and $(k,l)\in\{1,..,n_1,r,w\}x\{1,..,n_2,r\}$,

$$P_{rj,kl}(t) = \bar{G}_1(t)\ \exp(-\lambda t)\ \exp(T_2 t)_{jl}\ \delta_{kr} \tag{5.35}$$

$$+\sum_{h=1}^{n_2} \int_0^t g_1(u)\ \alpha_{2,j}^T\ (\exp(T_2 u))_{.h}\ \exp(-\lambda u)\ P_{oh,kl}(t-u)\ du$$

$$+\int_0^t \int_0^u f_{2,j}(v)\ \exp(-\lambda v)\ dv\ g_1(u)\ P_{or,kl}(t-u)\ du$$

$$+\int_0^t \int_0^u \lambda\exp(-\lambda v)\ \bar{F}_{2,j}(v)\ dv\ g_1(u)\ P_{or,kl}(t-u)\ du.$$

Analogously to $P_{ir,wr}(t)$, the function $P_{rj,rw}(t)$ is obtained replacing the first term in the rhs of (5.35) by

$$\bar{G}_1(t) \left\{ \int_0^t f_{2,j}(u)\ \exp(-\lambda u)\ du + \int_0^t \lambda\ \exp(-\lambda u)\ \bar{F}_{2,j}(u)\ du \right\}. \tag{5.36}$$

The stationary state probabilities

$$\pi_{kl} = \lim_{t\to\infty} P_{ij,kl}(t)$$

are given by

$$\pi_{kl} = \lim_{s\downarrow 0} s\ P^*_{ij,kl}(s),$$

where $P^*_{ij,kl}(s)$ is the Laplace transform of $P_{ij,kl}(t)$. As for the reliability and availability functions, taking the Laplace transforms of the system state probabilities results, for any pair (k,l), in a set of linear equations in $P^*_{ij,kl}(s)$:

$$P^*_{ij,kl}(s) = \mathfrak{L}\left(P_{2,jl}(t)\ \exp(T_1 t)_{ik}\ \exp(-\lambda t)\ (1-\delta_{kr})\right) \tag{5.37}$$

$$+\sum_{h=1}^{n_2} \mathfrak{L}\left[f_{1,i}(t)\ P_{2,jh}(t)\ \exp(-\lambda t)\right]\ P^*_{rh,kl}(s)$$

$$+\mathfrak{L}\left[\int_0^t \int_0^v f_{1,i}(v)\ P_{2,jr}(v,w)\ \exp(-\lambda v)\ \frac{g_2(t-v+w)}{\bar{G}_2(w)}\ dw\ dv\right]\ P^*_{ro,kl}(s)$$

$$+\sum_{h=1}^{n_2} \mathfrak{L}\left[\bar{F}_{1,i}(t)\ P_{2,jh}(t)\ \lambda\exp(-\lambda t)\right]\ \left(pP^*_{rw,kl}(s) + (1-p)P^*_{wr,kl}(s)\right)$$

$$+\mathfrak{L}\left[\int_0^t \int_0^v \bar{F}_{1,i}(v)\ P_{2,jr}(v,w)\ \lambda\exp(-\lambda v)\ \frac{g_2(t-v+w)}{\bar{G}_2(w)}\ dw\ dv\right]\ P^*_{ro,kl}(s),$$

where

$$P^*_{rw,kl}(s) = g_1^*(s)\ P^*_{or,kl}(s) + \bar{G}_1^*(s)\ \delta_{lw}$$

and

$$P^*_{wr,kl}(s) = g_2^*(s)\ P^*_{ro,kl}(s) + \bar{G}_2^*(s)\ \delta_{kw}.$$

Subsequently,

$$\begin{aligned} P^*_{ir,kl}(s) = {} & \mathfrak{L}\left(\exp(T_1 t)_{ik}\ \exp(-\lambda t)\ \bar{G}_2(t)\right)\ \delta_{lr} \\ & + \sum_{h=1}^{n_1} \mathfrak{L}\left[g_2(t)\ \alpha_{1,i}^T\ (\exp(T_1 t))_{.h}\ \exp(-\lambda t)\right]\ P^*_{ho,kl}(s) \\ & + \mathfrak{L}\left[g_2(t) \int_0^t \left(f_{1,i}(u) + \lambda \bar{F}_{1,i}(u)\right)\ \exp(-\lambda u)\ du\right]\ P^*_{ro,kl}(s) \end{aligned} \tag{5.38}$$

and

$$\begin{aligned} P^*_{rj,kl}(s) = {} & \mathfrak{L}\left(\exp(T_2 t)_{jl}\ \exp(-\lambda t)\ \bar{G}_1(t)\right)\ \delta_{kr} \\ & + \sum_{h=1}^{n_2} \mathfrak{L}\left[g_1(t)\ \alpha_{2,j}^T\ (\exp(T_2 t))_{.h}\ \exp(-\lambda t)\right]\ P^*_{oh,kl}(s) \\ & + \mathfrak{L}\left[g_1(t) \int_0^t \left(f_{2,j}(u) + \lambda \bar{F}_{2,j}(u)\right)\ \exp(-\lambda u)\ du\right]\ P^*_{or,kl}(s). \end{aligned} \tag{5.39}$$

In the appropriate cases, the Laplace transforms of (5.32), respectively (5.34) and (5.36) replace the first term in the rhs of (5.37)–(5.39). Further, as the expressions (5.31), (5.33) and (5.35) are obtained by considering exactly the same mutually exclusive and exhaustive events as in (5.19)–(5.21), for any pair (k,l) the coefficients of $P^*_{ij,kl}(s)$ in (5.37)–(5.39) are identical to the coefficients of $A^*_{ij}(s)$ in (5.22)–(5.24). Hence, the computation of π_{kl} imitates the computation of the limiting availability A: for any pair (k,l), the set of equations (5.37)–(5.39) is written as

$$\pi^*_{kl}(s) = M^*(s)\ \pi^*_{kl}(s) + b^*_{\pi,kl}(s), \tag{5.40}$$

where $\pi^*_{kl}(s)$ is a column vector containing the quantities $P^*_{ij,kl}(s)$, $b^*_{\pi,kl}(s)$ is the vector with the first term on the rhs of (5.37)–(5.39) and $M^*(s)$ is as in (5.27).

Subsequently, from (5.30)

$$\lim_{s \downarrow 0} s\ \pi^*_{kl}(s) = \frac{S^{-1}(0)_{.1}\ S(0)_{1.}\ b^*_{\pi,kl}(0)}{-\ \dot{\mu}_1(0)}. \tag{5.41}$$

Thus, $S^{-1}(0)_{.1}$, $S(0)_{1.}$ and $-\ \dot{\mu}_1(0)$ being calculated once to obtain the limiting availability A, the amount of work required for the computation of π_{kl} reduces to the computation of $(n_1+1)(n_2+1)+1$ vectors $b^*_{\pi,kl}(0)$. Using the diagonalisation $T_k=S_k^{-1}D_kS_k$, the components of $b^*_{\pi,kl}(s)$ can be found as

$$\mathfrak{L}\left(P_{2,jl}(t)\ \exp(T_1t)_{ik}\ \exp(-\lambda t)\right) = \left(S_1^{-1}\ P^*_{2,jl}((\lambda+s)I_1-D_1)\ S_1\right)_{ik},$$

$$\mathfrak{L}\left(\exp(T_1t)_{ik}\ \exp(-\lambda t)\ \bar{G}_2(t)\right) = \left(S_1^{-1}\ \bar{G}^*_2((\lambda+s)I_1-D_1)\ S_1\right)_{ik}$$

and

$$\mathfrak{L}\left(\exp(T_2t)_{jl}\ \exp(-\lambda t)\ \bar{G}_1(t)\right) = \left(S_2^{-1}\ \bar{G}^*_1((\lambda+s)I_2-D_2)\ S_2\right)_{jl}.$$

5.6 Interval reliability

The interval reliability $R_{ij}(t,\tau)$ of the *dependent* parallel system is

$$R_{ij}(t,\tau) = \Pr\{X(u)\neq(r,w)\ \wedge\ X(u)\neq(w,r),\ t\leq u\leq t+\tau\ \mid\ E_{(i,j)}\ \text{at}\ t=0\}.$$

Note that, in contrast to (5.1), the *interval* reliability of a system with dependent components does *not* equal the product of the interval reliability of a parallel system with two independent components and the interval reliability of the common cause component. Of course, the reason is that the availability of the independent parallel system consisting of the components C_1 and C_2, is influenced by the availability of the common cause component and *vice versa*.
However, given the derivation of $A_{ij}(t)$ and $P_{ij,kl}(t)$ in the sections 5.4 and 5.5, it is easily found, for $(i,j)\in\{o,1,..,n_1\}\times\{o,1,..,n_2\}$,

$$R_{ij}(t,\tau) = \bar{F}_{1,i}(t+\tau)\ \exp(-\lambda(t+\tau)) \tag{5.42}$$

$$+ \sum_{k=1}^{n_2} \int_0^t f_{1,i}(u)\ P_{2,jk}(u)\ \exp(-\lambda u)\ R_{rk}(t-u,\tau)\ du$$

$$+ \int_0^t \int_0^u \int_0^v f_{1,i}(v)\ P_{2,jr}(v,w)\ \exp(-\lambda v)\ \frac{g_2(u-v+w)}{\bar{G}_2(w)}\ R_{ro}(t-u,\tau)\ dw\ dv\ du$$

$$+ \sum_{k=1}^{n_2} \int_0^t \bar{F}_{1,i}(u)\ P_{2,jk}(u)\ \lambda\exp(-\lambda u)\ \left[p\ R_{rw}(t-u,\tau) + (1-p)\ R_{wr}(t-u,\tau)\right]\ du$$

$$+ \int_0^t \int_0^u \int_0^v \bar{F}_{1,i}(v)\ P_{2,jr}(v,w)\ \lambda\exp(-\lambda v)\ \frac{g_2(u-v+w)}{\bar{G}_2(w)}\ R_{ro}(t-u,\tau)\ dw\ dv\ du$$

$$+ \sum_{k=1}^{n_2} \int_t^{t+\tau} f_{1,i}(u)\ P_{2,jk}(u)\ \exp(-\lambda u)\ R_{rk}^{dep}(t+\tau-u)\ du,$$

where $R_{rk}^{dep}(.)$ is the reliability function as defined in section 5.2.

Obviously,

$$R_{rw}(t,\tau) = \int_0^t g_1(u)\ R_{or}(t-u,\tau)\ du$$

and

$$R_{wr}(t,\tau) = \int_0^t g_2(u)\ R_{ro}(t-u,\tau)\ du.$$

As in the expressions for $A_{ij}(t)$ and $P_{ij,kl}(t)$, the first term in the rhs of (5.42) gives the probability that C_1 and the common cause component do not fail and the next four terms represent the probability that the first failure of C_1 or the common cause component occurs in (0,t). In fact, the only term which needs any comment is the last term in the rhs of (5.42): it gives the probability that the first failure of C_1 occurs in $(t,t+\tau)$ while the common cause component survives C_1. Of course, failures of the common cause component in $(t,t+\tau)$ are not allowed.

Analogously, by joining one additional term, representing the probability that the first failure of C_1 occurs in $(t,t+\tau)$ and the common cause component does not fail before C_1, it follows for $i\in\{o,1,..,n_1\}$

$$R_{ir}(t,\tau) = \bar{G}_2(t+\tau)\ \bar{F}_{1,i}(t+\tau)\ \exp(-\lambda(t+\tau)) \tag{5.43}$$

$$+\int_0^t g_2(u) \int_0^u \left(f_{1,i}(v) + \lambda\ \bar{F}_{1,i}(v)\right)\ \exp(-\lambda v)\ dv\ R_{ro}(t-u,\tau)\ du$$

$$+\sum_{k=1}^{n_1} \int_0^t g_2(u)\ \alpha_{1,i}^T\ \exp(T_1 u)_{.k}\ \exp(-\lambda u)\ R_{ko}(t-u,\tau)\ du$$

$$+\sum_{k=1}^{n_1} \int_t^{t+\tau} g_2(u)\ \alpha_{1,i}^T\ \exp(T_1 u)_{.k}\ \exp(-\lambda u)\ R_{ko}^{dep}(t+\tau-u)\ du.$$

Interchanging the role of the components gives

$$R_{rj}(t,\tau) = \bar{G}_1(t+\tau)\ \bar{F}_{2,j}(t+\tau)\ \exp(-\lambda(t+\tau)) \tag{5.44}$$

$$+\int_0^t g_1(u) \int_0^u \left(f_{2,j}(v) + \lambda\ \bar{F}_{2,j}(v)\right)\ \exp(-\lambda v)\ dv\ R_{or}(t-u,\tau)\ du$$

$$+\sum_{k=1}^{n_2} \int_0^t g_1(u)\ \alpha_{2,j}^T\ \exp(T_2 u)_{.k}\ \exp(-\lambda u)\ R_{ok}(t-u,\tau)\ du$$

$$+\sum_{k=1}^{n_2} \int_t^{t+\tau} g_1(u)\ \alpha_{2,j}^T\ \exp(T_2 u)_{.k}\ \exp(-\lambda u)\ R_{ok}^{dep}(t+\tau-u)\ du.$$

As in section 3.6, a special point of interest is $\Re(\tau)$, the stationary interval reliability for an interval of length τ. From the regenerative nature of the events $E_{(i,j)}$ it is clear that $\Re(\tau)$ does not depend upon the initial condition of the system.

Defining $R^*_{ij}(s,\tau)$ as in section 3.6, it follows from (5.42)–(5.44)

$$R^*_{ij}(s,\tau) = \mathfrak{L}\left(\bar{F}_{1,i}(t+\tau)\ \exp(-\lambda(t+\tau))\right) \tag{5.45}$$

$$+\sum_{k=1}^{n_2} \mathfrak{L}\left[f_{1,i}(t)\ P_{2,jk}(t)\ \exp(-\lambda t)\right]\ R^*_{rk}(s,\tau)$$

$$+\mathfrak{L}\left[\int_0^t \int_0^v f_{1,i}(v)\ P_{2,jr}(v,w)\ \exp(-\lambda v)\ \frac{g_2(t-v+w)}{\bar{G}_2(w)}\ dw\ dv\right]\ R^*_{ro}(s,\tau)$$

$$+\sum_{k=1}^{n_2} \mathfrak{L}\left[\bar{F}_{1,i}(t)\ P_{2,jk}(t)\ \lambda\exp(-\lambda t)\right]\ \left(p\ R^*_{rw}(s,\tau)\ +(1-p)\ R^*_{wr}(s,\tau)\right)$$

$$+\mathfrak{L}\left[\int_0^t \int_0^v \bar{F}_{1,i}(v)\ P_{2,jr}(v,w)\ \lambda\exp(-\lambda v)\ \frac{g_2(t-v+w)}{\bar{G}_2(w)}\ dw\ dv\right]\ R^*_{ro}(s,\tau)$$

$$+\sum_{k=1}^{n_2} \mathfrak{L}\left[\int_t^{t+\tau} f_{1,i}(u)\ P_{2,jk}(u)\ \exp(-\lambda u)\ R^{dep}_{rk}(t+\tau-u)\ du\right],$$

where $R^*_{rw}(s,\tau) = g^*_1(s)\ R^*_{or}(s,\tau)$ and $R^*_{wr}(s,\tau) = g^*_2(s)\ R^*_{ro}(s,\tau)$,

$$R^*_{ir}(s,\tau) = \mathfrak{L}\left(\bar{G}_2(t+\tau)\ \bar{F}_{1,i}(t+\tau)\ \exp(-\lambda(t+\tau))\right) \tag{5.46}$$

$$+\mathfrak{L}\left[g_2(t) \int_0^t \left(f_{1,i}(v) + \lambda\ \bar{F}_{1,i}(v)\right)\ \exp(-\lambda v)\ dv\right]\ R^*_{ro}(s,\tau)$$

$$+\sum_{k=1}^{n_1} \mathfrak{L}\left[g_2(t)\ \alpha^T_{1,i}\ \exp(T_1 t)_{.k}\ \exp(-\lambda t)\right]\ R^*_{ko}(s,\tau)$$

$$+\sum_{k=1}^{n_1} \mathfrak{L}\left[\int_t^{t+\tau} g_2(u)\ \alpha^T_{1,i}\ \exp(T_1 u)_{.k}\ \exp(-\lambda u)\ R^{dep}_{ko}(t+\tau-u)\ du\right]$$

and

$$R^*_{rj}(s,\tau) = \mathfrak{L}\left(\bar{G}_1(t+\tau)\ \bar{F}_{2,j}(t+\tau)\ \exp(-\lambda(t+\tau))\right) \tag{5.47}$$

$$+\mathfrak{L}\left[g_1(t) \int_0^t \left(f_{2,j}(v) + \lambda\ \bar{F}_{2,j}(v)\right)\ \exp(-\lambda v)\ dv\right]\ R^*_{or}(s,\tau)$$

$$+\sum_{k=1}^{n_2} \mathfrak{L}\left[g_1(t)\ \alpha^T_{2,j}\ \exp(T_2 t)_{.k}\ \exp(-\lambda t)\right]\ R^*_{ok}(s,\tau)$$

$$+\sum_{k=1}^{n_2} \mathfrak{L}\left[\int_t^{t+\tau} g_1(u)\ \alpha^T_{2,j}\ \exp(T_2 u)_{.k}\ \exp(-\lambda u)\ R^{dep}_{ok}(t+\tau-u)\ du\right].$$

In order to find the stationary interval reliability, the set of equations (5.45)–(5.47) is written as

$$r^*(s,\tau) = M^*(s)\ r^*(s,\tau) + b^*_r(s,\tau). \tag{5.48}$$

As before, $r^*(s,\tau)$ is a vector containing quantities $R^*_{ij}(s,\tau)$, $M^*(s)$ is still as in section 5.4 and the vector $b^*_r(s,\tau)$ contains the appropriate terms of the rhs of (5.45)–(5.47). It is immediate that

$$\lim_{s\downarrow 0} s\, r^*(s,\tau) = \frac{S^{-1}(0)_{.1}\ S(0)_{1.}\ b^*_r(0,\tau)}{-\ \dot{\mu}_1(0)} \tag{5.49}$$

and hence only the computation of the vector $b^*_r(0,\tau)$ is needed for the calculation of $\Re(\tau)$, for fixed τ. However, remark that computation of $b^*_r(0,\tau)$ requires the calculation of the last term in the expressions (5.45)–(5.47), which can not be obtained explicitly by the diagonalisation of T_1 and T_2. The same holds for the first term in (5.46) and (5.47) and hence numerical integration is needed to obtain $b^*_r(0,\tau)$ and thereby $\Re(\tau)$.

5.7 Joint availability

The joint availability $A_{ij}(t,\tau)$ of the dependent parallel system is defined by

$$A_{ij}(t,\tau) = \Pr\{X(u)\neq(r,w) \wedge X(u)\neq(w,r),\ u=t,t+\tau \mid E_{(i,j)} \text{ at } t=0\}.$$

An equation for $A_{ij}(t,\tau)$, $(i,j)\in\{o,1,..,n_1\}x\{o,1,..,n_2\}$ is obtained, as in (5.19), by considering the moment at which the first failure of C_1 or the common cause component occurs. This moment can occur in the interval $(0,t)$, in $(t,t+\tau)$ or after time $t+\tau$. It is easily seen that these three mutually exclusive and exhaustive events yield:

$$A_{ij}(t,\tau) = \bar{F}_{1,i}(t+\tau)\ \exp(-\lambda(t+\tau)) \tag{5.50}$$

$$+\sum_{k=1}^{n_2}\int_0^t f_{1,i}(u)\ P_{2,jk}(u)\ \exp(-\lambda u)\ A_{rk}(t-u,\tau)\ du$$

$$+\sum_{k=1}^{n_2}\int_t^{t+\tau} f_{1,i}(u)\ P_{2,jk}(u)\ \exp(-\lambda u)\ A_{rk}(t+\tau-u)\ du$$

$$+\sum_{k=1}^{n_2}\int_0^t \bar{F}_{1,i}(u)\ P_{2,jk}(u)\ \lambda\exp(-\lambda u)\ \left[p\ A_{rw}(t-u,\tau) + (1-p)\ A_{wr}(t-u,\tau)\right]\ du$$

$$+\sum_{k=1}^{n_2} \int_t^{t+\tau} \bar{F}_{1,i}(u)\ P_{2,jk}(u)\ \lambda\exp(-\lambda u)\ \left[p\ A_{rw}(t+\tau-u) + (1-p)\ A_{wr}(t+\tau-u)\right]\ du$$

$$+\int_0^t\int_0^u\int_0^v \left[f_{1,i}(v)+\lambda \bar{F}_{1,i}(v)\right]\ P_{2,jr}(v,w)\ \exp(-\lambda v)\ \frac{g_2(u-v+w)}{\bar{G}_2(w)}\ A_{ro}(t-u,\tau)\ dw\ dv\ du$$

$$+\int_t^{t+\tau}\int_t^u\int_0^v \left[f_{1,i}(v)+\lambda \bar{F}_{1,i}(v)\right]\ P_{2,jr}(v,w)\ \exp(-\lambda v)\ \frac{g_2(u-v+w)}{\bar{G}_2(w)}\ A_{ro}(t+\tau-u)\ dw\ dv\ du$$

where

$$A_{rw}(t,\tau) = \int_0^t g_1(u)\ A_{or}(t-u,\tau)\ du$$

and

$$A_{wr}(t,\tau) = \int_0^t g_2(u)\ A_{ro}(t-u,\tau)\ du.$$

The function $A_{ij}(.)$ is the availability function, as defined in section 5.4.

Similarly, starting in state (i,r), the repair of component C_2 is completed in the interval (0,t), in $(t,t+\tau)$ or after time $t+\tau$. Hence, arguing as in the derivation of (5.20) gives

$$A_{ir}(t,\tau) = \bar{F}_{1,i}(t+\tau)\ \exp(-\lambda(t+\tau))\ \bar{G}_2(t+\tau) \tag{5.51}$$

$$+\sum_{k=1}^{n_1} \int_0^t g_2(u)\ \alpha_{1,i}^T\ \exp(T_1 u)_{.k}\ \exp(-\lambda u)\ A_{ko}(t-u,\tau)\ du$$

$$+\int_0^t g_2(u) \int_0^u \left[f_{1,i}(v) + \lambda\ \bar{F}_{1,i}(v)\right]\ \exp(-\lambda v)\ dv\ A_{ro}(t-u,\tau)\ du$$

$$+\sum_{k=1}^{n_1} \int_t^{t+\tau} g_2(u)\ \alpha_{1,i}^T\ \exp(T_1 u)_{.k}\ \exp(-\lambda u)\ A_{ko}(t+\tau-u)\ du$$

$$+\int_t^{t+\tau} g_2(u) \int_t^u \left[f_{1,i}(v) + \lambda\ \bar{F}_{1,i}(v)\right]\ \exp(-\lambda v)\ dv\ A_{ro}(t+\tau-u)\ du$$

and by analogy

$$A_{rj}(t,\tau) = \bar{F}_{2,j}(t+\tau)\ \exp(-\lambda(t+\tau))\ \bar{G}_1(t+\tau) \tag{5.52}$$

$$+ \sum_{k=1}^{n_2} \int_0^t g_1(u)\ \alpha_{2,j}^T\ \exp(T_2 u)_{.k}\ \exp(-\lambda u)\ A_{ok}(t-u,\tau)\ du$$

$$+ \int_0^t g_1(u) \int_0^u \left(f_{2,j}(v) + \lambda\ \bar{F}_{2,j}(v)\right)\ \exp(-\lambda v)\ dv\ A_{or}(t-u,\tau)\ du$$

$$+ \sum_{k=1}^{n_2} \int_t^{t+\tau} g_1(u)\ \alpha_{2,j}^T\ \exp(T_2 u)_{.k}\ \exp(-\lambda u)\ A_{ok}(t+\tau-u)\ du$$

$$+ \int_t^{t+\tau} g_1(u) \int_t^u \left(f_{2,j}(v) + \lambda\ \bar{F}_{2,j}(v)\right)\ \exp(-\lambda v)\ dv\ A_{or}(t+\tau-u)\ du.$$

For the stationary joint availability function $\mathfrak{A}(\tau)$, defined as in section 3.7, it is clear that

$$\mathfrak{A}(\tau) = \lim_{s\downarrow 0} s\ A^*_{ij}(s,\tau).$$

Secondly, it is clear that taking the Laplace transforms of (5.50)–(5.52) gives a set of linear equations in $A^*_{ij}(s,\tau)$, which can be written as

$$a^*(s,\tau) = M^*(s)\ a^*(s,\tau) + b^*_a(s,\tau). \tag{5.53}$$

Obviously, $a^*(s,\tau)$ is the vector containing quantities $A^*_{ij}(s,\tau)$, $M^*(s)$ is as in (5.27), (5.40) and (5.48) and $b^*_a(s,\tau)$ is the vector containing the appropriate terms of the rhs of the Laplace transforms of (5.50)–(5.52). Thirdly, from (5.53)

$$\lim_{s\downarrow 0} s\ a^*(s,\tau) = \frac{S^{-1}(0)_{.1}\ S(0)_{1.}\ b^*_a(0,\tau)}{-\ \dot{\mu}_1(0)} \tag{5.54}$$

and the only problem in the calculation of the limiting joint availability lies in the computation of the vector $b^*_a(0,\tau)$. The problems in obtaining $b^*_a(0,\tau)$ are identical to the problems in the computation of $b^*_r(0,\tau)$ in (5.49): explicit expressions for the Laplace transforms of the integrals with integration area $(t,t+\tau)$ in (5.50)–(5.52) are not available and hence numerical integration is needed to approximate $b^*_a(0,\tau)$.

5.8 Quasi-stationary distributions

In this section the quasi-stationary state probabilities and the limiting residual life time distribution are derived for the two-unit dependent parallel system. As explained earlier, both distributions are relevant if the system under consideration fails rarely. The two questions arising when system failures are exceptional, are: what is the distribution of the residual system life time, under the condition that no system failure has occurred until time t (for large t)? And what is the probability of being in state (i,j)? Both questions will be answered here.

The approach followed is more general than in chapter 3, where not only the components C_1 and C_2 have exponentially distributed life times, but also the common cause component C_3. In this section it is shown, straight from the definition of the two kinds of quasi-stationarity, that

i. the quasi-stationary state probabilities are independent of the distribution of the common cause component and

ii. the limiting residual life time distribution can be decomposed into the product of two terms: one concerning a parallel system with independent components and one concerning the common cause component.

Subsequently, explicit expressions are derived for the quasi-stationary distributions. Starting point of the analysis is the reliability function $R_{ij}^{dep}(t)$. From (5.1),

$$R_{ij}^{dep}(t) = R_{ij}^{ind}(t)\, R_{cc}(t), \tag{5.55}$$

where $R_{ij}^{ind}(t)$ is given by (5.2) and (5.6)–(5.8).

Further, define for the pairs (i,j) and $(k,l)\in\{o,1,..,n_1,r\}\mathrm{x}\{o,1,..,n_2,r\}$

$$\mathfrak{P}_{ij,kl}^{dep}(t) = \Pr\{X(t)=(k,l),\ X(u)\neq(r,w),\ X(u)\neq(w,r),\ 0\leq u\leq t \mid E_{(i,j)} \text{ at } t=0\},$$

i.e. $\mathfrak{P}_{ij,kl}^{dep}(t)$ gives the probability that the two-unit dependent parallel system is in state (k,l) at time t and no system failure has occurred in (0,t), given an occurrence of $E_{(i,j)}$ at t=0.Then

$$\mathfrak{P}_{ij,kl}^{dep}(t) = \mathfrak{P}_{ij,kl}^{ind}(t)\, R_{cc}(t), \tag{5.56}$$

where $\mathfrak{P}^{ind}_{ij,kl}(t)$ represents the probability that the two-unit parallel system with ***independent*** units is in state (k,l) at time t and no system failure has occurred in (0,t), conditioned on an $E_{(i,j)}$ event at the time origin.
Further, by definition the quasi-stationary state probabilities q_{kl} are

$$q_{kl} = \lim_{t\to\infty} q_{ij,kl}(t)$$

where

$$q_{ij,kl}(t) = \Pr\{X(t)=(k,l) \mid X(u)\neq(r,w),\ X(u)\neq(w,r),\ 0\leq u\leq t,\ E_{(i,j)} \text{ at } t=0\}.$$

The above expression for $q_{ij,kl}(t)$ represents the probability that the two-unit independent parallel system is in state (k,l) at time t, given that no system failure has occurred in (0,t) and given that $E_{(i,j)}$ has occurred at t=0. Clearly,

$$q_{ij,kl}(t) = \frac{\mathfrak{P}^{dep}_{ij,kl}(t)}{R^{dep}_{ij}(t)}$$

and hence

$$q_{kl} = \lim_{t\to\infty} \frac{\mathfrak{P}^{dep}_{ij,kl}(t)}{R^{dep}_{ij}(t)}. \tag{5.57}$$

On the other hand, the limiting residual lifetime distribution is

$$q(x) = \lim_{t\to\infty} q_{ij}(t,x),$$

where

$$q_{ij}(t,x) = 1 - \Pr\{X(u)\notin D,\ t\leq u\leq t+x \mid X(u)\notin D,\ 0\leq u\leq t,\ E_{(i,j)} \text{ at } t=0\},$$

and $D=\{(r,w),(w,r)\}$. Hence $q_{ij}(t,x)$ is the residual lifetime distribution at time t, conditioned on an $E_{(i,j)}$ event at the time origin.
Obviously

$$q_{ij}(t,x) = 1 - \frac{R^{dep}_{ij}(t+x)}{R^{dep}_{ij}(t)}$$

and thus

$$q(x) = 1 - \lim_{t\to\infty} \frac{R^{dep}_{ij}(t+x)}{R^{dep}_{ij}(t)}. \tag{5.58}$$

From the regenerative nature of the events $E_{(i,j)}$ and the fact that the states $(i,j)\in\{1,..,n_1,r\}x\{1,..,n_2,r\}$ communicate, it follows that the distributions q_{kl} and $q(x)$ are independent of the initial condition. Further, the probabilities q_{kl} do not depend upon the distribution of the common cause component, since from (5.55)-(5.57)

$$q_{kl} = \lim_{t\to\infty} \frac{\mathfrak{P}^{ind}_{ij,kl}(t)}{R^{ind}_{ij}(t)}. \tag{5.59}$$

Secondly, from (5.55) and (5.58) the limiting residual lifetime distribution $q(x)$ equals one minus the product of the limiting residual *survival* distributions of the system with independent components and the common cause component:

$$q(x) = 1 - \lim_{t\to\infty} \frac{R^{ind}_{ij}(t+x)}{R^{ind}_{ij}(t)} \lim_{t\to\infty} \frac{R_{cc}(t+x)}{R_{cc}(t)}. \tag{5.60}$$

In the following, expressions for the probabilities q_{kl} and the distribution $q(x)$ are derived. In principle the method followed is a generalisation in vector-matrix notation of the method followed in Pijnenburg *et al.* (1991). In short, the Laplace transform technique is used to determine the dominating terms in $R^{ind}_{ij}(t)$ and $\mathfrak{P}^{ind}_{ij,kl}(t)$. It is shown that the pole with mrp of $(I-\mathfrak{M}^*(s))^{-1}$ (where $\mathfrak{M}^*(s)$ is defined as in (5.14)) plays a key role in finding these dominant terms and thereby q_{kl} and $q(x)$.

The functions $R_{ij}(t)$ and their Laplace transforms have been obtained in section 5.2 and are given by (5.2)-(5.13). With respect to the functions $\mathfrak{P}^{ind}_{ij,kl}(t)$ (for ease of use $\mathfrak{P}^{ind}_{ij,kl}(t)$ is abbreviated to $\mathfrak{P}_{ij,kl}(t)$), it is immediate from the derivation of (5.2) and (5.6)-(5.8) that:

for $(i,j)\in\{o,1,..,n_1\}x\{o,1,..,n_2\}$ and $(k,l)\in\{1,2,..,n_1,r\}x\{1,2,..,n_2,r\}$,

$$\mathfrak{P}_{ij,kl}(t)=P_{2,jl}(t)\ \exp(T_1t)_{ik}\ (1-\delta_{kr}) + \sum_{h=1}^{n_2} \int_0^t f_{1,i}(u)\ P_{2,jh}(u)\ \mathfrak{P}_{rh,kl}(t-u)du \tag{5.61}$$

where δ is Kronecker's δ and $P_{2,jh}(t)$ is defined by (5.3).

Subsequently, for $i \in \{o,1,..,n_1\}$ and $(k,l) \in \{1,..,n_1,r\} \times \{1,..,n_2,r\}$,

$$\mathfrak{P}_{ir,kl}(t) = \bar{G}_2(t)\ \exp(T_1 t)_{ik}\ \delta_{lr} + \sum_{h=1}^{n_1} \int_0^t g_2(u)\ \alpha_{1,i}^T\ \exp(T_1 u)_{.h}\ \mathfrak{P}_{ho,kl}(t-u)du \tag{5.62}$$

and for $j \in \{o,1,..,n_2\}$ and $(k,l) \in \{1,..,n_1,r\} \times \{1,..,n_2,r\}$,

$$\mathfrak{P}_{rj,kl}(t) = \bar{G}_1(t)\ \exp(T_2 t)_{jl}\ \delta_{kr} + \sum_{h=1}^{n_2} \int_0^t g_1(u)\ \alpha_{2,j}^T\ \exp(T_2 u)_{.h}\ \mathfrak{P}_{oh,kl}(t-u)du \tag{5.63}$$

Note that $R_{ij}(t) = \sum \mathfrak{P}_{ij,kl}(t)$, where the summation is over the pairs $(k,l) \in \{1,..,n_1,r\} \times \{1,..,n_2,r\}$.
Taking the Laplace transforms yields

$$\mathfrak{P}^*_{ij,kl}(s) = \mathfrak{L}(P_{2,jl}(t)\ \exp(T_1 t)_{ik}\ (1-\delta_{kr})) + \sum_{h=1}^{n_2} \mathfrak{L}\left(f_{1,i}(t)\ P_{2,jh}(t)\right)\ \mathfrak{P}^*_{rh,kl}(s), \tag{5.64}$$

$$\mathfrak{P}^*_{ir,kl}(s) = \mathfrak{L}\left(\exp(T_1 t)_{ik}\ \bar{G}_2(t)\right)\ \delta_{lr} + \sum_{h=1}^{n_1} \mathfrak{L}\left[g_2(t)\ \alpha_{1,i}^T\ \exp(T_1 t)_{.h}\right]\ \mathfrak{P}^*_{ho,kl}(s) \tag{5.65}$$

and

$$\mathfrak{P}^*_{rj,kl}(s) = \mathfrak{L}(\exp(T_2 t)_{jl}\ \bar{G}_1(t))\ \delta_{kr} + \sum_{h=1}^{n_2} \mathfrak{L}\left[g_1(t)\ \alpha_{2,j}^T\ \exp(T_2 t)_{.h}\right]\ \mathfrak{P}^*_{oh,kl}(s). \tag{5.66}$$

Proceeding in matrix notation, from (5.14)

$$r^*(s) = \mathfrak{M}^*(s)\ r^*(s) + b_r^*(s),$$

where an appropriate choice for $r^*(s)$, $\mathfrak{M}^*(s)$ and $b_r^*(s)$ is given in section 5.2. Analogously, write for any pair (k,l),

$$p^*_{kl}(s) = \mathfrak{M}^*(s)\ p^*_{kl}(s) + b^*_{p,kl}(s),$$

where

$$p^*_{kl}(s) = [P^*_{o1,kl}(s)\ \dots\ P^*_{on_2,kl}(s)\ P^*_{r1,kl}(s)\ \dots\ P^*_{rn_2,kl}(s)]$$

and, from (5.64)–(5.66)

$$b^*_{p,kl}(s) = [\mathfrak{L}(P_{2,1l}(t)\alpha_1^T \exp(T_1 t)_{.k})\ \dots\ \mathfrak{L}(P_{2,n_2 l}(t)\alpha_1^T \exp(T_1 t)_{.k})\ 0\ \dots\ 0]^T,$$

$$b^*_{p,kr}(s) = [\mathfrak{L}(P_{2,1r}(t)\alpha_1^T\exp(T_1t)_{.k}) \ \dots \ \mathfrak{L}(P_{2,n_2r}(t)\alpha_1^T\exp(T_1t)_{.k}) \ 0 \ \dots \ 0]^T,$$

$$b^*_{p,rl}(s) = [0 \ \dots \ 0 \ \mathfrak{L}(\bar{G}_1(t)\exp(T_2t)_{1l}) \ \dots \ \mathfrak{L}(\bar{G}_1(t)\exp(T_2t)_{n_2l})]^T.$$

Then

$$r^*(s) = (I-\mathfrak{M}^*(s))^{-1}b^*_r(s) \tag{5.67}$$

and

$$p^*_{kl}(s) = (I-\mathfrak{M}^*(s))^{-1}b^*_{p,kl}(s). \tag{5.68}$$

Since the structure of the equations (5.67) and (5.68) is identical, attention is concentrated at (5.67) and especially on the matrix $(I-\mathfrak{M}^*(s))^{-1}$. Let $s=\delta$ be the pole with mrp of $r^*(s)$, then δ is negative and from Widder (1946, theorem 5b, p.58), δ is real. In appendix C it is proved that δ is a pole of $(I-\mathfrak{M}^*(s))^{-1}$ and not a pole of $b^*_r(s)$ or $\mathfrak{M}^*(s)$, thus

$$\delta = \max_{s\in\mathbb{R}} \{\det(I-\mathfrak{M}^*(s))=0\}.$$

More specifically,

$$\delta = \max_{s\in\mathbb{R}} \{m_1(s)=1\},$$

where $m_1(s)$ is the maximum eigenvalue of $\mathfrak{M}^*(s)$.
It is proved in appendix C that $\mathfrak{M}^*(s)$ exists and is irreducible for $s\geq\delta$. Hence from the Perron–Frobenius theorem $m_1(s)$ is real, positive, strictly decreasing in s and a simple root of the characteristic equation of $\mathfrak{M}^*(s)$, *i.e.* the eigenvalue $m_1(s)$ has multiplicity 1. Proceeding as in section 5.2 it is clear that

$$\lim_{s\downarrow\delta} (s-\delta)\ r^*(s) = \frac{S^{-1}(\delta)_{.1}\ S(\delta)_{1.}\ b^*_r(\delta)}{-\dot{m}_1(\delta)}, \tag{5.69}$$

where $S^{-1}(\delta)_{.1}$ and $S(\delta)_{1.}$ are the right and left eigenvectors respectively of $\mathfrak{M}^*(\delta)$ corresponding to the eigenvalue $m_1(\delta)$. Note that, replacing $M^*(s)$ by $\mathfrak{M}^*(s)$ and $\mu_1(s)$ by $m_1(s)$, it is immediate from appendix B that $m_1(s)$ is differentiable for $s\geq\delta$.

Since the Perron–Frobenius theorem states that the vectors $S^{-1}(\delta)_{.1}$ and $S(\delta)_{1.}$ can be chosen strictly positive, it follows that δ is a simple pole of $r^*(s)$ and hence, from (5.67)

$$r(t) = \frac{S^{-1}(\delta)_{.1}\, S(\delta)_{1.}\, b_r^*(\delta)}{-\dot{m}_1(\delta)} \exp(\delta t) + o(\exp(\delta t)),\ t\to\infty. \tag{5.70}$$

Although numerical experiments have shown that the ratio in the rhs of (5.70) equals one, a formal proof is not available.
Analogously to the reliability in (5.70), the conditional state probabilities satisfy

$$p_{kl}(t) = \frac{S^{-1}(\delta)_{.1}\, S(\delta)_{1.}\, b_{p,kl}^*(\delta)}{-\dot{m}_1(\delta)} \exp(\delta t) + o(\exp(\delta t)),\ t\to\infty. \tag{5.71}$$

and hence, from (5.59) and (5.60),

$$q_{kl} = \frac{\left[S^{-1}(\delta)_{.1}\, S(\delta)_{1.}\, b_{p,kl}^*(\delta)\right]_i}{\left[S^{-1}(\delta)_{.1}\, S(\delta)_{1.}\, b_r^*(\delta)\right]_i}, \tag{5.72}$$

for any $i\in\{1,..,2n_2\}$ and

$$q(x) = 1 - \exp(\delta x) \lim_{t\to\infty} \frac{R_{cc}(t+x)}{R_{cc}(t)}. \tag{5.73}$$

Note that it follows from probabilistic arguments, that the ratio in the rhs of (5.72) is independent of i.
Comparing the derivation of q_{kl} and the derivation of the stationary state probabilities π_{kl} in section 5.5, shows that the essential difference is that $M^*(0)$ in (5.27) is a stochastic matrix, as stated in (5.29). Hence, $S(0)_{1.}$ in (5.30) is, apart from a multiplicative constant, the stationary probability vector corresponding to $M^*(0)$ and $S^{-1}(0)_{.1}$ is a unit column vector. As a result, $S^{-1}(0)_{.1}S(0)_{1.}$ is a matrix with identical rows. On the other side, simple numerical examples show that generally $\mathfrak{M}^*(\delta)$ is *not* a stochastic matrix, *nor* a substochastic matrix and $S(\delta)_{1.}$ in (5.72) is *not* a vector with identical components. However, although $S^{-1}(\delta)_{.1}S(\delta)_{1.}$ is *not* a matrix with identical rows, the value of the quotient in (5.72) is the same for all $i\in\{1,..,2n_2\}$.

With respect to the computation of the pole δ it is shown in appendix C that $\delta \in (\max\{d_{1,1}, d_{2,1}+\sigma_1\}, 0]$, where $d_{k,1}$ is the eigenvalue with mrp of the matrix T_k and σ_k the pole with mrp of $g_k^*(s)$. Interchanging the role of the components C_1 and C_2 yields $\delta \in (\max\{d_{2,1}, d_{1,1}+\sigma_2\}, 0]$ and thus it follows that $\delta \in (\max\{d_{1,1}, d_{2,1}\}, 0]$. Finally, as $\mu_1(s)$ is strictly decreasing, a simple search method (*e.g.* the bisection or secant method) can be used to find δ.

Summarising this chapter, it has been shown how expressions can be derived for the system reliability and mean time to system failure, the transient and stationary interval reliability, (joint) availability and system state probabilities and two quasi-stationary distributions, under BVPH distributed life times and generally distributed repair times. Expressions for the performance measures are in terms of convolution integrals and, applying the techniques of this chapter, the ***stationary*** measures can be computed easily and quickly. On the other side, to obtain the ***transient*** behaviour of the system, it seems preferable to assume phase type distributed repair times. In the latter case the state description process becomes Markovian and the randomisation technique can be used to obtain the transient performance measures, instead of solving sets of integral equations.

6. AN OPPORTUNISTIC REPLACEMENT POLICY

6.1 Introduction

In the previous chapters the statistical dependence of two units in a parallel system has been modelled and methods have been described to compute a number of performance measures. Apart from dependence, maintenance of a system and replacement of units form important 'real world' aspects. Usually, the aim of a maintenance policy is to reduce the (expected) costs of a system per unit of time. In this chapter, the subject of analysis is an opportunistic replacement policy for a two-unit dependent parallel system.

There is an extensive literature on maintenance and replacement models. Thorough surveys are given by McCall (1965), Pierskalla *et al.* (1976), Sherif *et al.* (1981), Sherif (1982), Thomas (1986), Mine (1988) and Valdez-Flores *et al.* (1989). The overwhelming majority of papers mentioned in these surveys deals with one-unit systems or systems that can be considered as a single unit. The number of models and policies for such systems is numerous. The literature on multi-component systems is far less extensive. A survey of maintenance and replacement models for multi-component models is provided by Thomas (1986). The characteristic of a multi-component system is that it is worth considering repair or replacement of one unit in conjunction with what happens to the other units, if there is dependence between the components. Thomas classifies dependence into three categories:

i. Economic dependence. When a system shows economic dependence, the cost structure of replacement has interdependences between the units. Economic dependence reflects the presence of economies of scale, where the costs of a joint maintenance action are less than the sum of the costs of the separate maintenance actions.

ii. Structural dependence. If it is advantageous to replace several units (failed as well as unfailed) at the same time as *e.g.* working units have to be dismantled in order to replace or repair failed ones, the system has a structurally dependent configuration.

iii. Probabilistic or failure dependence occurs when the state of one unit can affect other units or their failure rate, *e.g.* when failure of one unit burdens other units.

Opportunistic policies exploit economics of scale in the maintenance activity: the necessity of performing at least one repair provides the economic justification to do several others at the same time. In this chapter a two-unit parallel redundant system with economic and failure dependence is examined. Clearly, the failure dependence is modelled by allowing the occurrence of common cause failures. The analysed replacement policy is a policy on component level and *not* on system level, *i.e.* the two-unit parallel system is *not* considered as a single unit.

A frequently used technique for the analysis of a multi-component system is to consider a discrete time model and to formulate the replacement problem as a Markov decision process, see *e.g.* Sethi (1977). Another well-known approach is to consider a system under Markovian deterioration, see *e.g.* Ohashi *et al.* (1981), who investigate a discrete time maintenance model, using the dynamic programming technique to find an optimal replacement policy. Mine *et al.* (1974) consider a continuous time maintenance model under Markovian deterioration and include maintenance times in their analysis. The stationary state probabilities are used in order to find an optimal maintenance policy. The aim of this chapter is to analyse a continuous time maintenance policy for a more general situation than that in Mine *et al.*, where in fact a special form of phase type distribution is chosen to model the deterioration. Without the assumption of Markovian deterioration, the role of the stationary state probabilities is played by the stationary joint pdf of the ages of the components. This latter approach is suggested and analysed by Bansard *et al.* (1969). Bansard suggested a trigger-off policy, an opportunistic continuous time replacement policy and used the stationary joint pdf of the ages of the units to compute the expected number of replacements of the units per unit time in the long run. This technique is also used by Berg (1978) who extended the work of Bansard and analysed a two-unit series system. Berg considers two replacement policies: an opportunistic failure replacement policy (OFRP) and an opportunistic age replacement policy (OARP). Under the OFRP, at any failure epoch of either of the two units, the unfailed unit is replaced too if its age exceeds a predetermined control limit L. On the other hand, under the OARP a unit is replaced at failure *or* when its age reaches a predetermined critical age S. At planned replacements as well as at failure epochs, the unfailed unit is replaced too if its age exceeds the control limit L. The OARP is the continuous time analogue of an (n,N)-policy, a well-known replacement policy

in Markov decision theory (Van der Duyn-Schouten *et al.*, 1990).
For the OFRP as well as the OARP, Berg deduces an implicit expression for the stationary joint pdf and shows that explicit expressions can be obtained for the particular choice of Erlang distributions for the life times of the individual units. As the system considered by Berg is a two-unit series system, either both units are operating or one unit is in repair and the other is waiting. Since Berg assumes that simultaneous failures do not occur, the state 'waiting' is used to denote that a unit is waiting to restart operating instead of being queued for repair. The system's one-step transition diagram is given in figure 6.1.

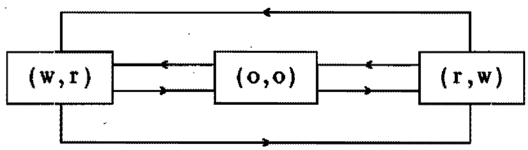


Fig. 6.1: One-step transition diagram under the OFRP and OARP

A second assumption made by Berg is that repairs are instantaneous. Hence, not only the deterioration of a unit during a repair of the other one is neglected, but also the effect of restarting an unfailed unit. Implicitly it is supposed that restarting a unit does not affect the residual life time of a unit.
In the case of a parallel system, neglecting repair times means that a unit can no longer fail during a repair of the other unit. Since this approach would ignore a fundamental aspect of a parallel system, the repair times are involved in the analysis here. On the other hand, the importance of incorporating repair times in the model should not be exaggerated. It is plausible that the effects of neglecting repair times are small when the duration of a repair is relatively small compared to the length of a life time of a component. It seems without doubt that Berg's approach (which considerably simplifies the analysis) gives a good approximation in the latter case. The effect of repair time durations will be subject of study in section 6.8, where numerical examples illustrate the OFRP.

Including repair times in the analysis gives rise to some slight modifications in the OFRP and OARP: obviously, the failure of a single unit provides a natural replacement opportunity for the unfailed unit, but only after repair completion of the failed one. Hence, the opportunistic replacement procedure to follow is:

i. Under the OFRP, after repair completion of a failed unit, an operating unit is replaced too if its age exceeds the control limit L.

ii. Under the OARP the OFRP is extended as follows. If *both* units are working, a unit is replaced if its operating time reaches the predetermined critical operating time S. On the other hand, if after repair completion of a unit, the operating time of the other unit exceeds the control limit L (or the critical age S), the latter unit is replaced as well (as under the OFRP).Clearly, the OFRP is obtained by setting $S=\infty$.

The one-step transition diagram under both the OFRP and OARP is given in figure 6.2, where, as the system under consideration is a parallel system, in state 'w' a unit is waiting for repair.

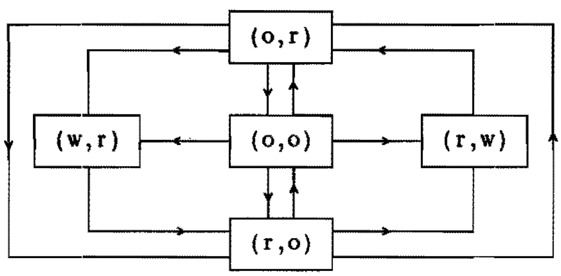


Fig. 6.2: One-step transition diagram under the OFRP and OARP

Compared to the previous chapters, a further generalisation is made here with respect to the probability distributions, since both life *and* repair times are supposed to be generally distributed. It appears that the use of discrete supplementary variables and phase type distributions does not offer any analytical advantages over the use of arbitrary distribution functions. Hence, starting from the model in section 1.4, the life time of component C_i (i=1,2) is assumed to be distributed with arbitrary (but continuous) cdf $F_i(.)$.

The pdf is denoted by $f_i(.)$ and the hazard rate by $h_i(.)$, *i.e.*

$$h_i(.) = \frac{f_i(.)}{\bar{F}_i(.)}.$$

As in chapter 5, it is assumed that the common cause effect is external and can be modelled by an exponential distribution with constant hazard rate h_3. Hence, $F_3(t)=1-\exp(-h_3t)$, $f_3(t)=h_3\exp(-h_3t)$ and $h_3(t)=h_3$.
Further, the repair time of component C_i $(i=1,2)$ has (continuous) cdf $G_i(.)$, pdf $g_i(.)$ and hazard rate $r_i(.)$. Obviously the joint survival function $\bar{F}(.,.)$ of the life times of two new units is

$$\bar{F}(x_1,x_2) = \bar{F}_1(x_1)\ \bar{F}_2(x_2)\ \bar{F}_3(\max(x_1,x_2)).$$

Note that in the general case of arbitrary distribution functions $F_i(.)$, it is easily verified that the lack-of-memory properties formulated in (3.2), (3.3) and (4.8) are lost.
A key role in the analysis is played by the stationary joint pdf of the process $\{X(t),t\geq 0\}$ which describes the operating, repair and waiting time (for repair) of both units at time t. In section 6.2 an integral equation is derived for the stationary joint pdf of the state description process under the OFRP. In section 6.3 some long run operating characteristics are computed and a cost function is constructed in order to determine the optimal choice of the policy parameter L. The optimal value of L minimises the expected cost per unit time in the long run. A similar analysis is performed in section 6.4 for the OARP. In order to reduce the (down) costs, an extension of the number of repair facilities is considered in section 6.5. The situation considered by Berg is object of study in section 6.6, where repairs are assumed to be instantaneous. Finally, the numerical evaluation of the stationary joint pdf is investigated in section 6.7 and in 6.8 numerical examples illustrate the techniques.

6.2 The stationary joint pdf under the OFRP

As described before, in the opportunistic failure replacement policy (OFRP) at any failure epoch of either of the two units, the failed unit is repaired and the unfailed unit is replaced if, after repair completion, the age of the unfailed unit exceeds a predetermined control limit. For simplicity, repairs and replacements are assumed to be identically distributed. Since the units are not necessarily identical, the control limits are supposed to be unit dependent: the control limit of unit i (i=1,2) is denoted by L_i. To find the stationary joint pdf, a state description process $\{X(t),t\geq 0\}$ is constructed which denotes not only the state of the system at time t, but also the sojourn time in the particular state. Therefore, let

$$X(t)=\{X_1(t),X_2(t),Y_1(t),Y_2(t),W_1(t),W_2(t)\},$$

where $X_i(t)$ represents the operating time of unit i at time t if unit i is operating and $X_i(t)=0$ otherwise. Analogously, $Y_i(t)$ is the repair time if the unit is in repair and $W_i(t)$ the waiting time if the unit is waiting for repair at time t. Note that at any time t, unit i is in one of the states 'o', 'r' or 'w'. Hence, at any time t, for each unit (at least) two quantities of the triple $X_i(t)$, $Y_i(t)$, $W_i(t)$ are zero and the remaining one is nonnegative. Let $p(x_1,x_2,y_1,y_2,w_1,w_2,t)$, denote the joint probability density function of the process $X(t)$. Then, at any time t, probability mass is concentrated in five planes, according to the five states in figure 6.2, *viz.* the (x_1,x_2), (x_1,y_2), (x_2,y_1), (y_1,w_2) and (y_2,w_1) plane. The Kolmogorov forward equations in these planes yield, for $\Delta\downarrow 0$,

$$p(\underline{x},t)=p(x_1-\Delta,x_2-\Delta,0,0,0,0,t-\Delta)\ (1-h_1(x_1-\Delta)\Delta)\ (1-h_2(x_2-\Delta)\Delta)\ (1-h_3\Delta)\ +\ o(\Delta),$$

$$\underline{x}\in D_{1,t}=\{(x_1,x_2,0,0,0,0)\,|\,0\leq x_1,x_2\leq t,\max(0,x_1-L_1)\leq x_2\leq \min(t,x_1+L_2)\}. \qquad (6.1)$$

The domain $D_{1,t}$ is obtained from $X_1(t)<X_2(t)+L_1$, $X_2(t)<X_1(t)+L_2$ and $0\leq X_i(t)\leq t$, i=1,2. Figure 6.3 shows $D_{1,t}$ in the case $t>\max(L_1,L_2)$.

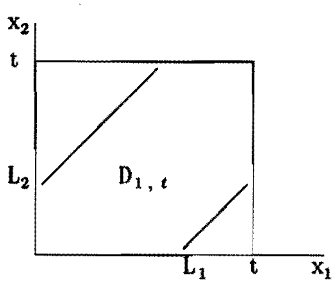


Fig. 6.3: $D_{1,t}$, $t>\max(L_1,L_2)$

Further,

$$p(\underline{x},t)=p(x_1-\Delta,0,0,y_2-\Delta,0,0,t-\Delta)\ (1-h_1(x_1-\Delta)\Delta)\ (1-r_2(y_2-\Delta)\Delta)\ (1-h_3\Delta)\ +\ o(\Delta),$$

$$\underline{x}\in D_{2,t}=\{(x_1,0,0,y_2,0,0)\,|\,0\le y_2\le x_1\le t\} \tag{6.2}$$

$$p(\underline{x},t)=p(0,x_2-\Delta,y_1-\Delta,0,0,0,t-\Delta)\ (1-h_2(x_2-\Delta)\Delta)\ (1-r_1(y_1-\Delta)\Delta)\ (1-h_3\Delta)\ +\ o(\Delta),$$

$$\underline{x}\in D_{3,t}=\{(0,x_2,y_1,0,0,0)\,|\,0\le y_1\le x_2\le t\} \tag{6.3}$$

$$p(\underline{x},t)=p(0,0,y_1-\Delta,0,0,w_2-\Delta,t-\Delta)\ (1-r_1(y_1-\Delta)\Delta)\ +\ o(\Delta),$$

$$\underline{x}\in D_{4,t}=\{(0,0,y_1,0,0,w_2)\,|\,0\le w_2\le y_1\le t\} \tag{6.4}$$

$$p(\underline{x},t)=p(0,0,0,y_2-\Delta,w_1-\Delta,0,t-\Delta)\ (1-r_2(y_2-\Delta)\Delta)\ +\ o(\Delta),$$

$$\underline{x}\in D_{5,t}=\{(0,0,0,y_2,w_1,0)\,|\,0\le w_1\le y_2\le t\} \tag{6.5}$$

$$p(\underline{x},t)=0,\ \text{elsewhere.} \tag{6.6}$$

Defining $\dot{p}(\underline{x},t)\ =\ dp/dt$, it follows that

$$\dot{p}(\underline{x},t)\ =\ \frac{\partial p}{\partial x_1}+\frac{\partial p}{\partial x_2}+\frac{\partial p}{\partial y_1}+\frac{\partial p}{\partial y_2}+\frac{\partial p}{\partial w_1}+\frac{\partial p}{\partial w_2}+\frac{\partial p}{\partial t},$$

since $dx_i/dt=1$, $dy_i/dt=1$ and $dw_i/dt=1$, because the stochastic processes $X_i(t)$, $Y_i(t)$ and $W_i(t)$ $(i=1,2)$ increase linearly with slope 1.

The following set of partial differential equations is obtained from (6.1)–(6.6):

$$\dot{p}(\underline{x},t)\ =\ \begin{cases} -\{h_1(x_1)+h_2(x_2)+h_3\}\ p(\underline{x},t), & \underline{x}\in D_{1,t} \\ -\{h_1(x_1)+r_2(y_2)+h_3\}\ p(\underline{x},t), & \underline{x}\in D_{2,t} \\ -\{h_2(x_2)+r_1(y_1)+h_3\}\ p(\underline{x},t), & \underline{x}\in D_{3,t} \\ -r_1(y_1)\ p(\underline{x},t), & \underline{x}\in D_{4,t} \\ -r_2(y_2)\ p(\underline{x},t), & \underline{x}\in D_{5,t} \end{cases}$$

Setting

$$p(\underline{x})\ =\ \lim_{t\to\infty}\ p(\underline{x},t),$$

$$\dot{p}(\underline{x})\ =\ \lim_{t\to\infty}\ \dot{p}(\underline{x},t),$$

$$D_i\ =\ \lim_{t\to\infty}\ D_{i,t},$$

yields (assumed that the above limits exist)

$$\dot{p}(\underline{x}) = \begin{cases} -\{h_1(x_1)+h_2(x_2)+h_3\}\ p(\underline{x}), & \underline{x}\in D_1 \\ -\{h_1(x_1)+r_2(y_2)+h_3\}\ p(\underline{x}), & \underline{x}\in D_2 \\ -\{h_2(x_2)+r_1(y_1)+h_3\}\ p(\underline{x}), & \underline{x}\in D_3 \\ -r_1(y_1)\ p(\underline{x}), & \underline{x}\in D_4 \\ -r_2(y_2)\ p(\underline{x}), & \underline{x}\in D_5 \end{cases} \tag{6.7}$$

As in Berg (1978), the general solution of the partial differential equations (6.7) is given by

$$p(\underline{x}) = \begin{cases} \bar{F}_1(x_1)\ \bar{F}_2(x_2)\ \bar{F}_3((x_1+x_2)/2)\ H_1(x_1-x_2), & \underline{x}\in D_1 \\ \bar{F}_1(x_1)\ \bar{G}_2(y_2)\ \bar{F}_3(x_1)\ H_2(x_1-y_2), & \underline{x}\in D_2 \\ \bar{G}_1(y_1)\ \bar{F}_2(x_2)\ \bar{F}_3(x_2)\ H_3(x_2-y_1), & \underline{x}\in D_3 \\ \bar{G}_1(y_1)\ H_4(y_1-w_2), & \underline{x}\in D_4 \\ \bar{G}_2(y_2)\ H_5(y_2-w_1), & \underline{x}\in D_5 \end{cases} \tag{6.8}$$

where the functions $H_i(.)$ $(i=1,..,5)$ are determined by the boundary conditions.

Further, the process $\mathbf{X}(t)$ jumps back to the origin when a common cause failure occurs, *i.e.* when a transition from state (o,o) to (w,r) or (r,w) is made in figure 6.2. Thus, probability mass is concentrated on the paths

$$D'_{4,t}=\{(0,0,y_1,0,0,w_2)\,|\,0\le y_1=w_2\le t\}\subset D_{4,t}$$

and

$$D'_{5,t}=\{(0,0,0,y_2,w_1,0)\,|\,0\le y_2=w_1\le t\}\subset D_{5,t}.$$

Moreover, it is clear that the process $\mathbf{X}(t)$ also jumps back to the origin when a transition is made from (r,w) or (r,o) to (o,r) and when a transition is made from (w,r) or (o,r) to (r,o). Hence, there is also probability mass concentrated on

$$D'_{2,t}=\{(x_1,0,0,y_2,0,0)\,|\,0\le y_2=x_1\le t\}\subset D_{2,t}$$

and

$$D'_{3,t}=\{(0,x_2,y_1,0,0,0)\,|\,0\le y_1=x_2\le t\}\subset D_{3,t}.$$

Let $q_i(\underline{x},t)$ be the probability mass on $D'_{i,t}$ and define for i=2,3,4,5

$$q_i(\underline{x}) = \lim_{t\to\infty} q_i(\underline{x},t),$$

$$\dot{q}_i(\underline{x},t) = \frac{\partial q_i}{\partial x} + \frac{\partial q_i}{\partial t},$$

$$\dot{q}_i(\underline{x}) = \lim_{t\to\infty} \dot{q}_i(\underline{x},t)$$

and

$$D'_i = \lim_{t\to\infty} D'_{i,t}.$$

Next, take $\underline{x}\in D'_{2,t}$, say $\underline{x}=(x,0,0,x,0,0)$. Then the Kolmogorov forward equations yield

$$q_2(\underline{x},t) = q_2(x-\Delta,0,0,x-\Delta,0,0,t-\Delta)\ (1-h_1(x-\Delta)\Delta)\ (1-r_2(x-\Delta)\Delta)\ (1-h_3\Delta) + o(\Delta).$$

Hence

$$\dot{q}_2(\underline{x},t) = -\{h_1(x)+r_2(x)+h_3\}\ q_2(\underline{x},t), \quad \underline{x}\in D'_{2,t}$$

and

$$\dot{q}_2(\underline{x}) = -\{h_1(x)+r_2(x)+h_3\}\ q_2(\underline{x}), \quad \underline{x}\in D'_2. \tag{6.9}$$

The general solution of (6.9) is

$$q_2(\underline{x}) = \bar{F}_1(x)\ \bar{G}_2(x)\ \bar{F}_3(x)\ H_2, \tag{6.10}$$

where H_2 is a positive constant.

Analogously, let $\underline{x}=(0,x,x,0,0,0)\in D'_3$, then

$$q_3(\underline{x}) = \bar{F}_2(x)\ \bar{G}_1(x)\ \bar{F}_3(x)\ H_3. \tag{6.11}$$

Further, for $\underline{x}=(0,0,x,0,0,x)\in D'_{4,t}$,

$$q_4(\underline{x},t)=q_4(0,0,x-\Delta,0,0,x-\Delta,t-\Delta)\ (1-r_1(x-\Delta)\Delta) + o(\Delta),$$

thus

$$\dot{q}_4(\underline{x},t) = -r_1(x)\ q_4(\underline{x},t), \quad \underline{x}\in D'_{4,t}$$

and

$$\dot{q}_4(\underline{x}) = -r_1(x)\ q_4(\underline{x}), \quad \underline{x}\in D_4^1. \tag{6.12}$$

The general solution of (6.12) is

$$q_4(\underline{x}) = \bar{G}_1(x)\ H_4, \tag{6.13}$$

where H_4 is a positive constant.

Similarly, for $\underline{x}=(0,0,0,x,x,0)\in D_5^1$,

$$q_5(\underline{x}) = \bar{G}_2(x)\ H_5. \tag{6.14}$$

Expressions for the functions $H_i(.)$ are found by considering the boundary conditions, *i.e.* the probability density along the x_i and y_i axes. First the probability density along the x_1-axis is considered. From figure 6.2 it is clear that the x_1-axis is crossed whenever a transition from (o,o) to (o,r) or from (o,r) to (o,o) occurs. Both transitions are represented symbolically in figure 6.4 which shows the (x_1,x_2,y_2)-space.

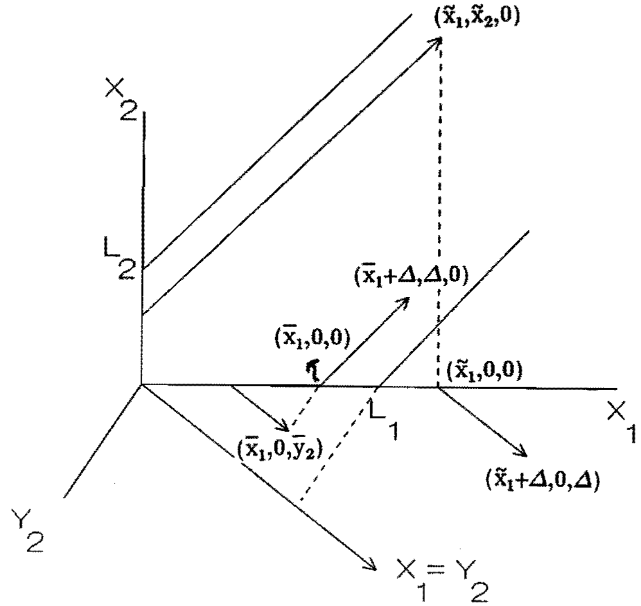


Fig. 6.4: Transitions from (o,o) to (o,r) and vice versa

Observing transitions from (o,r) to (o,o) gives,

$$F_1(x_1)\ F_3(x_1/2)\ H_1(x_1) = \int_0^{x_1} p(x_1,y_2)\ r_2(y_2)\ dy_2,\quad 0<x_1<L_1. \tag{6.15}$$

For convenience, only the non-zero arguments of p(.) are represented: *e.g.* $p(x_1,0,0,y_2,0,0)$ is abbreviated as $p(x_1,y_2)$.

Equation (6.15) is explained as follows. Consider a system in state (o,r). Then the x_1-axis is crossed at $(\bar{x}_1,0,0)$, $0<\bar{x}_1<L_1$, if the repair of unit 2 is completed after (say) $\bar{y}_2$ units of time, $\bar{y}_2\leq\bar{x}_1$, and unit 1 has been operating for $\bar{x}_1$ units of time at the moment of repair completion. This explains the right hand side. Alternatively, if unit 1 has operated for $\bar{x}_1+\Delta$ and unit 2 for Δ units of time, it is clear that a repair of unit 2 has been completed Δ units of time ago, *i.e.* the x_1-axis has been crossed at $(\bar{x}_1,0,0)$. So, by setting $x_2=0$ in the first equation of (6.8) the left hand side of (6.15) is obtained. Substituting $p(x_1,y_2)$ from (6.8) and dividing both sides by $F_1(x_1)\ F_3(x_1/2)$ yields

$$H_1(x) = F_3(x/2) \int_0^{x^-} g_2(y)\ H_2(x-y)\ dy + F_3(x/2)\ g_2(x)\ H_2,\quad 0<x<L_1. \tag{6.16}$$

Secondly, observing transitions from (o,o) to (o,r) gives, for $x_1>0$

$$F_1(x_1)\ F_3(x_1)\ H_2(x_1) = \int_{\substack{\max(0,x_1-L_1)\\ x_1\neq x_2}}^{x_1+L_2} p(x_1,x_2)\ h_2(x_2)\ dx_2. \tag{6.17}$$

The right hand side of equation (6.17) is obtained by examining transitions from (o,o) to (o,r). Starting in (o,o) the x_1-axis is crossed at $(\tilde{x}_1,0,0)$, $\tilde{x}_1>0$, if unit 2 fails after it has been operating for (say) $\tilde{x}_2$ units of time, $\max(0,\tilde{x}_1-L_1)\leq\tilde{x}_2\leq\tilde{x}_1+L_2$, and at the moment of failure unit 1 has operated for a time period of length $\tilde{x}_1$. On the other hand, if unit 1 has operated for $\tilde{x}_1+\Delta$ and unit 2 has been in repair for Δ units of time, it is clear that unit 2 failed Δ units of time ago, and so the x_1-axis has been crossed at $(\tilde{x}_1,0,0)$. Now, by setting $y_2=0$ in the second equation of (6.8), the left hand side of (6.17) follows. Note that the right hand side of (6.17) has the restriction $x_1\neq x_2$, since the function $H_1(.)$, which appears in $p(x_1,x_2)$, is not necessarily continuous at the origin.

After substitution of $p(x_1,x_2)$ it is found that

$$H_2(x) = \int_{\substack{\max(0,x-L_1)\\ x\neq y}}^{x+L_2} f_2(y)\ \bar{F}_3((y-x)/2)\ H_1(x-y)\ dy, \quad x>0, \tag{6.18}$$

where throughout this chapter $\bar{F}_3(x)=\exp(-h_3x)$, for *all* $x\in\mathbb{R}$.
Analogously, by interchanging the roles of unit 1 and unit 2

$$H_1(-x) = \bar{F}_3(x/2) \int_0^{x^-} g_1(y)\ H_3(x-y)\ dy + \bar{F}_3(x/2)\ g_1(x)\ H_3, \quad 0<x<L_2 \tag{6.19}$$

and

$$H_3(x) = \int_{\substack{\max(0,x-L_2)\\ x\neq y}}^{x+L_1} f_1(y)\ \bar{F}_3((y-x)/2)\ H_1(y-x)\ dy, \quad x>0. \tag{6.20}$$

Similarly, the y_1-axis is crossed if a transition occurs from state (r,o) to (r,w) and obviously, for $y_1>0$

$$\bar{G}_1(y_1)\ H_4(y_1) = \int_{y_1}^{\infty} p(x_2,y_1)\ \left(h_2(x_2)+h_3\right)\ dx_2.$$

Substituting $p(x_2,y_1)$ and dividing both sides by $\bar{G}_1(y_1)$ yields $(x>0)$,

$$H_4(x) = \int_{x^+}^{\infty} \left(f_2(y)\bar{F}_3(y)+\bar{F}_2(y)f_3(y)\right)\ H_3(y-x)\ dy + \left(f_2(x)\bar{F}_3(x)+\bar{F}_2(x)f_3(x)\right)\ H_3. \tag{6.21}$$

By symmetry

$$\bar{G}_2(y_2)\ H_5(y_2) = \int_{y_2}^{\infty} p(x_1,y_2)\ \left(h_1(x_1)+h_3\right)\ dx_1$$

and hence $(x>0)$

$$H_5(x) = \int_{x^+}^{\infty} \left(f_1(y)\bar{F}_3(y)+\bar{F}_1(y)f_3(y)\right)\ H_2(y-x)\ dy + \left(f_1(x)\bar{F}_3(x)+\bar{F}_1(x)f_3(x)\right)\ H_2. \tag{6.22}$$

For known constants H_i, equations (6.16) and (6.18)–(6.22) determine the functions $H_i(.)$, $i=1,..,5$, and hence the stationary joint pdf given by (6.8). In order to find the constants H_i, note that $q_i(\underline{0})=H_i$. Expressions for $q_i(\underline{0})$ are found by investigating the possibilities by which the process X(t) can jump to the origin and continue on the path D_i^1. Remember that the process X(t) restarts in the origin whenever a transition occurs from

i. (o,o) to (w,r)

ii. (o,o) to (r,w)

iii. (r,w) or (r,o) to (o,r)

iv. (w,r) or (o,r) to (r,o).

The cases *i* and *ii* concern the occurrence of a common cause failure, while unit 1 (2) is the first to be repaired, *i.e.* the process X(t) jumps back to the origin and continues on D_4^1 (D_5^1). Hence,

$$H_4 = p_1 \int_0^{\infty} \int_{\substack{\max(0,x_1-L_1) \\ x_1 \neq x_2}}^{x_1+L_2} p(x_1,x_2)\ h_3\ dx_2\ dx_1 \tag{6.23}$$

and

$$H_5 = p_2 \int_0^{\infty} \int_{\substack{\max(0,x_1-L_1) \\ x_1 \neq x_2}}^{x_1+L_2} p(x_1,x_2)\ h_3\ dx_2\ dx_1. \tag{6.24}$$

Case *iii* incorporates the repair completion of unit 1, while unit 2

i. is waiting for repair or

ii. has been operating for more than L_2 units of time at the moment of repair completion of unit 1.

In both cases the X(t) process jumps to the origin and continues on D_2^1.
Thus,

$$H_2 = \int_0^{\infty} \int_0^{y_1} p(y_1,w_2)\ r_1(y_1)\ dw_2\ dy_1 + \int_{L_2}^{\infty} \int_0^{x_2} p(x_2,y_1)\ r_1(y_1)\ dy_1\ dx_2. \tag{6.25}$$

and analogously

$$H_3 = \int_0^{\infty} \int_0^{y_2} p(y_2,w_1)\ r_2(y_2)\ dw_1\ dy_2 + \int_{L_1}^{\infty} \int_0^{x_1} p(x_1,y_2)\ r_2(y_2)\ dy_2\ dx_1. \tag{6.26}$$

The functions $H_i(.)$ and the constants H_i are determined up to a constant by the equations (6.16)–(6.26). This constant is obtained from the normalisation equation

$$\sum_{i=1}^{5} \iint_{D_i - D_i'} p(x,y)\, dx\, dy + \sum_{i=2}^{5} \int_{D_i'} q_i(x)\, dx = 1. \tag{6.27}$$

In section 6.7 it will be shown how to compute the stationary joint pdf (6.8) from the set of equations (6.16)–(6.26).

6.3 Operating characteristics under the OFRP

Once the stationary joint pdf (6.8) has been obtained, the system's long run operating characteristics under the OFRP can be computed. The main interest here is in the expected costs and hence in the expected number of failures and replacements (per unit of time) in the long run. Apart from these quantities, expressions are derived for the limiting availability and unavailability. Define

N_i^O = The expected number of failures per unit of time of unit i, while the other unit is operating at the moment the failure of unit i occurs,

N_i^R = The expected number of failures per unit of time of unit i, while the other unit is in repair at the moment the failure of unit i occurs,

N^{CC} = The expected number of times per unit of time that both units fail simultaneously by a common cause failure,

N^S = The expected number of failures per unit of time on system level,

N_i^{PR} = The expected number of preventive replacements per unit time of unit i,

N_i^{MA} = The expected number of maintenance actions (defined by a repair or replacement) per unit of time of unit i.

Expressions for the above measures are easily obtained in terms of the stationary joint pdf. To start with, it can be seen that N_2^O equals the expected number of transitions per unit time from (o,o) to (o,r) in the stationary case. In other words: N_2^O equals the number of times the x_1–axis is crossed, due to a transition from (o,o) to (o,r).

Hence, by (6.17)

$$N_2^O = \int_0^\infty \int_{\max(0,x_1-L_1)}^{x_1+L_2} p(x_1,x_2)\ h_2(x_2)\ dx_2\ dx_1 \tag{6.28}$$

and an expression for N_1^O is obtained by interchanging the roles of the units. Similarly, N_2^R equals the number of the times per unit time the y_1–axis is crossed. From the derivation of (6.21) it is clear that

$$N_2^R = \int_0^\infty \int_{y_1}^\infty p(x_2,y_1)\ \left[h_2(x_2)+h_3\right]\ dx_2\ dy_1. \tag{6.29}$$

Further, N_1^{PR} is the expected number of times state (o,r) is left by a repair completion of unit 2, while the operating time of unit 1 exceeds L_1 at the moment of repair completion.
As a result (*cf.* (6.15)),

$$N_1^{PR} = \int_{L_1}^\infty \int_0^{x_1} p(x_1,y_2)\ r_2(y_2)\ dy_2\ dx_1. \tag{6.30}$$

Notice that N_1^{PR} equals the second term of (6.26), as H_3 contains, among others, the expected number of transitions per unit time from (o,r) to (r,o).
Considering the number of simultaneous failures, it is clear that N^{CC} is the expected number of transitions per unit time from (o,o) to (w,r) or (r,w), *i.e.* the number of times per unit time the process $\mathbf{X}(t)$ jumps to the origin and continues on the path D_4' or D_5'. Thus, by (6.23) and (6.24),

$$N^{CC} = \frac{H_4}{p_1} = \frac{H_5}{p_2}. \tag{6.31}$$

With respect to N^S and N_i^{MA} it follows immediately,

$$N^S = N_1^R + N_2^R + N^{CC} \tag{6.32}$$

and

$$N_i^{MA} = N_i^O + N_i^R + N_i^{PR} + N^{CC}. \tag{6.33}$$

The total number of maintenance actions performed by the repair facility per unit time, denoted by N^{MA}, is

$$N^{MA} = \sum_{i=1}^{2} (N_i^O + N_i^R + N_i^{PR} + N^{CC}). \tag{6.34}$$

However, the number of visits to the system (per unit of time), made by the repair facility, is less than N^{MA}, since the number of maintenance actions executed per visit can be arbitrarily large. To illustrate this, suppose that the system follows the path (o,o)–(o,r)–(r,o)–(r,w)–(o,r)–(o,o), which results in three maintenance actions during one visit of the repair facility. Being interested in the number of maintenance actions performed per visit, define N^{RF} as

N^{RF} = The expected number of times per unit of time that the system is visited by the repair facility.

Then N^{RF} equals the expected number of transitions per unit of time from (o,r) or (r,o) to (o,o). Thus,

$$N^{RF} = \int_0^{L_1}\int_0^{x_1} p(x_1,y_2)\ r_2(y_2)\ dy_2\ dx_1 + \int_0^{L_2}\int_0^{x_2} p(x_2,y_1)\ r_1(y_1)\ dy_1\ dx_2 \tag{6.35}$$

and hence the average number of maintenance actions per visit is N^{MA}/N^{RF}. On the other hand, it is easily seen that N^{RF} equals the expected number of transitions per unit of time from state (o,o) to the set {(o,r),(r,o),(r,w), (w,r)}. Consequently,

$$N^{RF} = N_1^O + N_2^O + N^{CC}.$$

The next point of interest is the number of 'single' and 'multiple' maintenance actions. Let N_i^{SMA} denote the expected number of single maintenance actions per unit of time on unit i, *i.e.*

N_i^{SMA} = The expected number of times per unit of time that the repair facility visits the system and performs only a single maintenance action on unit i during the visit.

The expected number of multiple maintenance actions per unit of time (N^{MMA}) is defined as

N^{MMA} = The expected number of times per unit of time that the repair facility visits the system and performs more than one maintenance action during the visit.

Obviously, N_2^{SMA} equals the expected number of transitions from (o,r) to (o,o), where the transition has to be made from the space D_2–D_2' to D_1. Hence,

$$N_2^{SMA} = \int_0^{L_1}\int_0^{x_1^-} p(x_1,y_2)\ r_2(y_2)\ dy_2\ dx_1$$

and similarly N_1^{SMA} follows from transitions from D_3–D_3' to D_1.
Secondly, a multiple maintenance action ends by a transition from D_2' or D_3' to D_1, so

$$N^{MMA} = \int_0^{L_1} q_2(x)\ r_2(x)\ dx + \int_0^{L_2} q_3(x)\ r_1(x)\ dx.$$

Of course,

$$N^{RF} = N_1^{SMA} + N_2^{SMA} + N^{MMA}.$$

Further, by definition, the limiting availability A of the system is given by the probability the system is in (o,o), (o,r) or (r,o) in the long run. Hence,

$$A = \iint_{D_1} p(x_1,x_2)\ dx_2\ dx_1 + \iint_{D_2} p(x_1,y_2)\ dy_2\ dx_1 + \iint_{D_3} p(x_2,y_1)\ dy_1\ dx_2,$$

and the limiting unavailability U, U=1–A, is

$$U = \iint_{D_4} p(y_1,w_2)\ dw_2\ dy_1 + \iint_{D_5} p(y_2,w_1)\ dw_1\ dy_2.$$

In fact, U is composed of two contributions:

i. U_{fr}, the unavailability caused by a failure of one unit during the repair of the other unit,

ii. U_{cc}, the unavailability caused by common cause failures.

Clearly

$$U_{fr} = \iint\limits_{D_4 - D_4'} p(y_1,w_2)\ dw_2\ dy_1 + \iint\limits_{D_5 - D_5'} p(y_2,w_1)\ dw_1\ dy_2.$$

and

$$U_{cc} = \int_0^\infty q_4(x)\ dx + \int_0^\infty q_5(x)\ dx$$

Apart from giving insight into the system's failure behaviour, the above operating characteristics are useful in determining the optimal replacement policy. A pair of control limits (L_1^*,L_2^*) is said to be optimal if it minimises the expected costs per unit time in the long run. The cost function to be minimised here is an extension of the cost structure suggested by Berg (1978). It is a variant which contains specific terms in order to make it suitable for the analysis of a parallel system. Two types of costs are incorporated in the model: costs due to a repair or preventive replacement of a unit and the costs of staying in a particular system state. With respect to the replacement costs it is supposed that the fixed costs corresponding to a call for the repair facility are k_0. Further, let k_i^R be the repair or replacement costs of unit i (i=1,2) and k_i^{PR} the costs of a preventive replacement. Then the costs of repairs and preventive replacements per unit time in the long run, denoted by $C_{rpr}(L_1,L_2)$, are

$$C_{rpr}(L_1,L_2) = N^{RF}\ k_0 + \sum_{i=1}^{2} \left[(N_i^O + N_i^R + N^{CC})\ k_i^R + N_i^{PR}\ k_i^{PR} \right]. \tag{6.36}$$

The state dependent costs are expressed in the marginal 'running' costs per unit time. Define

$c_i^O(x)$:the marginal operating costs of unit i, when operating x units of time,

$c^R(x)$:the marginal repair costs at repair time x,

$c^D(x)$:the marginal down costs of the system, being down for x units of time.

From the above definitions, the expected running costs per unit time in the long run, denoted by $C_{run}(L_1,L_2)$, are

$$C_{run}(L_1,L_2) = \iint_{D_1} p(x_1,x_2)\ (c_1^O(x_1)+c_2^O(x_2))\ dx_2dx_1 \tag{6.37}$$

$$+\iint_{D_2} p(x_1,y_2)\ (c_1^O(x_1)+c^R(y_2))\ dy_2dx_1 + \iint_{D_3} p(x_2,y_1)\ (c_2^O(x_2)+c^R(y_1))\ dy_1dx_2$$

$$+\iint_{D_4} p(y_1,w_2)\ (c^R(y_1)+c^D(w_2))\ dw_2dy_1 + \iint_{D_5} p(y_2,w_1)\ (c^R(y_2)+c^D(w_1))\ dw_1dy_2.$$

Hence the total expected costs per unit time in the long run, $C(L_1,L_2)$, are

$$C(L_1,L_2) = C_{run}(L_1,L_2) + C_{rpr}(L_1,L_2). \tag{6.38}$$

In principle, the optimal pair of control limits (L_1^*,L_2^*) can be found minimising $C(L_1,L_2)$ over L_1 and L_2. However, without any assumptions on the marginal cost functions or the life and repair time distributions, it is not clear how the costs (6.38) behave as a function of L_1 and L_2. Obviously, further research is needed to find out if assumptions on the structure of the marginal cost functions can be exploited in the minimisation of (6.38). Without further knowledge of the cost function, only general optimisation procedures, which use function values or a numerical approximation of the partial derivatives of (6.38) with respect to L_1 and L_2, can be used to obtain the optimum (L_1^*,L_2^*). For a survey of optimisation methods and techniques, the reader is referred to Scales (1985).

6.4 The opportunistic age replacement policy

Under the opportunistic *failure* replacement policy (OFRP), at any failure epoch of either of the two units, the unfailed unit is replaced as well, if at repair completion of the failed unit, the age of the unfailed unit exceeds a predetermined control limit. Under the opportunistic *age* replacement policy (OARP), a unit is also replaced if its age reaches a predetermined critical age. Thus under the OARP it is not necessary to wait for a failure of either of the units to replace a unit. Let S_i denote the critical age of unit i (i=1,2). Then, compared to the OFRP, the only modification made is that an additional restriction prevents the operating time x_i of unit i from exceeding S_i in the (x_1,x_2) plane.

As a result, under the OARP

$$D_1=\{(x_1,x_2,0,0,0,0)\,|\,0\le x_1\le \min(S_1,S_2+L_1),\ \max(0,x_1-L_1)\le x_2\le \min(S_2,x_1+L_2)\}.$$

To simplify the expressions for the region of integration of the various integrals, the following assumptions are made:

i. It is supposed that $S_i>L_i$, since in the special case where $S_i\le L_i$ only age replacement of unit i occurs. Besides, if both $S_1\le L_1$ and $S_2\le L_2$, the OARP reduces to an individual age replacement policy for the two units as in Berg (1978).

ii. It is assumed that $S_1\le S_2+L_1$ and $S_2\le S_1+L_2$. Hence the cases in figure 6.5 and 6.6, where the control limit S_1, respectively S_2, is never reached, are excluded from the analysis.

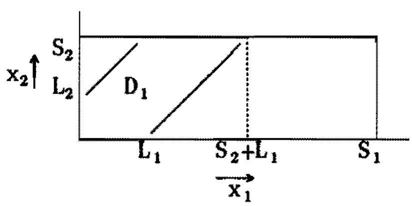


Fig. 6.5: $S_1>S_2+L_1$

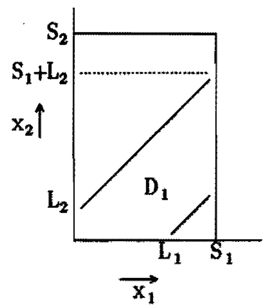


Fig. 6.6: $S_2>S_1+L_2$

Under the above assumptions,

$$D_1=\{(x_1,x_2,0,0,0,0)\,|\,0\le x_1\le S_1,\ \max(0,x_1-L_1)\le x_2\le \min(S_2,x_1+L_2)\}.$$

The upper bound S_1 for x_1 in the (x_1,x_2) plane leads to a second modification compared to the OFRP in the (x_1,y_2) plane: as illustrated in figure 6.7, $y_2\le x_1<y_2+S_1$. By symmetry $y_1\le x_2<y_1+S_2$ in the (x_2,y_1) plane.

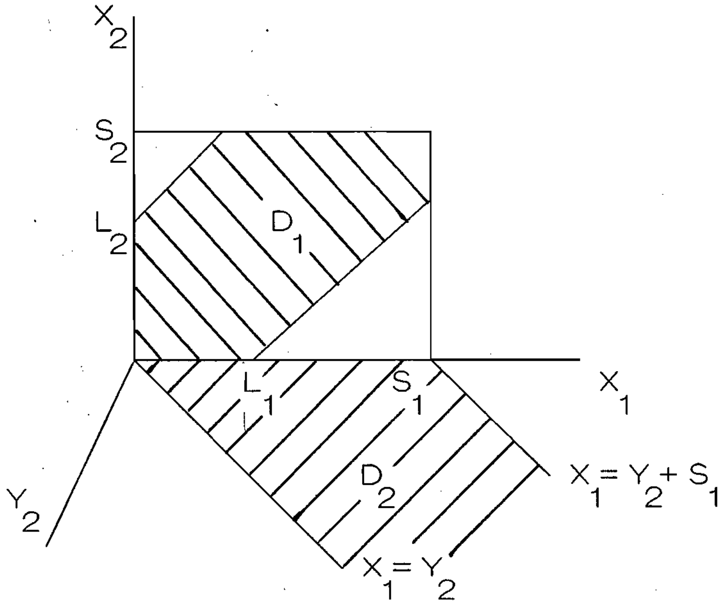


Fig. 6.7: D_1 and D_2 under the OARP

Thus, under the OARP

$$D_1=\{(x_1,x_2,0,0,0,0)\,|\,0\le x_1\le S_1,\ \max(0,x_1-L_1)\le x_2\le \min(S_2,x_1+L_2)\},$$

$$D_2=\{(x_1,0,0,y_2,0,0)\,|\,0\le y_1\le x_2<y_1+S_2\},$$

$$D_3=\{(0,x_2,y_1,0,0,0)\,|\,0\le y_2\le x_1<y_2+S_1\},$$

$$D_4=\{(0,0,y_1,0,0,w_2)\,|\,0\le w_2\le y_1\},$$

$$D_5=\{(0,0,0,y_2,w_1,0)\,|\,0\le w_1\le y_2\}.$$

The general solution of the partial differential equations, obtained by the Kolmogorov forward equations, is again given by (6.8), with the above modifications for the domains D_i. Expressions for the probability density along the paths D_i^j, defined as in section 6.2, are as in (6.10)–(6.14). The functions $H_i(.)$ and the constants H_i are obtained by a similar reasoning as under the OFRP. To start with, the functions $H_1(.)$, $H_4(.)$ and $H_5(.)$ are, apart from some slight modifications in the integration area, as in section 6.2.

Clearly,

$$H_1(x) = \bar{F}_3(x/2) \int_0^{x^-} g_2(y)\ H_2(x-y)\ dy + \bar{F}_3(x/2)\ g_2(x)\ H_2, \quad 0<x<L_1,$$

$$H_1(-x) = \bar{F}_3(x/2) \int_0^{x^-} g_1(y)\ H_3(x-y)\ dy + \bar{F}_3(x/2)\ g_1(x)\ H_3, \quad 0<x<L_2,$$

and for $x>0$,

$$H_4(x) = \int_{x^+}^{x+S_2} \big(f_2(y)\bar{F}_3(y)+\bar{F}_2(y)f_3(y)\big)\ H_3(y-x)\ dy + \big(f_2(x)\bar{F}_3(x)+\bar{F}_2(x)f_3(x)\big)\ H_3,$$

$$H_5(x) = \int_{x^+}^{x+S_1} \big(f_1(y)\bar{F}_3(y)+\bar{F}_1(y)f_3(y)\big)\ H_2(y-x)\ dy + \big(f_1(x)\bar{F}_3(x)+\bar{F}_1(x)f_3(x)\big)\ H_2.$$

An expression for $H_2(.)$ is derived by analysing transitions from (o,o) to (o,r). Under the OARP these transitions occur by a failure of unit 2 *or* when unit 2 has reached the critical age S_2. Hence, *cf.* (6.17), for $0<x_1<S_1$

$$\bar{F}_1(x_1)\ \bar{F}_3(x_1)\ H_2(x_1) = \int_{\substack{\max(0,x_1-L_1) \\ x_1 \neq x_2}}^{\min(x_1+L_2,S_2)^-} p(x_1,x_2)\ h_2(x_2)\ dx_2 + p(x_1,S_2).$$

Note that $p(x_1,S_2)=0$ if $x_1 \notin (S_2-L_2,S_1)$.

Substitution of $p(x_1,x_2)$ yields, for $0<x<S_1$

$$H_2(x) = \int_{\substack{\max(0,x-L_1) \\ x \neq y}}^{\min(x+L_2,S_2^-)} f_2(y)\ \bar{F}_3((y-x)/2)\ H_1(x-y)\ dy + \bar{F}_2(S_2)\ \bar{F}_3((S_2-x)/2)\ H_1(x-S_2).$$

By symmetry, for $0<x<S_2$

$$H_3(x) = \int_{\substack{\max(0,x-L_2) \\ x \neq y}}^{\min(x+L_1,S_1^-)} f_1(y)\ \bar{F}_3((y-x)/2)\ H_1(y-x)\ dy + \bar{F}_1(S_1)\ \bar{F}_3((S_1-x)/2)\ H_1(S_1-x).$$

With respect to the constants H_i it is easily verified that

$$H_2 = \int_0^{\infty}\int_0^{y_1} p(y_1,w_2)\ r_1(y_1)\ dw_2\ dy_1 + \int_{L_2}^{\infty}\int_{\max(0,x_2-S_2)}^{x_2} p(x_2,y_1)\ r_1(y_1)\ dy_1\ dx_2,$$

$$H_3 = \int_0^{\infty}\int_0^{y_2} p(y_2,w_1)\ r_2(y_2)\ dw_1\ dy_2 + \int_{L_1}^{\infty}\int_{\max(0,x_1-S_1)}^{x_1} p(x_1,y_2)\ r_2(y_2)\ dy_2\ dx_1,$$

$$H_4 = p_1 \iint_{D_1} p(x_1,x_2)\ h_3\ dx_2\ dx_1$$

and

$$H_5 = p_2 \iint_{D_1} p(x_1,x_2)\ h_3\ dx_2\ dx_1.$$

As in section 6.2, the above equations determine the functions $H_i(.)$ and the constants H_i up to a constant, which is obtained from the normalisation equation (6.27), with the appropriate domains D_i and D_i^*. The derivation of the operating characteristics follows an identical reasoning. Apart from some modifications in the limits of integration, the operating characteristics are as in section 6.3. However, under the OARP one more operating characteristic is defined by the expected number of age replacements per unit time of unit i, denoted by N_i^{AR}.

If a transition from state (o,r) to (r,o), while the operating time of unit 1 exceeds S_1 at the moment that the repair of unit 2 is completed, is called a *preventive* replacement (because $S_1>L_1$) and *not* an *age* replacement, then it is easily seen that

$$N_1^{AR} = \int_{S_1-L_1}^{S_2} p(S_1,x_2)\ dx_2.$$

A similar expression follows for N_2^{AR} by interchanging the roles of the units 1 and 2. As a consequence, an additional term occurs in the expression for N_i^{MA} and instead of (6.33) and (6.34),

$$N_i^{MA} = N_i^{O} + N_i^{R} + N_i^{PR} + N_i^{AR} + N_{cc}$$

and

$$N^{MA} = \sum_{i=1}^{2} (N_i^{O} + N_i^{R} + N_i^{PR} + N_i^{AR} + N^{CC}).$$

Finally, let k_0 be fixed costs, k_i^R the repair costs of a unit of type i and k_i^{REPL} the costs of an age replacement or a preventive replacement, then the total expected costs per unit time in the long run, denoted by $C(L_1,L_2,S_1,S_2)$ are

$$C(L_1,L_2,S_1,S_2) = C_{run}(L_1,L_2,S_1,S_2) + C_{rr}(L_1,L_2,S_1,S_2) \tag{6.39}$$

where the running costs are given by (6.37) (with the appropriate domains D_i) and the repair and replacement costs $C_{rr}(.)$ are given by

$$C_{rr}(L_1,L_2,S_1,S_2) = N^{RM}\ k_0 + \sum_{i=1}^{2} \left[(N_i^R + N_i^O + N_{cc})\ k_i^R + (N_i^{PR} + N_i^{AR})\ k_i^{REPL}\right].$$

With respect to the minimisation of the total expected costs per unit time, the same conclusion holds as for (6.38): further research is needed to investigate whether assumptions on the structure of the cost functions, or assumptions on the life and repair time distributions, can be exploited to find the optimal tuple of control limits $(L_1^*,L_2^*,S_1^*,S_2^*)$.

6.5 Two repair facilities

Assume that when the system is down, an emergency call is made and a second (emergency) repair facility becomes active. It is clear that the presence of a second repair man reduces the system down time and hence the down costs. However, the number of calls for a repair man (N^{RF}) increases as do the fixed costs in (6.36). To decide whether the availability of a second repair facility is economically worthwhile, the case with two repair crews is analysed for the OFRP. (The modifications that have to be made under the OARP are obvious.) To simplify the analysis, the assumption is made that both repair facilities are identical, since in the other case the repair time distributions may depend upon the type of repair facility. This dependence would necessitate an extension of the state space, since the identity of the repair facility is needed whenever a repair or replacement is executed. However, the assumption of identical repair crews simplifies the analysis: compared to figure 6.2, the states (r,w) and (w,r) join together into a new state (r,r) and the number of states reduces to four. The one-step transition diagram is shown in figure 6.8.

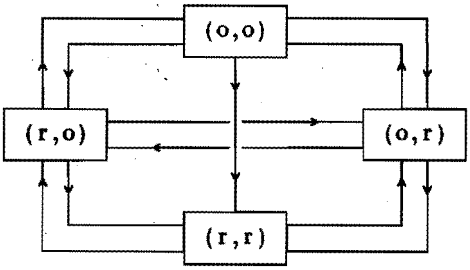


Fig. 6.8: One-step transition diagram, two repair facilities

As a failed unit is switched to the repair state instantaneously, the state description process $\{X(t),t\geq 0\}$ is given by

$$X(t)=\{X_1(t),X_2(t),Y_1(t),Y_2(t)\},$$

where $X_i(t)$ and $Y_i(t)$ are as in section 6.2. Again, let $p(\underline{x})$, $\underline{x}=(x_1,x_2,y_1,y_2)$, be the stationary joint pdf of $X(t)$, then probability mass is concentrated in four planes, viz. the (x_1,x_2), (x_1,y_2), (x_2,y_1) and (y_1,y_2) plane. Moreover, probability mass is concentrated on the paths $x_1=y_2$, $y_1=x_2$ and $y_1=y_2$. Let

$$D_1=\{(x_1,x_2,0,0)\,|\,x_1\geq 0,\ \max(0,x_1-L_1)\leq x_2\leq x_1+L_2\}$$

$$D_2=\{(x_1,0,0,y_2)\,|\,x_1,y_2\geq 0\},\quad D_2'=\{(x,0,0,x)\,|\,x\geq 0\}\subset D_2$$

$$D_3=\{(0,x_2,y_1,0)\,|\,x_2,y_1\geq 0\},\quad D_3'=\{(0,x,x,0)\,|\,x\geq 0\}\subset D_3$$

$$D_4=\{(0,0,y_1,y_2)\,|\,y_1,y_2\geq 0\},\quad D_4'=\{(0,0,x,x)\,|\,x\geq 0\}\subset D_4$$

then it is easily verified

$$p(\underline{x}) = \begin{cases} \bar{F}_1(x_1)\ \bar{F}_2(x_2)\ \bar{F}_3((x_1+x_2)/2)\ H_1(x_1-x_2) & \underline{x}\in D_1 \\ \bar{F}_1(x_1)\ \bar{G}_2(y_2)\ \bar{F}_3(x_1)\ H_2(x_1-y_2) & \underline{x}\in D_2 \\ \bar{G}_1(y_1)\ \bar{F}_2(x_2)\ \bar{F}_3(x_2)\ H_3(x_2-y_1) & \underline{x}\in D_3 \\ \bar{G}_1(y_1)\ \bar{G}_2(y_2)\ H_4(y_1-y_2) & \underline{x}\in D_4 \end{cases}$$

and along the paths D_i',

$$q_2(\underline{x}) = \bar{F}_1(x)\ \bar{G}_2(x)\ \bar{F}_3(x)\ H_2 \qquad \underline{x}\in D_2'$$
$$q_3(\underline{x}) = \bar{G}_1(x)\ \bar{F}_2(x)\ \bar{F}_3(x)\ H_3 \qquad \underline{x}\in D_3'$$
$$q_4(\underline{x}) = \bar{G}_1(x)\ \bar{G}_2(x)\ H_4 \qquad \underline{x}\in D_4'$$

As before, the integral equations for the functions $H_i(.)$ and the constants H_i are obtained by a careful analysis of the various transition possibilities. From figure 6.8 there are 11 types of transitions, *viz.* transitions from

i. (o,r) to (o,o)
ii. (r,o) to (o,o)
iii. (o,o) to (o,r)
iv. (r,r) to (o,r)
v. (o,o) to (r,o)
vi. (r,r) to (r,o)
vii. (r,o) to (r,r)
viii. (o,r) to (r,r)
ix. (r,o) to (o,r)
x. (o,r) to (r,o)
xi. (o,o) to (r,r)

Examination of the above possibilities yields the following set of integral equations for the functions $H_i(.)$ and the constants H_i.

i. $$\bar{F}_1(x_1)\ \bar{F}_3(x_1/2)\ H_1(x_1) = \int_0^\infty p(x_1,y_2)\ r_2(y_2)\ dy_2$$

ii. $$\bar{F}_2(x_2)\ \bar{F}_3(x_2/2)\ H_1(-x_2) = \int_0^\infty p(x_2,y_1)\ r_1(y_1)\ dy_1$$

iii. $$\bar{F}_1(x_1)\ \bar{F}_3(x_1)\ H_2(x_1) = \int_{\max(0,\,x_1-L_1)}^{x_1+L_2} p(x_1,x_2)\ h_2(x_2)\ dx_2$$

iv. $$\bar{G}_2(y_2)\ H_2(-y_2) = \int_0^\infty p(y_1,y_2)\ r_1(y_1)\ dy_1$$

v. $$\bar{F}_2(x_2)\ \bar{F}_3(x_2)\ H_3(x_2) = \int_{\max(0,x_2-L_2)}^{x_2+L_1} p(x_1,x_2)\ h_1(x_1)\ dx_1$$

vi. $$\bar{G}_1(y_1)\ H_3(-y_1) = \int_0^{\infty} p(y_1,y_2)\ r_2(y_2)\ dy_2$$

vii. $$\bar{G}_1(y_1)\ H_4(y_1) = \int_0^{\infty} p(x_2,y_1)\ \left[h_2(x_2)+h_3\right]\ dx_2$$

viii. $$\bar{G}_2(y_2)\ H_4(-y_2) = \int_0^{\infty} p(x_1,y_2)\ \left[h_1(x_1)+h_3\right]\ dx_1$$

ix. $$H_2 = \int_{L_2}^{\infty}\int_0^{\infty} p(x_2,y_1)\ r_1(y_1)\ dy_1\ dx_2$$

x. $$H_3 = \int_{L_1}^{\infty}\int_0^{\infty} p(x_1,y_2)\ r_2(y_2)\ dy_2\ dx_1$$

xi. $$H_4 = \int_0^{\infty}\int_{\max(0,x_1-L_1)}^{x_1+L_2} p(x_1,x_2)\ h_3\ dx_2\ dx_1$$

The above equations, together with the normalisation equation, determine the functions $H_i(.)$ and the constants H_i. With respect to the operating characteristics it is clear that the expressions (6.28)–(6.30) and (6.32)–(6.33) still hold, since the *failure* mechanism has not changed: only the *repair* mechanism is different from the original situation. Further, it is easily seen that, instead of (6.32),

$$N^{CC} = H_4.$$

Moreover, N^{RF} equals the number of transitions per unit time from (o,r) or (r,o) to (o,o) *plus* the number of transitions from (r,r) to (o,r) or (r,o), and hence, *cf.* (6.35),

$$N^{RF} = \int_0^{L_1}\int_0^{\infty} p(x_1,y_2)\ r_2(y_2)\ dy_2\ dx_1 + \int_0^{L_2}\int_0^{\infty} p(x_2,y_1)\ r_1(y_1)\ dy_1\ dx_2$$

$$+ \int_0^{\infty}\int_0^{\infty} p(y_1,y_2)\ \left[r_1(y_1)+r_2(y_2)\right]\ dy_2\ dy_1.$$

For the limiting availability and unavailability, it is quite trivial that

$$A = \iint_{D_1} p(x_1,x_2)\ dx_2\ dx_1 + \iint_{D_2} p(x_1,y_2)\ dy_2\ dx_1 + \iint_{D_3} p(x_2,y_1)\ dy_1\ dx_2,$$

and

$$U = \iint_{D_4} p(y_1,y_2)\ dy_2\ dy_1.$$

As in section 6.3 the unavailability can be decomposed as

$$U = U_{fr} + U_{cc},$$

where

$$U_{fr} = \iint_{D_4-D_4'} p(y_1,y_2)\ dy_2\ dy_1$$

and

$$U_{cc} = \int_0^{\infty} q_4(x)\ dx.$$

Other operating characteristics can be computed when a decision rule is known for the assignment of maintenance actions to the repair facilities. An example of a decision rule is:

i. if the system is in state (o,o) and unit i (i=1,2) fails, the unit is assigned to repair facility i with probability p_i and to repair facility 3–i with probability $1-p_i$,

ii. if both units fail simultaneously:
- unit 1 is assigned to repair facility 1 with probability q_1 and to repair facility 2 with probability $1-q_1$,
- unit 2 is assigned to the repair facility which remains vacant.

Given the decision rule, interesting questions concern:

i. The expected number of times repair facility i is called for,

ii. The expected number of times a maintenance action on unit i is executed by repair facility j (i,j=1,2).

iii The expected number of maintenance actions done by repair facility i during one call,

iv. Etc.

However, attention will not be focussed here on these questions.

The cost criterion can be used to compare the situations with either one or two repair facilities. As in section 6.3, the total expected costs per unit of time in the long run are given by

$$C(L_1,L_2) = C_{run}(L_1,L_2) + C_{rpr}(L_1,L_2),$$

where the repair and replacement costs $C_{rpr}(L_1,L_2)$ are as in (6.36) and the running costs $C_{run}(L_1,L_2)$ are

$$\begin{aligned} C_{run}(L_1,L_2) = & \iint_{D_1} p(x_1,x_2)\ (c_1^O(x_1)+c_2^O(x_2))\ dx_2dx_1 \\ & +\iint_{D_2} p(x_1,y_2)\ (c_1^O(x_1)+c^R(y_2))\ dy_2dx_1 \\ & +\iint_{D_3} p(x_2,y_1)\ (c_2^O(x_2)+c^R(y_1))\ dy_1dx_2 \\ & +\iint_{D_4} p(y_1,y_2)\ [c^R(y_1)+c^R(y_2)]\ dy_2dy_1. \end{aligned}$$

Obviously, the availability of a second repair facility is worthwhile if the minimum expected costs per unit time in the long run are less in the case with two repair facilities than in the case with one repair facility.

6.6 Instantaneous repairs

As mentioned in section 6.1, it is plausible that ignoring repair times will give a good approximation of the situation with non-zero repair times when the repair time durations are relatively small compared to the life times. To compare the results of an analysis with and without repairs, the total expected costs per unit of time are used as a criterion in section 6.8, where numerical examples illustrate the techniques in both cases. The situations studied have relatively small and relatively large repair times. However, first it is shown how to perform the analysis when the repair times can be neglected.

Berg (1978) analyses a two-unit series system and assumes that repairs are instantaneous, *i.e.* deterministic with length zero. In that case, the states (r,w) and (w,r) in figure 6.1 are instantaneous. Moreover, the stochastic analysis of a two-unit *series* system is identical to the analysis of a two-unit *parallel* system with zero repair times, since the states (o,r), (r,o), (r,w) and (w,r) in figure 6.2 are then all instantaneous. The only modification that has to be made in Berg's analysis, to study a *dependent* parallel system, is to include the common cause effect. The following analysis summarises the results obtained by Berg, adjusted for a two-unit dependent parallel system.

The state description process $\{X(t),t\geq 0\}$ under consideration is

$$X(t)=\{X_1(t),X_2(t)\},$$

where $X_i(t)$ represents the operating time of unit i at time t. Let $p(x_1,x_2,t)$ denote the joint probability density function of the process $X(t)$. Then conditioning on time $t-\Delta$ yields, for $\Delta\downarrow 0$,

$$p(x_1,x_2,t) = p(x_1-\Delta,x_2-\Delta,t-\Delta)\ (1-h_1(x_1-\Delta)\Delta)\ (1-h_2(x_2-\Delta)\Delta)\ (1-h_3\Delta) + o(\Delta),$$

$$(x_1,x_2)\in D_t = \{(x_1,x_2)\ |\ 0\leq x_1,x_2\leq t,\ \max(0,x_1-L_1)\leq x_2\leq \min(t,x_1+L_2)\}. \tag{6.40}$$

By similar reasoning to that in section 6.2, the stationary joint pdf $p(x_1,x_2)$ is obtained from (6.40) as

$$p(x_1,x_2) = \bar{F}_1(x_1)\ \bar{F}_2(x_2)\ \bar{F}_3((x_1+x_2)/2)\ H(x_1-x_2), \tag{6.41}$$

$$(x_1,x_2)\in D = \{(x_1,x_2)\ |\ x_1,x_2\geq 0,\ \max(0,x_1-L_1)\leq x_2\leq x_1+L_2\}.$$

The function H(.) is determined by the boundary conditions.
Further, the process X(t) jumps back to the origin when a common cause failure or a preventive replacement occurs. Hence, probability mass is concentrated on the path

$$D_t' = \{(x,x) \mid 0 \le x \le t\} \subset D_t.$$

Let q(x,t) be the probability mass on D_t', then it is easily verified that

$$q(x,t) = q(x-\Delta,t-\Delta)\ (1-h_1(x-\Delta)\Delta)\ (1-h_2(x-\Delta)\Delta)\ (1-h_3\Delta) + o(\Delta), \quad \Delta\downarrow 0$$

and in the stationary case the probability mass q(x) on $D=\{(x,x) \mid x\ge 0\}$ is given by

$$q(x) = \bar{F}_1(x)\ \bar{F}_2(x)\ \bar{F}_3(x)\ H_0, \quad x>0 \tag{6.42}$$

where H_0 is a positive constant.

Observation of the boundary conditions yields, after some manipulations

$$H(x) = \int_{0}^{x^-} f_2(y)\ \bar{F}_3(y/2)\ H(x-y)\ dy + f_2(x)\ \bar{F}_3(x/2)\ H_0, \quad 0<x<L_1 \tag{6.43}$$

and

$$H(-x) = \int_{0}^{x^-} f_1(y)\ \bar{F}_3(y/2)\ H(y-x)\ dy + f_1(x)\ \bar{F}_3(x/2)\ H_0, \quad 0<x<L_2. \tag{6.44}$$

The function H(.) and the constant H_0 are determined up to a constant by the above equations. This constant is obtained from the normalisation equation

$$\iint_{D-D'} p(x_1,x_2)\ dx_2\ dx_1 + \int_{D'} q(x)\ dx = 1.$$

With respect to the operating characteristics, it is clear that $N_i^R=0$ (i=1,2), since repairs are instantaneous. Further, it is easily verified that

$$N_1^O = \int_{0}^{\infty} \int_{\max(0,x_1-L_1)}^{x_1+L_2} p(x_1,x_2)\ h_1(x_1)\ dx_2\ dx_1, \tag{6.45}$$

$$N_2^{PR} = \int_{L_2}^{\infty} \int_{x_2-L_2}^{x_2+L_1} p(x_1,x_2)\ h_1(x_1)\ dx_1\ dx_2, \tag{6.46}$$

$$N_1^{SMA} = \int_0^{L_2}\int_0^{x_2+L_1} p(x_1,x_2)\ h_1(x_1)\ dx_1\ dx_2 \tag{6.47}$$

and

$$N^{CC} = \int_0^{\infty} \int_{\max(0,x_1-L_1)}^{x_1+L_2} p(x_1,x_2)\ h_3\ dx_2\ dx_1. \tag{6.48}$$

Expressions for N_2^O, N_1^{PR} and N_2^{SMA} are found by interchanging the roles of the components C_1 and C_2. As in section 6.2, the expected number of visits per unit of time, made by the repair man to the system, satisfies

$$N^{RF} = N_1^{SMA} + N_2^{SMA} + N^{MMA} \tag{6.49}$$

and, since repairs are instantaneous,

$$N^{MMA} = N_1^{PR} + N_2^{PR} + N^{CC}. \tag{6.50}$$

Moreover,

$$N^{MMA} = H_0,$$

since N^{MMA} is the expected number of times per unit of time that the process $\{X(t),t\geq0\}$ crosses the origin. In other words: $N^{MMA} = q(0)$.
Further, it follows from (6.45)–(6.47) that the operating characteristics N_i^{SMA}, N_i^{PR} and N_i^O are related by

$$N_i^{SMA} + N_{3-i}^{PR} = N_i^O,\ i=1,2. \tag{6.51}$$

Substitution of (6.50) and (6.51) in (6.49) yields an alternative expression for N^{RF}, *viz.*

$$N^{RF} = N_1^O + N_2^O + N^{CC}.$$

Apart from the availability, the other operating characteristics are obtained as in section 6.2, with the substitution $N_i^R=0$. Formally, ignoring repair times yields a limiting availability of 100 %, *i.e.* $A=1$ and $U_{cc}=U_{fr}=0$. However, in section 6.8 it will be shown how an approximation of the limiting (un)availability can be obtained, when the model with instantaneous repairs is used as an approximation of reality. The given approximation will be used to estimate the running costs (6.37). Subsequently, the repair and replacement costs being defined as in (6.36), the total expected costs per unit of time are calculated in a couple of situations to investigate the effect of incorporating repair times in the analysis. However, first it will be shown in section 6.7 how to compute the stationary joint pdf numerically.

6.7 Numerical evaluation of the stationary joint pdf under the OFRP

Consider the situation described in section 6.2, with one repair facility and non-zero repair times. Then the stationary joint pdf is given by (6.8), with the provision that the probability density along the paths D_i' (i=2,..,5) is as in (6.10)–(6.14). The functions $H_i(.)$ (i=1,..,5) and the constants H_i (i=2,..,5) are determined up to a constant by the expressions (6.15)–(6.26). This constant, say C_{norm}, is obtained from the normalisation equation (6.27). In this section it is illustrated how to compute the stationary joint pdf (6.8) numerically from (6.15)–(6.27).

Considering the set of equations (6.15)–(6.27), an important observation concerns the role of the functions $H_2(x)$ $(x>L_1)$, $H_3(x)$ $(x>L_2)$, $H_4(x)$, $H_5(x)$ and the constants H_4 and H_5 in the numerical computation of the stationary joint pdf. The solution of these integral equations depends on finding the functions $H_1(x)$ $(-L_2<x<L_1)$, $H_2(x)$ $(0<x<L_1)$, $H_3(x)$ $(0<x<L_2)$ and the constants H_2 and H_3. Clearly, these functions completely determine $H_2(x)$ $(x>L_1)$, $H_3(x)$ $(x>L_2)$, $H_4(x)$, $H_5(x)$, H_4 and H_5. In fact, these latter functions and constants are only used to compute the normalisation constant C_{norm} by (6.27). Therefore, attention is concentrated on $H_1(x)$ $(-L_2<x<L_1)$, $H_2(x)$ $(0<x<L_1)$ and $H_3(x)$ $(0<x<L_2)$, which are determined by (6.16)–(6.20) for *known* constants H_2 and H_3.

Firstly, the case with identical units is considered, which simplifies the analysis by reasons of symmetry. It is shown how discretisation of the integral equations yields the stationary joint pdf (6.8). Secondly, the more complicated situation with non-identical units is considered and an iterative procedure is suggested for finding the above functions and constants and thereby the stationary joint pdf (6.8).

When the units are identical, the life time distributions are identical, as well as the repair time distributions and the control limits, *i.e.* $f_1(t)=f_2(t)$, $g_1(t)=g_2(t)$ $(t\geq 0)$ and $L_1=L_2$.

Further, by symmetry it follows that

$$\begin{aligned}
H_1(x) &= H_1(-x), \quad 0<x<L_1 \\
H_2(x) &= H_3(x), \quad x>0 \\
H_4(x) &= H_5(x), \quad x>0 \\
H_2 &= H_3, \\
H_4 &= H_5.
\end{aligned}$$

Hence, from (6.16), (6.18) and (6.21),

$$H_1(x) = F_3(x/2)\int_{0}^{x^-} g_2(y)\, H_2(x-y)\, dy + F_3(x/2)\, g_2(x)\, H_2, \quad 0<x<L_1 \tag{6.52}$$

$$H_2(x) = \int_{\substack{\max(0,x-L_1) \\ x\neq y}}^{x+L_1} f_2(y)\, F_3((y-x)/2)\, H_1(|x-y|)\, dy, \quad x>0, \tag{6.53}$$

$$H_4(x) = \int_{x^+}^{\infty} \left(f_2(y)F_3(y)+F_2(y)f_3(y)\right) H_2(y-x)\, dy + \left(f_2(x)F_3(x)+F_2(x)f_3(x)\right) H_2, \quad x>0. \tag{6.54}$$

Next, from (6.23), (6.25), (6.8), (6.10) and (6.13),

$$H_2 = \int_{0}^{\infty}\int_{0}^{y_1^-} g_1(y_1)\, H_4(y_1-w_2)\, dw_2\, dy_1 + \int_{0}^{\infty} g_1(x)\, H_4\, dx \tag{6.55}$$

$$+ \int_{L_2}^{\infty}\int_{0}^{x_2^-} g_1(y_1)\, F_2(x_2)\, F_3(x_2)\, H_2(x_2-y_1) + \int_{L_2}^{\infty} g_1(x)\, F_2(x)\, F_3(x)\, H_2\, dx$$

and

$$H_4 = 0.5\int_{0}^{\infty}\int_{\substack{\max(0,x_1-L_1) \\ x_1\neq x_2}}^{x_1+L_2} F_1(x_1)\, F_2(x_2)\, f_3((x_1+x_2)/2)\, H_1(|x_1-x_2|)\, dx_2\, dx_1, \tag{6.56}$$

since for identical units $p_1=p_2=0.5$ in (6.23) and (6.24).

Note that, for given H_2, the functions $H_1(x)$ and $H_2(x)$ $(x<L_1)$ can be obtained numerically by discretisation of (6.52) and (6.53) (appendix D). Subsequently, $H_2(x)$ $(x>L_1)$, $H_4(x)$, H_2, H_4 and the normalisation constant C_{norm} can be computed by numerical integration, using *e.g.* the trapezium rule or Simpson's rule. Truncation of the infinite integrals is discussed in appendix D. Finally, the stationary joint pdf follows from (6.8), (6.10) and (6.13).

Hence, the strategy is:

i. Initialise H_2, *e.g.* $H_2:=1$,

ii. Solve $H_1(x)$ $(x<L_1)$ and $H_2(x)$ $(x<L_1)$ by discretisation of (6.52), (6.53),

iii. Compute $H_2(x)$ $(x>L_1)$ and $H_4(x)$ by numerical integration of (6.53),(6.54),

iv. Compute H_2 and H_4 by numerical integration of (6.55) and (6.56),

v. Compute the normalisation constant C_{norm} from (6.27), using (6.8), (6.10) and (6.13),

vi. Update the functions $H_1(x)$, $H_2(x)$, $H_4(x)$ and the constants H_2, H_4, dividing them by C_{norm},

vii. Compute the stationary joint pdf by (6.8), (6.10) and (6.13).

For non-identical units, the constants H_2 *and* H_3 have to be initialised and the crucial point is that the above procedure will only give the right result when the ratio H_2/H_3 is known. However, although a formal proof is not presented here, numerical experiments have shown that the stationary joint pdf can be computed successfully by the following iterative version of the above algorithm. Let k denote iteration step k and use the suffix k to denote the approximation of the functions $H_i(.)$ and the constants H_i in step k. Then compute the stationary joint pdf (6.8) as follows:

i. $k:=1$, $H_2^{(k)}:=1$, $H_3^{(k)}:=1$,

ii. Compute $H_1^{(k)}(x)$ $(-L_2<x<L_1)$, $H_2^{(k)}(x)$ $(x<L_1)$ and $H_3^{(k)}(x)$ $(x<L_2)$ by discretisation of (6.16) and (6.18)–(6.20).

iii. Compute $H_2^{(k)}(x)$ $(x>L_1)$, $H_3^{(k)}(x)$ $(x>L_2)$, $H_4^{(k)}(x)$ $(x>0)$ and $H_5^{(k)}(x)$ $(x>0)$ from (6.18) and (6.20)–(6.22) by numerical integration,

iv. Compute the constants $H_4^{(k)}$ and $H_5^{(k)}$ by numerical integration of (6.23) and (6.24), using (6.8) and (6.10)–(6.14),

v. Compute the normalisation constant C_{norm} by (6.27), using (6.8) and (6.10)–(6.14). If $|1-C_{norm}|<\varepsilon$ (for fixed $\varepsilon>0$), then go to *viii*, else go to *vi*,

vi. Update $H_i^{(k)}(x)$ $(i=1,..,5)$ and $H_i^{(k)}$ $(i=1,..,4)$ by setting $k:=k+1$, $H_i^{(k)}(x):=H_i^{(k-1)}(x)/C_{norm}$, $H_i^{(k)}:=H_i^{(k-1)}/C_{norm}$,

vii. Re-calculate $H_2^{(k)}$ and $H_3^{(k)}$ by computing the right hand side of (6.25) and (6.26) and go to *ii*,

viii. Compute the stationary joint pdf, using (6.8) and (6.10)–(6.14).

The iteration method has been tested for identically exponentially distributed life, respectively repair times and an exponentially distributed common cause effect. In this case the functions $H_i(.)$ and the constants H_i can be obtained analytically and hence the results of the iteration procedure can be checked. Since only a few numerical experiments have been executed, further research is needed with respect to the convergence of the iteration method and the accuracy of the results. However, it appeared that:

i. the iteration method is not sensitive to the initial values of the constants H_2 and H_3,

ii. within 3 or 4 steps $|1-C_{norm}|<0.001$, for identically exponentially distributed life and repair times,

iii. the relative error in the functions $H_i(.)$ and the constants H_i was less than 2 %.

Hence, it seems that the above iterative procedure is a useful tool to obtain the stationary joint pdf when the units are non-identical.

6.8 Numerical examples

Consider a two-unit dependent parallel system with identical units and one repair facility. To investigate the effects of non-zero repair times, the OFRP is illustrated numerically in the following cases.

The life times of the components C_1 and C_2 are assumed to have a Weibull distribution with scale parameter α ($\alpha>0$) and a shape parameter β ($\beta>0$), denoted by $W(\alpha,\beta)$. Hence,

$$f_1(t) = \frac{\beta}{\alpha}\left[\frac{t}{\alpha}\right]^{\beta-1} \exp(-(t/\alpha)^{\beta}),\ t\geq 0$$

and $f_2(t)=f_1(t)$. The Weibull distribution is a popular distribution to model life times, since its hazard rate is

i. increasing if $\beta>1$,

ii. decreasing if $\beta<1$,

iii. constant if $\beta=1$.

The mean μ_L and variance σ_L^2 of $W(\alpha,\beta)$ distributed life times are

$$\mu_L = \alpha\ \Gamma(1+1/\beta)$$

and

$$\sigma_L^2 = \alpha^2 \cdot \{\Gamma(1+2/\beta)-\Gamma^2(1+1/\beta)\},$$

where $\Gamma(.)$ represents the gamma function.
The repair times are assumed to follow a lognormal distribution with scale parameter μ ($\mu \in \mathbb{R}$) and shape parameter σ ($\sigma>0$), denoted by $\mathcal{LN}(\mu,\sigma)$. Thus,

$$g_1(t) = \frac{1}{\sqrt{2\pi}\ \sigma t} \exp\left[-\tfrac{1}{2}\left[\frac{\ln(t)-\mu}{\sigma}\right]^2\right], \ t>0$$

and $g_2(t)=g_1(t)$. The popularity of the lognormal distribution for modelling repair times is due to its heavy tail. The mean μ_R and variance σ_R^2 of $\mathcal{LN}(\mu,\sigma)$ distributed repair times are

$$\mu_R = \exp(\mu+\sigma^2/2)$$

and

$$\sigma_R^2 = (\exp(\sigma^2)-1) \exp(2\mu+\sigma^2).$$

Further, the common cause effect is supposed to be exponentially distributed with parameter λ, *i.e.*

$$f_3(t) = \lambda \exp(-\lambda t), \ t \geq 0.$$

To compare the situations with non-zero and instantaneous repairs and to investigate the effect of non-zero repair times on the OFRP, the cost criterion is used and the total expected costs per unit of time (6.38) are calculated in the following cases.
The life times of the components C_1 and C_2 have a $\mathcal{W}(5,3)$ distribution (μ_L=4.46, σ_L^2=2.63), which means that they are approximately normally distributed. With respect to the repair times, four cases are considered, *viz.* repairs which have a

i. $\mathcal{LN}(-1.5,0.9)$ distribution (μ_R=0.33, σ_R^2=0.14),
ii. $\mathcal{LN}(-0.5,0.9)$ distribution (μ_R=0.91, σ_R^2=1.03),
iii. $\mathcal{LN}(0.5,0.9)$ distribution (μ_R=2.47, σ_R^2=7.63),
and
iv. instantaneous repairs.

Hence the expected repair time varies from relatively small to relatively large values. Further, the common cause effect is exponentially distributed with λ=0.10.

The following choices are made for the cost factors in (6.36) and (6.37). Since the ratio k_i^{PR}/k_i^R appears to influence strongly the optimal value of the control limit, three cases are considered, *viz.*

i. $k_i^{PR}=800$, $k_i^R=2000$, $i=1,2$,

ii. $k_i^{PR}=1200$, $k_i^R=2000$, $i=1,2$,

iii. $k_i^{PR}=1600$, $k_i^R=2000$, $i=1,2$.

Note that $k_i^{PR}/k_i^R<1$ in the above cases, since the failure of a unit generally causes damage on the system, in contrast to a preventive replacement.

The fixed costs k_0 in (6.36) are chosen relatively small ($k_0=10$), since they represent the call-out fee when the repair facility visits the system. The marginal operating, repair and down costs in (6.37) are taken $c_i^O(x)=50$, $c^R(x)=100$ and $c^D(x)=5000$ per unit of time ($i=1,2$, $x>0$). The marginal operating costs include small maintenance actions as an oil change, etc. The marginal repair costs consist of the salary of the repair man and the marginal down costs of the loss of production.

The figures 6.9–6.20 show plots of the expected repair and replacement costs per unit of time (6.36), the expected running costs per unit of time (6.37) and the total expected costs per unit of time (6.38), as a function of the control limit. Note that the control limits L_1 and L_2 are identical: $L_1=L_2=L$. The value of L varies from 1 to 6 (step size 0.2), *i.e.* from relatively small to relatively large, compared to the mean life time of the components C_1 and C_2. The cases considered are: $k_i^{PR}=800$ in fig. 6.9–6.12, $k_i^{PR}=1200$ in fig. 6.13–6.16 and the figures 6.17–6.20 represent the case with $k_i^{PR}=1600$ ($i=1,2$). For each value of k_i^{PR}, the effect of the repair time distribution is examined and the cost functions (6.36)–(6.38) are plotted for lognormally distributed repair times with shape parameter 0.9 and a scale parameter of respectively −1.5, −0.5 and 0.5 and for instantaneous repairs.

The last case needs some comments. Under instantaneous repairs, the analysis proceeds as in section 6.6 and the stationary joint pdf (6.41) is found by discretisation of the integral equations (6.43) and (6.44). However, of special interest are the questions if the results for instantaneous repairs can be used to approximate the optimum value L^* of the control limit and the total expected costs per unit of time in the case with non-zero repair times. Note that the answer on the latter question will be negative, since the marginal down costs are relatively high and the limiting availability is 100 % under instantaneous repairs. Therefore, when the model with instantaneous repairs is used to approximate the situation with non-zero repair times, the stationary availability is approximated by

$$\hat{A}(\mu_R) = 1 - \frac{N^{CC}\mu_R}{1 + N^{CC}\mu_R},$$

where μ_R is the mean repair time and N^{CC} is given by (6.48). Note that, as a result, the running costs $C_{run}(L,L)$ are independent of L, since the marginal operating and down costs are constant and the limiting availability $\hat{A}(\mu_R)$ does not depend on L. More specific, by (6.37)

$$C_{run}(L,L) = (1-\hat{A}(\mu_R))\ (c^R(0)+c^D(0)) + \hat{A}(\mu_R)\ (c_1^O(0)+c_2^O(0)).$$

This explains the three cases with constant running costs in the figures 6.12b, 6.16b and 6.20b. On the other hand, the repair and replacement costs $C_{rpr}(L,L)$, defined by (6.36) and plotted in the figures 6.12a, 6.16a and 6.20a, *depend* on L but *not* on the repair time distribution. Note that the total expected costs per unit of time are not plotted for instantaneous repairs, since the optimum value of the control limit is determined by the costs $C_{rpr}(L,L)$.

The figures 6.9–6.20 show that:

i. $C_{run}(L,L)$ is almost constant as a function of L, because the limiting availability (and hence the unavailability) is hardly influenced by the value of L. Consequently, L^* is mainly determined by the repair and replacement costs $C_{rpr}(L,L)$ (when the costs $C_{rpr}(L,L)$ and $C_{run}(L,L)$ are of the same order of magnitude).

ii. When $C_{run}(L,L) \gg C_{rpr}(L,L)$, the total expected costs per unit of time are almost constant.

iii. The optimum L^* heavily depends on the ratio k_i^{PR}/k_i^R. Clearly, L^* increases when the ratio k_i^{PR}/k_i^R increases. It seems that the optimum replacement policy by and large can be formulated as:

1. Never perform a preventive replacement if k_i^{PR}/k_i^R is relatively large, *i.e.* when $k_i^{PR}/k_i^R \approx 1$,
2. Always perform a preventive replacement if k_i^{PR}/k_i^R is relatively small, *i.e.* $k_i^{PR}/k_i^R \approx 0.5$ (or less),
3. If $0.5 < k_i^{PR}/k_i^R < 1$, then perform a numerical analysis to determine the exact value of L^*.

Life times : W(5,3)
Repair times : LN(−1.5,0.9)
Costs : k_i^R=2000, k_i^{PR}=800

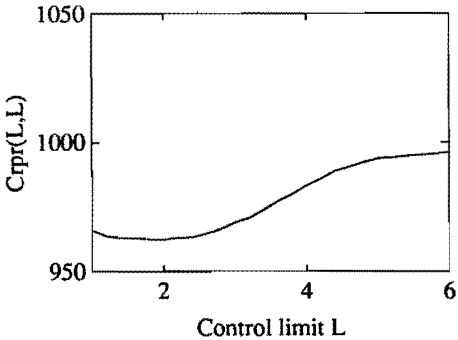


Fig. 6.9a: C_{rpr}(L,L)

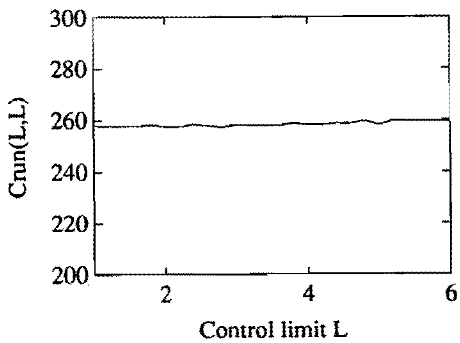


Fig. 6.9b: C_{run}(L,L)

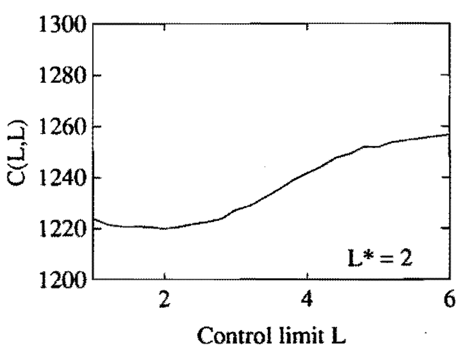


Fig. 6.9c: C(L,L)

Life times : W(5,3)
Repair times : LN(−0.5,0.9)
Costs : k_i^R=2000, k_i^{PR}=800

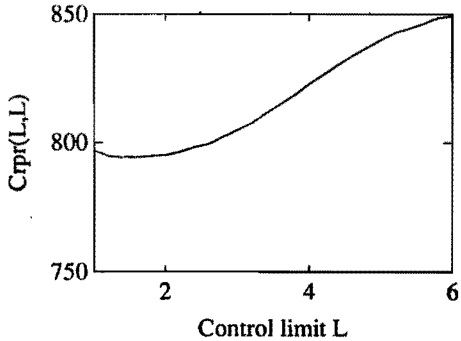


Fig. 6.10a: C_{rpr}(L,L)

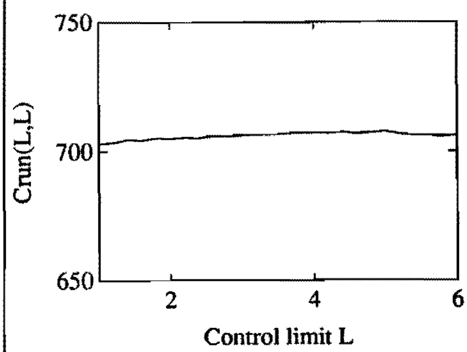


Fig. 6.10b: C_{run}(L,L)

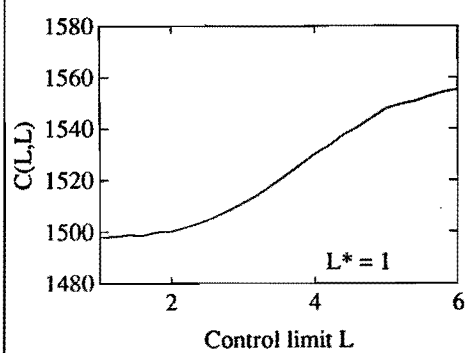


Fig. 6.10c: C(L,L)

Life times : W(5,3)

Repair times : LN(0.5,0.9)

Costs : k_i^R=2000, k_i^{PR}=800

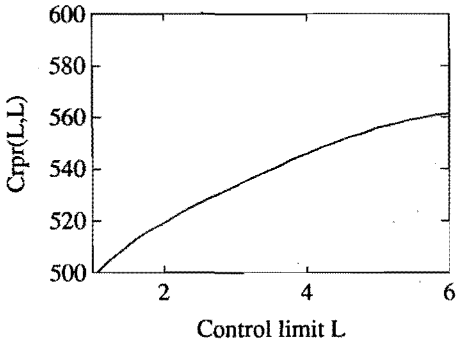


Fig. 6.11a: C_{rpr}(L,L)

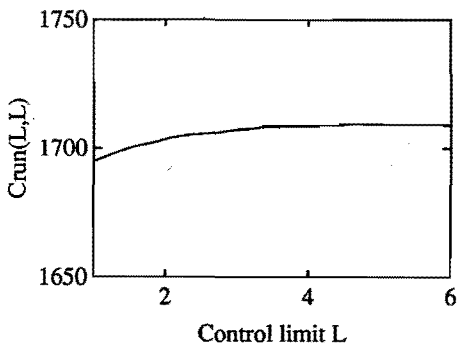


Fig. 6.11b: C_{run}(L,L)

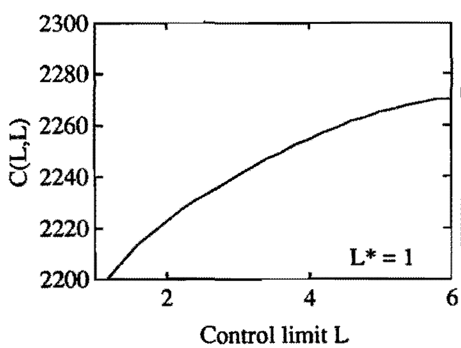


Fig. 6.11c: C(L,L)

Life times : W(5,3)

Instantaneous repairs

Costs : k_i^R=2000, k_i^{PR}=800

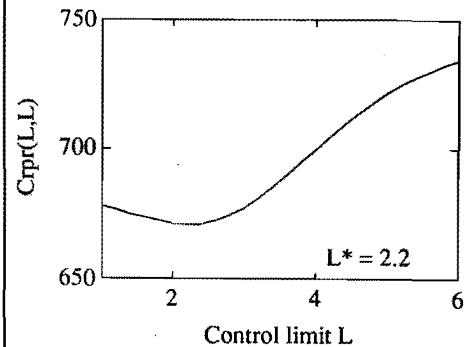


Fig. 6.12a: C_{rpr}(L,L)

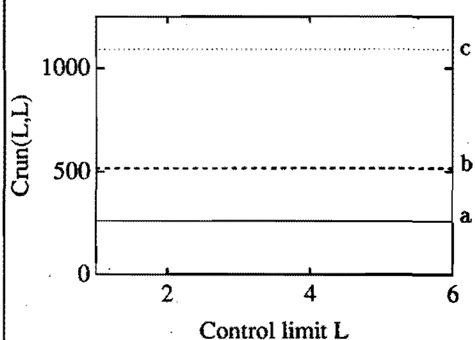


Fig. 6.12b: C_{run}(L,L)

Cases a, b, c: availability approximated by respectively $\hat{A}(0.33)$, $\hat{A}(0.91)$ and $\hat{A}(2.47)$.

Life times : W(5,3)
Repair times : LN(−1.5,0.9)
Costs : $k_i^R=2000$, $k_i^{PR}=1200$

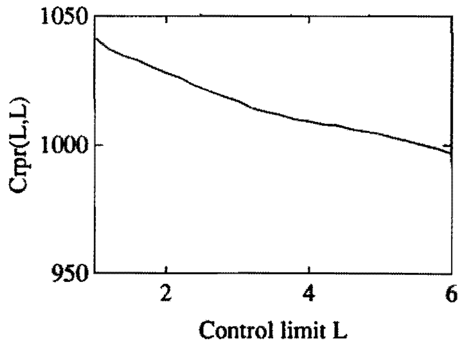


Fig. 6.13a: $C_{rpr}(L,L)$

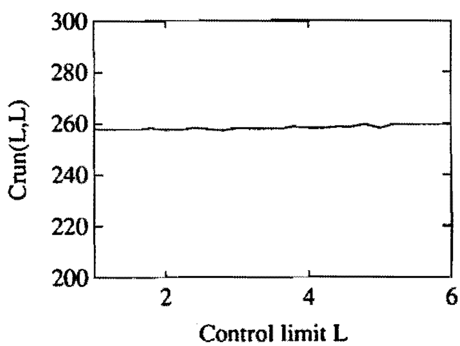


Fig. 6.13b: $C_{run}(L,L)$

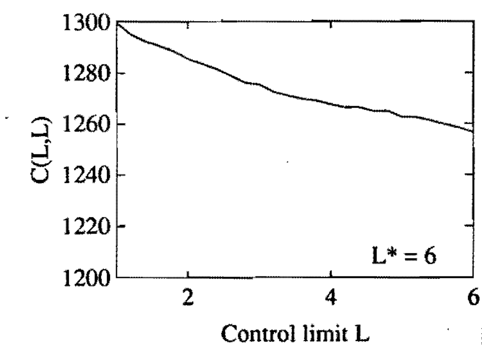


Fig. 6.13c: C(L,L)

Life times : W(5,3)
Repair times : LN(−0.5,0.9)
Costs : $k_i^R=2000$, $k_i^{PR}=1200$

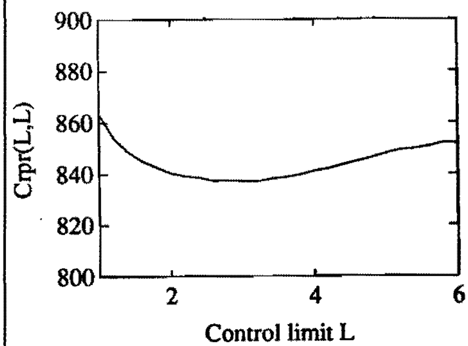


Fig. 6.14a: $C_{rpr}(L,L)$

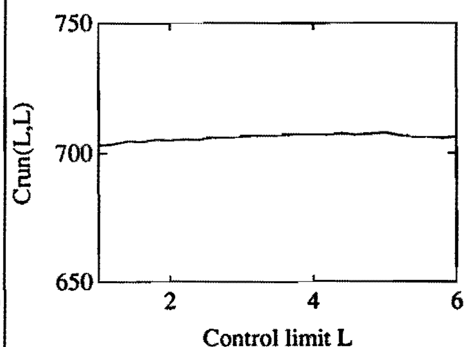


Fig. 6.14b: $C_{run}(L,L)$

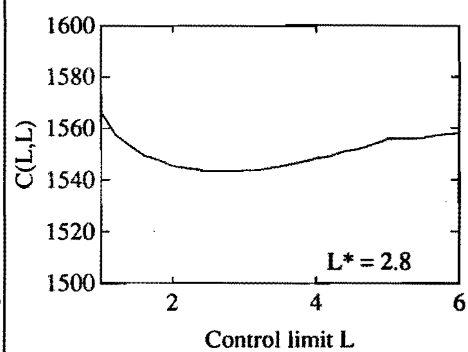


Fig. 6.14c: C(L,L)

Life times : W(5,3)

Repair times : LN(0.5,0.9)

Costs : k_i^R=2000, k_i^{PR}=1200

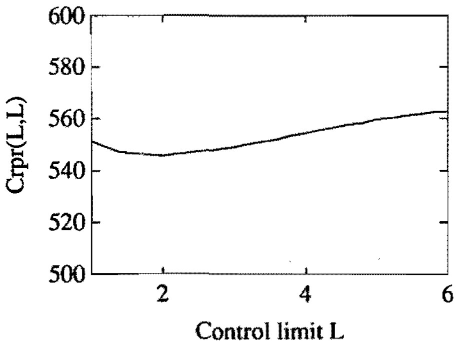


Fig. 6.15a: C_{rpr}(L,L)

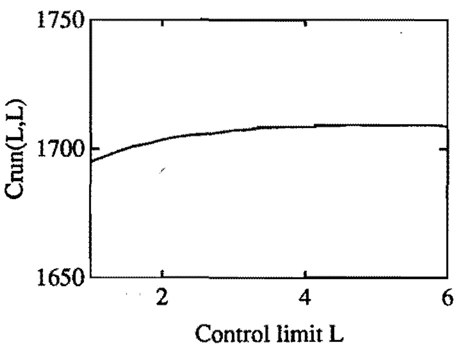


Fig. 6.15b: C_{run}(L,L)

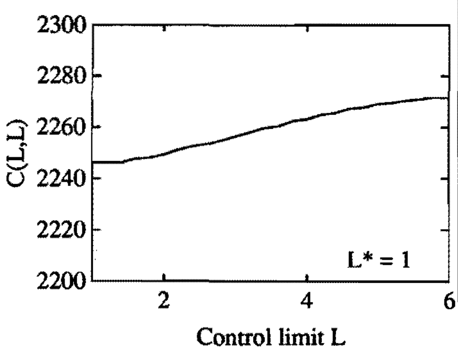


Fig. 6.15c: C(L,L)

Life times : W(5,3)

Instantaneous repairs

Costs : k_i^R=2000, k_i^{PR}=1200

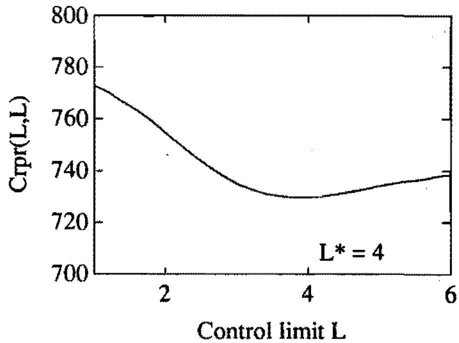


Fig. 6.16a: C_{rpr}(L,L)

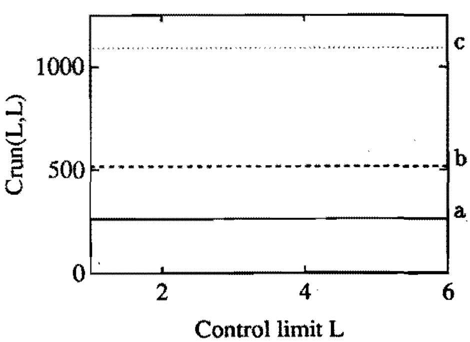


Fig. 6.16b: C_{run}(L,L)

Cases a, b, c: availability approximated by respectively $\hat{A}(0.33)$, $\hat{A}(0.91)$ and $\hat{A}(2.47)$

Life times : W(5,3)

Repair times : LN(−1.5,0.9)

Costs : k_i^R=2000, k_i^{PR}=1600

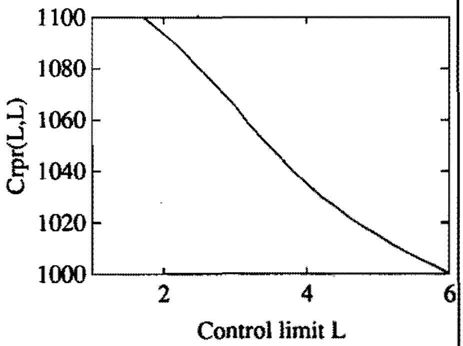


Fig. 6.17a: C_{rpr}(L,L)

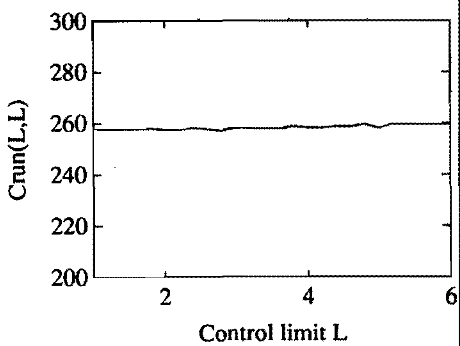


Fig. 6.17b: C_{run}(L,L)

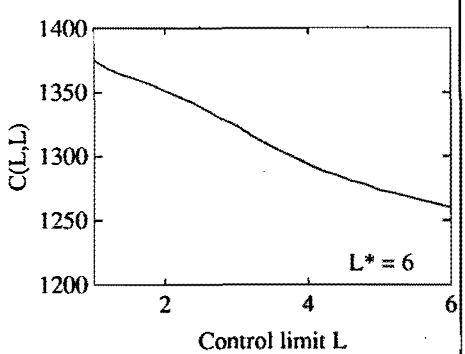


Fig. 6.17c: C(L,L)

Life times : W(5,3)

Repair times : LN(−0.5,0.9)

Costs : k_i^R=2000, k_i^{PR}=1600

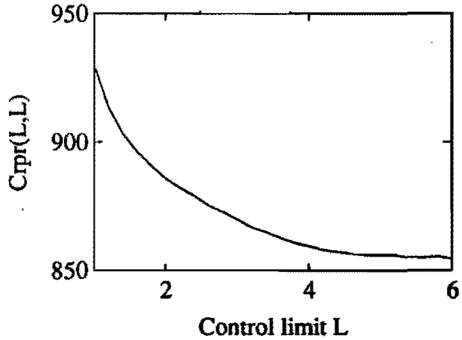


Fig. 6.18a: C_{rpr}(L,L)

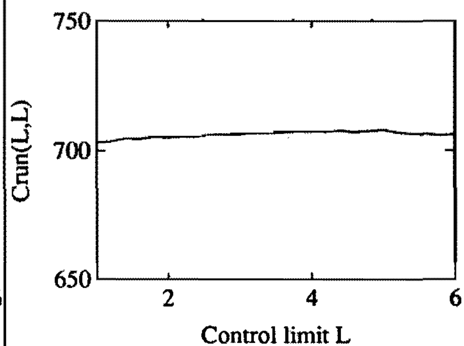


Fig. 6.18b: C_{run}(L,L)

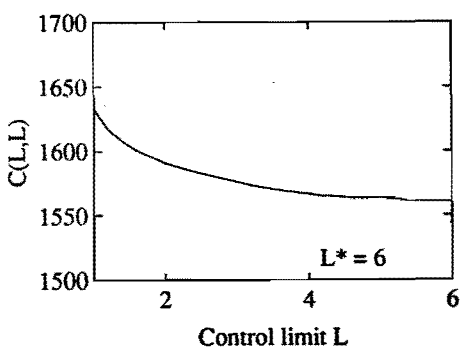


Fig. 6.18c: C(L,L)

Life times : W(5,3)

Repair times : LN(0.5,0.9)

Costs : k_i^R=2000, k_i^{PR}=1600

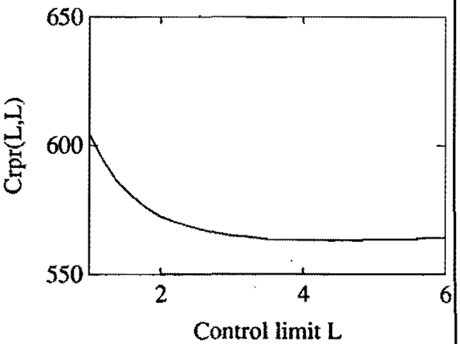


Fig. 6.19a: C_{rpr}(L,L)

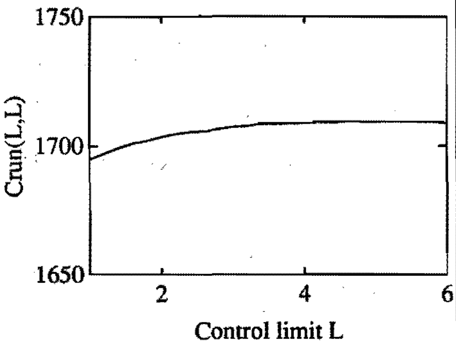


Fig. 6.19b: C_{run}(L,L)

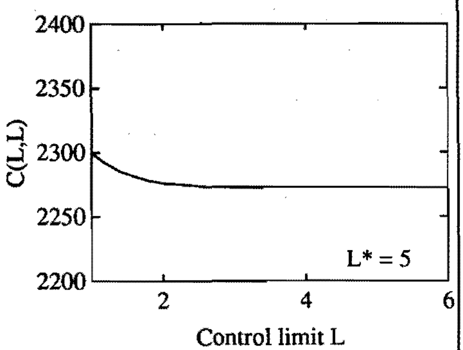


Fig. 6.19c: C(L,L)

Life times : W(5,3)

Instantaneous repairs

Costs : k_i^R=2000, k_i^{PR}=1600

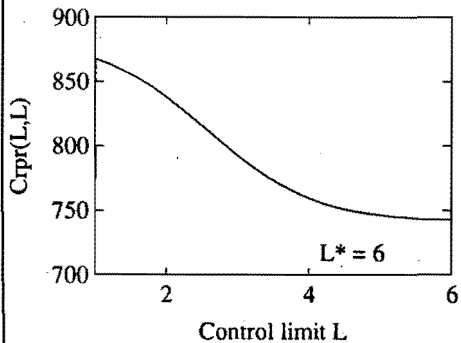


Fig. 6.20a: C_{rpr}(L,L)

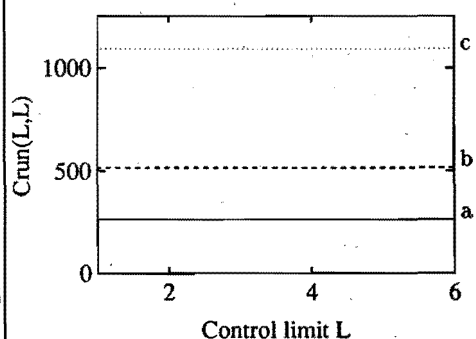


Fig. 6.20b: C_{run}(L,L)

Cases a, b, c: availability approximated by respectively $\hat{A}(0.33)$, $\hat{A}(0.91)$ and $\hat{A}(2.47)$

iv. The optimum L^*, obtained by an approximate analysis with instantaneous repairs, appears to give a good indication of the optimum for non-zero repair times if k_i^{PR}/k_i^R is relatively small or relatively large. For the cases in between these extremes, it seems important to include the repair times in the model, in order to obtain the exact value of L^*. Note that it is just the latter case which is of particular interest, since it is understandable that L^* is small, respectively large, when a preventive replacement is relatively cheap, respectively expensive. Thus this provides a reason to perform an analysis with non-zero repair times.

v. The order of magnitude of the costs $C_{rpr}(L,L)$, as well as $C_{run}(L,L)$, heavily depends on the repair time distribution. Consequently, the model with instantaneous repairs is not adequate to approximate the expected costs per unit of time.

vi. The reduction of costs which can be obtained in the optimum L^* is a few percent in the cases under consideration. Although this reduction is relatively small, the absolute value can be quite large in practice. On the other side, it seems plausible that the profits which can be obtained under the OARP are more interesting, since under the OARP it is not necessary to wait until a failure occurs, in order to perform a preventive replacement.

To give an impression of the operating characteristics under the OFRP, some are plotted as a function of L in figure 6.21 and 6.22 (for repair which have a $\mathcal{LN}(-0.5,0.9)$ distribution and life times which follow a $\mathcal{W}(5,3)$ distribution). Figure 6.21 represents the behaviour of N_i^O, N_i^R, N_i^{PR}, N_i^{SMA} and N^{CC}, as defined in section 6.3, and figure 6.22 shows the availability and unavailability. As in section 6.3 the unavailability U is decomposed into $U=U_{fr}+U_{cc}$ and with respect to the availability the following decomposition is made. The availability A is written as $A=A_{oo}+A_{or}+A_{o=r}$, where A_{oo} represents the probability mass in D_1, A_{or} the mass in $(D_2-D_2^i)\cup(D_3-D_3^i)$ and $A_{o=r}$ the probability mass concentrated in $D_2^i\cup D_3^i$.

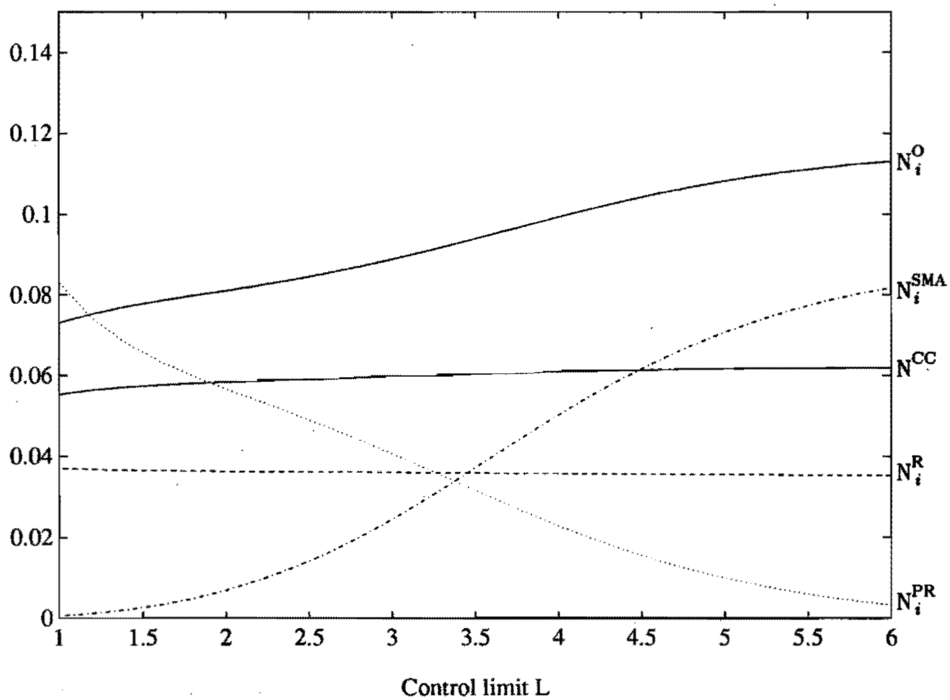


Fig. 6.21: The operating characteristics N_i^O, N_i^R, N_i^{PR}, N_i^{SMA} and N^{CC}.

For obvious reasons the operating characteristics N_i^O and N_i^{SMA} are increasing in L, whereas N_i^{PR} is decreasing. N^{CC} is more or less constant, since the availability is constant and the common cause effect is exponentially distributed. Further, figure 6.21 shows that N_i^R is almost constant as a function of L. Figure 6.22 shows that U_{fr}, U_{cc} and hence the limiting unavailability U are approximately constant, which also holds for the A_{oo} (not plotted in fig. 6.22). However, since N_i^{PR} is decreasing in L and N_i^R is approximately constant, the quantities $A_{o=r}$ and A_{or} are decreasing, respectively increasing in L.

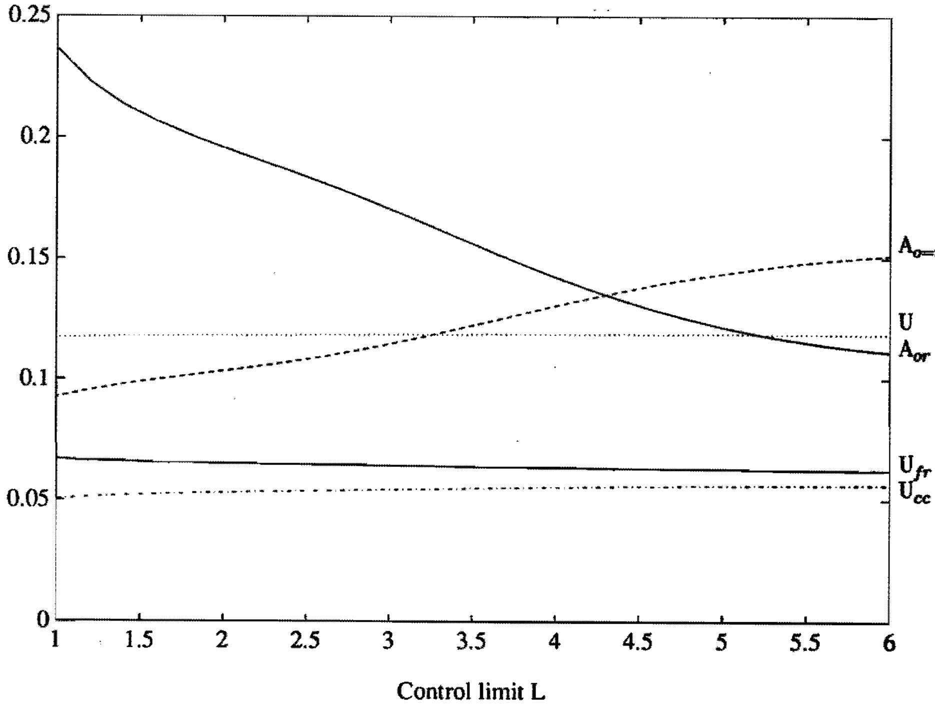


Fig. 6.22: The performance measures U, U_{fr}, U_{cc}, A_{or} and $A_{o=r}$.

Summarising chapter 6, an opportunistic replacement policy of the control limit type is analysed for the two-unit dependent parallel system. A set of differential equations is derived for the stationary joint pdf of the state description process $\{X(t),t\geq0\}$ and it is shown how a solution can be found from the boundary conditions. Once the stationary joint pdf has been obtained, it is shown how the system's long run operating characteristics are computed and (given the costs of repairs, replacements, etc.) the expected costs per unit of time in the long run. The optimal pair of control limits can be found by minimising the total expected costs per unit of time in the long run. To decide whether the presence of a second repair facility is worthwhile, a similar analysis is performed for the situation with two repair men. The cost criterion can be used to compare both situations. Further, the situation with instantaneous repairs is analysed, to investigate whether the expected costs per unit of time and the optimum value of the control limit can be approximated by assuming zero repair times. Numerical examples illustrate the techniques and it appears that the assumption of instantaneous repairs often results in bad approximations.

7. AGEING

7.1 Introduction

A repairable system was defined in section 1.2 as a system which can be restored to fully satisfactory performance by any method other than replacement of the entire system. On the other hand, a nonrepairable system is discarded after its first failure. Since repairs were assumed to be perfect and to restore the normal operational efficiency of the units, the parallel system under consideration can be modelled as a repairable system with two nonrepairable units. In this case ageing at *component* level can be modelled by an appropriate choice of the life time distribution of a component. Distributions with increasing hazard rate are suitable for modelling the wearout of parts and decreasing hazard functions can be used to describe burn–in phenomena. However, the methods and models described in this study can not be applied when ageing occurs at *system* level, as expressed by an increasing or decreasing ROCOF. In fact, the assumption of perfect repairs makes it possible to give a neat analysis of the parallel system, using the technique of regenerative point processes, but the presence of an imbedded renewal process renders the system's ROCOF either periodic or constant. Hence the models in the previous chapters are not appropriate to describe deterioration at system level.

In this chapter the ageing of a repairable system is discussed. Ageing can result in failures which tend to occur more frequently with increasing operating time, in increasing operating or repair costs, or increasing repair times. The method discussed applies to both deteriorating and improving systems, but the terminology will be in terms of deterioration. Although deterioration is plausible in many practical situations, models for deteriorating repairable systems are largely ignored in the literature. Reasons for this state of affairs include the need for *non*stationary models and the fact that the time dependent behaviour is of major importance for a repairable system. Hence, the models and techniques are in general more complicated and consume more computing time than in the case of a nonrepairable system.

Among the models for repairable systems are:

i. The nonhomogeneous Poisson process (NHPP). As mentioned in section 2.3, the NHPP models a bad–as–old situation.

ii. Differential equations models (Ascher *et al.*, 1984), which describe the relationship between the rate of deterioration and the operating time. The rate of deterioration is expressed in *e.g.* the number of failures at time t or the mean time between failures. Reliability growth is overemphasised in the literature, but many of the models used to depict reliability growth are also applicable to deteriorating systems.

iii. Markov models (Ross, 1970, Tijms, 1986). Well known in the literature is the situation of Markovian deterioration, where a system's life consists of a number of exponentially distributed stages of degradation. Mine *et al.* (1974) and Van der Duyn-Schouten *et al.* (1989) consider a two-unit parallel system with Markovian degradation of the operating units.

iv. The proportional hazards model (PHM), described in section 2.9, and in particular the version introduced by Prentice *et al.* (1981).

v. The branching Poisson process (BPP), see section 2.6. However, the only reported application of a BPP is by Lewis (1964), who analysed the failure patterns of three computers.

Apart from the above models time series techniques (Box *et al.*, 1970) are potentially applicable to repairable systems failure data analysis. The advantage of the techniques is the great flexibility of *e.g.* the autoregressive integrated moving average model. A disadvantage is the need for a large number of failure data to implement the model. The only reported application of time series techniques to repairable systems seems to be by Singpurwalla (1978), who investigates the relationship between successive up and down times.

It is not intended to develop here an all embracing model for repairable systems, able to handle all the real world factors described by Ascher *et al.* (1984). Instead, an extension of the models in the previous chapters towards the subject of ageing is developed and it is shown how the regeneration point technique can be used to obtain the system's performance measures. The deterioration of the system is modelled by embedding the two-unit parallel system in a larger system, called the system body. Repairs of the units of the parallel system are assumed to be perfect, *i.e.* after repair the units are as good as new. On the other hand, the system body is not necessarily as good as new after repair and in this way, degradation of the system body represents the deterioration of the system as a whole.

The model is described in greater detail in section 7.2. The performance measures are derived in section 7.3 and in section 7.4 the advantages of the use of phase type distributions are outlined.

7.2 Model description

In order to model deterioration, the two-unit dependent parallel system is supposed to be part of a larger system, called the system body. The failure behaviour of the system body is modelled by an additional component, denoted as the system component C_S. The parallel system is assumed to be connected in series with the system component, as shown in figure 7.1. Hence, the system is down if either the system component or the parallel system is down.

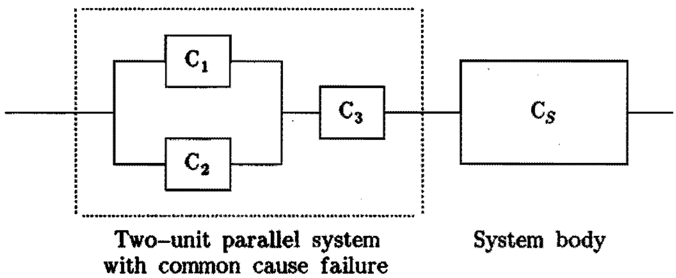


Fig. 7.1: System configuration

The stochastic behaviour of the parallel system is as described in section 1.4. With respect to the system component, some additional assumptions are made, to create points in time which show a lack-of-memory property, as will become clear later on in this chapter. The model assumptions are:

i. The i^{th} life time of the system component has pdf $f_i^s(.)$ and the i^{th} repair time pdf $g_i^s(.)$, $i=1,2,...$

ii. When the system component fails during a repair of component C_1 or C_2, the repair of C_S has priority.

iii. At a failure of the system component, the components C_1 and C_2 are both overhauled after the completion of the repair of C_S. For simplicity, overhauls are assumed to be perfect and the time required is distributed identically to the repair times of C_1 and C_2.

iv. When the parallel system is down and the system component is not in repair, the system component is hot standby, *i.e.* it behaves as if it is operating.

v. After N failures of the system component, the entire system is discarded and replaced by an identical but new system. The time needed to replace the system is a random variable with pdf $g_N^s(.)$.

The state description process of the model is constructed largely as in the chapters 3, 4 and 5. Note that at any time t, the component C_i (i=1,2) is operating, under repair or waiting for repair and the common cause component is either operating or waiting to restart operating. The state of the common cause component follows from the state of C_1 and C_2. Further, the system component is either operating or under repair. However, since the life times and repair times of C_S are not necessarily identically distributed, the number of the life and repair times has to be included in the state description of the system component. The i^{th} life and repair are denoted by o_i, respectively r_i (i=1,2,..). Now, triples (x_1,x_2,x_s) can be used to denote the system state at any time t, where x_i represents the state of C_i (i=1,2) and x_s the state of C_S. Let period i cover the time period starting at the moment that C_S begins its i^{th} life time and ending at the moment of the i^{th} repair completion of C_S. Then, the state space S_X^i of the system state description process $\{X(t),t\geq 0\}$ in period i is

$$S_X^i = \{(o,o,o_i),(r,o,o_i),(o,r,o_i),(r,w,o_i),(w,r,o_i),(w,w,r_i)\}$$

and under the above assumptions, the system's one-step transition diagram during period i is given in figure 7.2. Obviously, the system state description process $\{X(t),t\geq 0\}$ has state space S_X, where

$$S_X = \bigcup_{i=1}^{N} S_X^i$$

and the one-step transition diagram is obtained by connecting N single period transition diagrams, as shown in figure 7.3. By definition, state (w,w,r_N) represents a replacement on system level, which ends by a transition to state (o,o,o_1). For obvious reasons, the length of N periods is called a (system) life cycle. The pdf of the life cycle, denoted by $\gamma(.)$, is

$$\gamma(t) = f_1^s(t)*g_1^s(t)*...*f_N^s(t)*g_N^s(t).$$

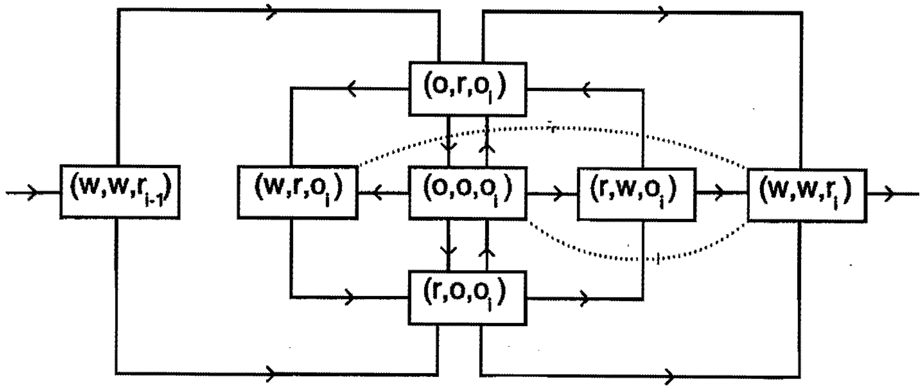


Fig. 7.2: One step transition diagram of period i

Theoretically, life cycles form a renewal process, but the practical relevance of this renewal process is doubtful. It is plausible that only the transient behaviour in the first life cycle is important in practice.

Remark that the introduction of overhauls is achieved artificially: it is a trick to generate a kind of regenerative events, which allow an analysis of the system without needing supplementary variables. Strictly, the process $\{X(t),t\geq0\}$ is not regenerative with respect to a repair completion of the system component, as the life times of C_S are not identically distributed. However, note that the continuation of the process $X(t)$ *beyond* a repair completion is independent of the history *until* the repair completion, since:

i. the system component *starts* the next life time and

ii. the repair facility *starts* an overhaul of the dependent parallel system

at the moment of a repair completion.
This lack-of-memory property will be exploited in the analysis of the performance measures: as in chapter 5, time epochs which show the lack-of-memory property are used to derive recurrence relations in terms of convolution integrals for the (interval) reliability and (joint) availability of the system.

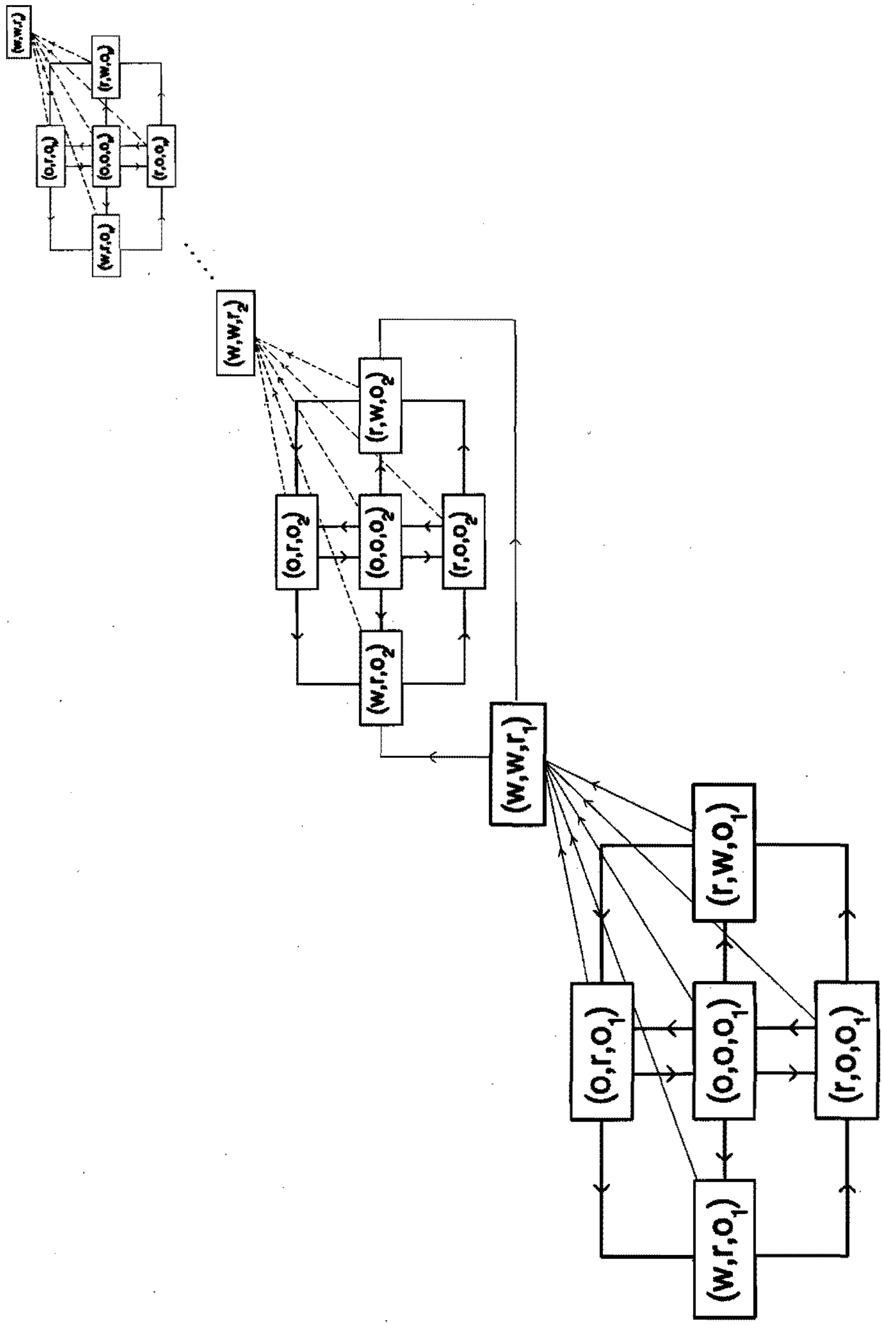


Fig. 7.3: The system's one-step transition diagram of one life cycle

An additional advantage of the introduction of overhauls is the possibility of performing an analysis per period, initiated by a repair completion of the system component. Moreover, the occurrence of overhauls allows the bivariate distribution of the dependent parallel system to differ per period, without giving analytical problems, as will be clear in the next section. This situation is attractive when the deterioration of the system component influences the efficiency of the parallel system. On the other hand, treating the system component as an artificial component (without repair time), offers the possibility of modelling internal and external common cause failures separately. In this case, failures of the system component mark the stages of degradation of the system.

7.3 Performance measures

In general the process $\{X(t),t\geq 0\}$ is nonstationary when the system deteriorates and hence only the transient performance measures are of interest. Expressions for the reliability, availability, interval reliability and joint availability are derived here and the role of the regenerative events in the previous chapters is played by events which show the lack-of-memory property, mentioned in section 7.2. Thus, the performance measures are obtained, conditioned on an entry into state (o,o,o_1), (r,w,o_i), (w,r,o_i) or (w,w,r_i).

To start with, an expression is obtained for the system reliability. It is easily seen that $R_{ooo_1}(t)$, the reliability at time t given a start in (o,o,o_1), is

$$R_{ooo_1}(t) = \bar{F}_1^s(t)\, R_{oo}(t), \tag{7.1}$$

where $R_{oo}(t)$ is the reliability of the two-unit dependent parallel subsystem (see chapters 3, 4 and 5).

Secondly, the reliability conditioned on an entry into state (r,o,o_i) from state (w,r,o_i) is obtained. Two situations are considered, depending on whether the system component or the two-unit dependent parallel system has failed. When the system restarts operating after a failure of the system component, the process X(t) moves directly from (w,w,r_{i-1}) via (w,r,o_i) to (r,o,o_i) and the operating time of C_S equals the repair time of component C_2 at the moment state (r,o,o_i) is entered. Denote the reliability, conditioned on an entry at the time origin into (r,o,o_i), while the visit is made directly from (w,w,r_{i-1}) via (w,r,o_i) to (r,o,o_i), by $R^{s}_{roo_i}(.)$. Then, clearly

$$R^{s}_{roo_i}(t) = \int_0^{\infty} \frac{F^{s}_{i}(u+t)}{F^{s}_{i}(u)} \, g_2(u) \, du \, R_{ro}(t). \tag{7.2}$$

An expression for $R^{s}_{oro_i}(t)$ is obtained by interchanging the roles of the components C_1 and C_2.

However, the situation is different when state (r,o,o_i) is entered after two or more visits to (w,r,o_i), *i.e.* after one or more failures of the parallel system during life time i of C_S. Two situations are distinguished now. In a practical application, when the operating time of C_S is known at the moment the system restarts working, computation of the reliability is trivial. Let $R_{roo_i}(t|u)$ denote the system reliability at time t, with the provision that state (r,o,o_i) is entered at t=0 and given that C_S has operated for u units of time at t=0. Then it is immediate that

$$R_{roo_i}(t|u) = \frac{F^{s}_{i}(u+t)}{F^{s}_{i}(u)} \, R_{ro}(t). \tag{7.3}$$

Secondly, when the operating time of C_S is unknown at the moment of entrance into (r,o,o_i), the expression for the reliability is in terms of the backward recurrence time distribution of the life time of component C_S. Let $b^{s}_{i}(.)$ be the pdf of the backward recurrence time of the i^{th} life time of C_S and let $R^{P}_{roo_i}(.)$ denote the reliability under the above condition that state (r,o,o_i) is entered after two or more visits to (w,r,o_i), then

$$R^{P}_{roo_i}(t) = \int_0^{\infty} \frac{F^{s}_{i}(u+t)}{F^{s}_{i}(u)} \, b^{s}_{i}(u) \, du \, R_{ro}(t) \tag{7.4}$$

and a similar expression follows for $R^{P}_{oro_i}(t)$. With respect to the backward recurrence time, explicit expressions seem to be available, only for the limit distribution (Feller, 1966, pp. 354–357). Hence, the practical relevance of formula (7.4) seems minimum.

Further, note that the mean time to system failure, given a start in a particular state, can be computed by integration of the corresponding reliability function.

The system availability is obtained as in section 3.4, with the life cycle playing the role of the interval between two successive E_1 events. Hence, first the system availability at time t is computed, provided that the last life cycle started at time t–u and subsequently the number of completed life cycles at t–u is varied from zero to infinity. Thus the pdf $f_{11}(t)$ in equation (3.14) is replaced by $\gamma(t)$ and $\alpha(t)$ by the 'single cycle availability' $A^c_{ooo_1}(t)$, defined as

$$A^c_{ooo_1}(t) = F^s_1(t)\ A_{oo}(t) + \int_0^t f^s_1(u)\ A^c_{wwr_1}(t-u)\ du. \qquad (7.5)$$

In equation (7.5):

i. $A_{oo}(t)$ represents the availability of the two–unit dependent parallel subsystem, conditioned on a start in state (o,o) at the time origin. Expressions for $A_{oo}(t)$ have been derived in the sections 3.4 and 5.4.

ii. $A^c_{wwr_1}(t)$ is the system availability conditioned on a start in (w,w,r_1) at the time origin and given that the system is still in its first life cycle at time t.

In other words: $A^c_{ooo_1}(t)$ represents the system availability at time t conditioned on a start in (o,o,o_1) and given that the length of the first life cycle exceeds time t.

Further, starting in state (w,w,r_i), the repair of the system component has to be completed before time t, in order to be available at time t. Thus, the single cycle availability $A^c_{wwr_i}(.)$ satisfies

$$A^c_{wwr_i}(t) = \int_0^t g^s_i(u) \left[p\ A^c_{rwo_{i+1}}(t-u) + (1-p)\ A^c_{wro_{i+1}}(t-u)\right] du, \quad i=1,..,N-1$$

and, from the mutually exclusive events that life time i of the system component exceeds time t or not,

$$A^c_{rwo_i}(t) = F^s_i(t)\ A_{rw}(t) + \int_0^t f^s_i(u)\ A^c_{wwr_i}(t-u)\ du, \quad i=1,..,N,$$

where, by definition, $A^c_{wwr_N}(.)=0$.

However, the number of life cycles completed at time t is not bounded. Hence, a formal expression for $A_{ooo_1}(t)$, the system availability at time t, conditioned on an entrance into state (o,o,o_1) at the time origin, is (*cf.* (3.14))

$$A_{ooo_1}(t) = A^c_{ooo_1}(t) + \sum_{n=1}^{\infty} \gamma^{(n)}(t)*A^c_{ooo_1}(t). \tag{7.6}$$

Further, starting in state (r,w,o_i) at $t=0$, the first life cycle ends with probability $f^s_i(t)*g^s_i(t)*...*f^s_N(t)*g^s_N(t)$ before time t and hence (*cf.* (3.26))

$$A_{rwo_i}(t) = A^c_{rwo_i}(t) + f^s_i(t)*g^s_i(t)*...*f^s_N(t)*g^s_N(t)*A_{ooo_1}(t), \; i=1,..,N. \tag{7.7}$$

Analogously, for i=1,..,N

$$A_{wwr_i}(t) = A^c_{wwr_i}(t) + g^s_i(t)*f^s_{i+1}(t)*g^s_{i+1}(t)*...*f^s_N(t)*g^s_N(t)*A_{ooo_1}(t). \tag{7.8}$$

Note that the (time dependent) behaviour in the *first* cycle is of special interest when t is relatively small or when replacements at system level do not occur. In the latter case, the formulas (7.6)–(7.8) have no practical relevance anymore.

Expressions for the interval reliability and joint availability are obtained in a similar way to the availability function. First, a recurrence relation is derived for the interval reliability or the joint availability in the first life cycle and secondly the number of life cycles is varied. As before, the superscript 'c' is used to denote the single cycle interval reliability or joint availability.

The mutually exclusive events that the first failure of C_S occurs before time t or not, yield (*cf.* (7.5))

$$R^c_{ooo_1}(t,\tau) = \bar{F}^s_1(t+\tau)\, R_{oo}(t,\tau) + \int_0^t f^s_1(u)\, R^c_{wwr_1}(t-u,\tau)\, du, \tag{7.9}$$

where $R_{oo}(t,\tau)$ is the interval reliability of the two-unit dependent parallel subsystem. Obviously,

$$R^c_{wwr_i}(t,\tau) = \int_0^t g^s_i(u) \left[p\, R^c_{rwo_{i+1}}(t-u,\tau) + (1-p)\, R^c_{wro_{i+1}}(t-u,\tau)\right] du, \; i=1,..,N-1 \tag{7.10}$$

and

$$R^c_{rwo_i}(t,\tau) = \bar{F}^s_i(t+\tau)\, R_{rw}(t,\tau) + \int_0^t f^s_i(u)\, R^c_{wwr_i}(t-u,\tau)\, du, \quad i=1,..,N, \tag{7.11}$$

where $R^c_{wwr_N}(.)=0$, by definition.

Formally, from (7.9)–(7.11),

$$R_{ooo_1}(t,\tau) = R^c_{ooo_1}(t,\tau) + \sum_{n=1}^{\infty} \gamma^{(n)}(t)*R^c_{ooo_1}(t,\tau),$$

$$R_{rwo_i}(t,\tau) = R^c_{rwo_i}(t,\tau) + f^s_i(t)*g^s_i(t)*...*f^s_N(t)*g^s_N(t)*R_{ooo_1}(t,\tau)$$

and

$$R_{wwr_i}(t,\tau) = R^c_{wwr_i}(t,\tau) + g^s_i(t)*f^s_{i+1}(t)*g^s_{i+1}(t)*...*f^s_N(t)*g^s_N(t)*R_{ooo_1}(t,\tau),$$

where $\gamma(t)*R^c_{ooo_1}(t,\tau)$ is defined as

$$\gamma(t)*R^c_{ooo_1}(t,\tau) = \int_0^t \gamma(u)\ R^c_{ooo_1}(t-u,\tau)\ du.$$

The single cycle joint availability $A^c_{ooo_1}(t,\tau)$ is obtained by considering the mutually exclusive events that the first failure of C_S occurs:

i. before time t,

ii. in the interval $(t,t+\tau)$ or

iii. after time $t+\tau$.

It is easily seen that

$$A^c_{ooo_1}(t,\tau)=\bar{F}^s_1(t+\tau)\ A_{oo}(t,\tau) + \int_0^t f^s_1(u)\ A^c_{wwr_1}(t-u,\tau)du + \int_t^\tau f^s_1(u)\ A^c_{wwr_1}(\tau-u)du,$$

$$A^c_{rwo_i}(t,\tau)=\bar{F}^s_i(t+\tau)\ A_{rw}(t,\tau) + \int_0^t f^s_i(u)\ A^c_{wwr_i}(t-u,\tau)du + \int_t^\tau f^s_i(u)\ A^c_{wwr_i}(\tau-u)du$$

and

$$A^c_{wwr_i}(t,\tau) = \int_0^t g^s_i(u)\ \left[p\ A^c_{rwo_{i+1}}(t-u,\tau) + (1-p)\ A^c_{wro_{i+1}}(t-u,\tau)\right]du.$$

Consequently,

$$A_{ooo_1}(t,\tau) = A^c_{ooo_1}(t,\tau) + \sum_{n=1}^{\infty} \gamma^{(n)}(t)*A^c_{ooo_1}(t,\tau),$$

$$A_{rwo_i}(t,\tau) = A^c_{rwo_i}(t,\tau) + f^s_i(t)*g^s_i(t)*...*f^s_N(t)*g^s_N(t)*A_{ooo_1}(t,\tau)$$

and

$$A_{wwr_i}(t,\tau) = A^c_{wwr_i}(t,\tau) + g^s_i(t)*f^s_{i+1}(t)*g^s_{i+1}(t)*...*f^s_N(t)*g^s_N(t)*A_{ooo_1}(t,\tau).$$

So far, no assumptions have been made with respect to the life or repair time distributions of the components C_1, C_2, C_3 and C_s. Nevertheless, it is shown that expressions can be derived for the (interval) reliability and (joint) availability in this general situation. However, in the next section it is shown that the analysis is simplified considerably by the introduction of phase type distributions. Phase type distributions render the state description process Markovian and hence the randomisation technique can be used to compute the transient performance measures, which are of particular interest. The possibilities of Markov models will be illustrated by an example.

7.4 Phase type distributed life and repair times

Under phase type distributed life and repair times, the state description process $\{X(t), t \geq 0\}$ becomes Markovian when it includes the phases of the life and repair time distributions. In this section, the generator of the Markov process is constructed, while assumption *iii* in section 7.2 is relaxed. More precisely, consider the system in figure 7.1 with the following assumptions:

i. The dependent parallel subsystem has a bivariate phase type life time distribution with exponentially distributed common cause, *i.e.* component C_i ($i=1,2$) has a phase type distribution with representation (α_i, T_i) and dimension n_i and the common cause component has a phase type distribution with representation $(1, -\lambda)$ and dimension 1.

ii. The components in the parallel system have a phase type repair time distribution with representation (β_i, S_i) and dimension m_i, $i=1,2$.

iii. Life time i and repair time i of the system component are phase type distributed with representation (α_{3i}, T_{3i}), respectively (β_{3i}, S_{3i}), and dimension n_{3i}, respectively m_{3i}.

iv. The repair facility operates with FIFO (first–in–first–out) repair policy. Repairs are not interrupted anymore. Repairs of C_1 and C_2 are perfect.

v. The system component does not fail when the parallel system is down and *vice versa*.

vi. When a failure at system level occurs, the non-failed components behave in a bad as old fashion, *i.e.* they are bad as old when the system restarts operating.

vii. After N failures of the system component, the entire system is discarded and replaced by an identical but new system.

In order to satisfy assumption *vi*, the state 'memory' is added to the triple 'operating', 'repair' and 'waiting for repair'. Its function is to store the state of a non-failed component at the moment a failure at system level occurs. Hence, the components of the parallel system are in 'memory' (abbreviated as 'm') during a repair of the system component and the system component is in 'memory' when the parallel system is down.
Using ordered triples (x_1,x_2,x_s) to denote the state of the system and the subscript i to denote period i, the state space S_X^i of period i is

$$S_X^i=\{(o,o,o_i),(r,o,o_i),(o,r,o_i),(r,w,m_i),(w,r,m_i),(m,r,w_i),(r,m,w_i),(m,m,r_i)\}$$

and the one-step transition diagram of the state description process $\{X(t),t\geq 0\}$ is given in figure 7.4.

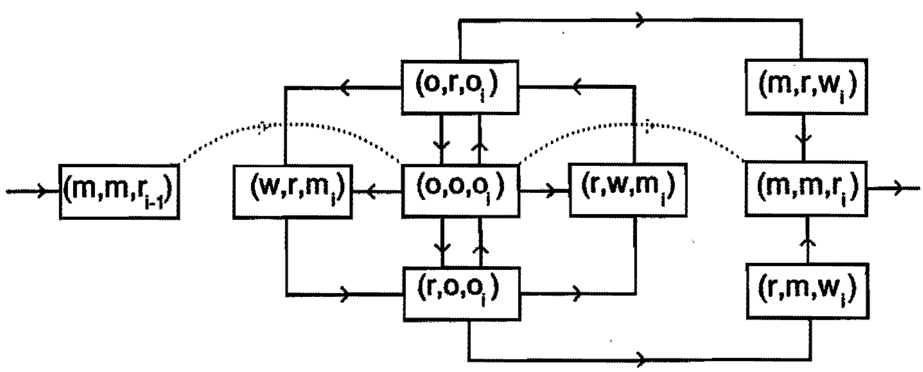


Fig. 7.4: One-step transition diagram of period i

Subsequently, to make the state description process $\{X(t),t\geq 0\}$ Markovian, the state space is extended to include the phases of the life and repair time distributions. As in section 4.6, phase i of a component's life time is denoted by ℓ_i, phase i of a repair time by $\imath_i$ and ϖ represents a component waiting for repair. A non-failed component, which was in phase i of its life time distribution at the moment the system went down, is denoted by m_i. Further, the system component has an additional index to denote the period number.

Hence, the state space S_X^p of period p $(1\leq p\leq N)$ is extended to

$$S_X^p=(o,o,o_p)\cup(r,o,o_p)\cup(o,r,o_p)\cup(r,w,m_p)\cup(w,r,m_p)\cup(m,r,w_p)\cup(r,m,w_p)\cup(m,m,r_p)$$

where

$$\begin{aligned}
(o,o,o_p) &= \{(\ell_i,\ell_j,\ell_{kp})\}_{1\leq i\leq n_1,1\leq j\leq n_2,1\leq k\leq n_3}\\
(o,r,o_p) &= \{(\ell_i,\imath_j,\ell_{kp})\}_{1\leq i\leq n_1,1\leq j\leq m_2,1\leq k\leq n_3}\\
(r,o,o_p) &= \{(\imath_i,\ell_j,\ell_{kp})\}_{1\leq i\leq m_1,1\leq j\leq n_2,1\leq k\leq n_3}\\
(m,m,r_p) &= \{(m_i,m_j,\imath_{kp})\}_{1\leq i\leq n_1,1\leq j\leq n_2,1\leq k\leq m_3}\\
(r,w,m_p) &= \{(\imath_i,\varpi,m_{kp})\}_{1\leq i\leq m_1,1\leq k\leq n_3}\\
(w,r,m_p) &= \{(\varpi,\imath_j,m_{kp})\}_{1\leq j\leq m_2,1\leq k\leq n_3}\\
(r,m,w_p) &= \{(\imath_i,m_j,\varpi_p)\}_{1\leq i\leq m_1,1\leq j\leq n_2}\\
(m,r,w_p) &= \{(m_i,\imath_j,\varpi_p)\}_{1\leq i\leq n_1,1\leq j\leq m_2}.
\end{aligned}$$

The generator Q of the Markov process under consideration is constructed as follows. The matrix Q is partitioned as

$$Q=\begin{bmatrix} Q_{11} & Q_{12} & & & \\ & Q_{22} & Q_{23} & & \\ & & \ddots & \ddots & \\ & & & Q_{N-1N-1} & Q_{N-1N} \\ Q_{N1} & & & & Q_{NN}\end{bmatrix}, \tag{7.12}$$

where

i. the matrices Q_{ii} (i=1,..,N) contain the rates belonging to transitions from state u_i to v_i, where $(u_i,v_i)\in S_X^i x S_X^i$.

ii. the matrices Q_{ii+1} (i=1,..,N−1) have entries corresponding to transitions from the set (m,m,r_i) to (o,o,o_{i+1}).

iii. Q_{N1} contains the transition rates corresponding to a renewal on system level, *i.e.* transitions from the set (m,m,r_N) to (o,o,o_1).

Let the above sets of states be lexicographically ordered (per period) on the indices of the phases of the life and repair time distributions. Then, using the notation from section 4.6, Q_{ii} is given by figure 7.5. Secondly, the matrix Q_{ii+1} has dimensions which equal the cardinality of S_X^i (rows) and S_X^{i+1} (columns) and which has non-zero entries $I_1 \otimes I_2 \otimes S_{3i}^0 \alpha_{3i+1}^T$ on the positions corresponding to transitions from (m,m,r_i) to (o,o,o_{i+1}). The entries of the matrices Q_{ii} and Q_{ii+1} are obtained in a similar way as the entries of the generator in 4.6.

Notice that it is immediate from (7.12) that the generator Q of the Markov process is sparse, which is reinforced by the fact that the matrices Q_{ii} and Q_{ii+1} are all sparse as well. Hence the randomisation technique is useful for computing the performance measures. When the sparseness of Q is exploited to the fullest degree and the generator is stored in an economical way, the performance measures can be computed without difficulties, even when the state space is very large and contains (say) 1000 states or more (Grassmann, 1977, Gross *et al.*, 1984).

Finally, the flexibility of Markov models is illustrated by considering some simple modifications of the transition mechanism, resulting in interesting models. The generator Q above is considered with the following modification. The matrix Q_{N1} is deleted and Q_{NN} is replaced by an identical matrix, with the elements of $I_1 \otimes I_2 \otimes S_{3N}^0 \alpha_{3N}^T$ at the positions corresponding to transitions from (m,m,r_N) to (o,o,o_N). In this case, the set S_X^N is an absorbing set which can be considered as an equilibrium state: after N−1 periods, or stages of degradation, the system reaches its steady state. The time dependent behaviour is now determined mainly by the transient set $S_X^1 \cup .. \cup S_X^{N-1}$ and its corresponding transition rates and the limiting or steady state behaviour by the Markov process with state space S_X^N and generator Q_{NN}. The methods and techniques described in this study can now be used to analyse the system's time dependent as well as its steady state failure behaviour.

	(o,o,o_i)	(r,o,o_i)	(o,r,o_i)	(r,w,m_i)	(w,r,m_i)	(m,m,r_i)	(r,m,w_i)	(m,r,w_i)
(o,o,o_i)	$T_1\otimes I_2\otimes I_{3i}+$ $I_1\otimes T_2\otimes I_{3i}+$ $I_1\otimes I_2\otimes T_{3i}+$ $-\lambda I_1\otimes I_2\otimes I_{3i}$	$T_1^0\beta_1^T\otimes I_2\otimes I_{3i}$	$I_1\otimes T_2^0\beta_2^T\otimes I_{3i}$	$p_1\lambda e_1\beta_1^T\otimes e_2\otimes I_{3i}$	$p_2\lambda e_1\otimes e_2\beta_2^T\otimes I_{3i}$	$I_1\otimes I_2\otimes T_{3i}^0\beta_{3i}^T$	0	0
(r,o,o_i)	$S_1^0\alpha_1^T\otimes I_2\otimes I_{3i}$	$S_1\otimes I_2\otimes I_{3i}+$ $I_4\otimes T_2\otimes I_{3i}+$ $I_4\otimes I_2\otimes T_{3i}+$ $-\lambda I_4\otimes I_2\otimes I_{3i}$	0	$I_4\otimes T_2^0\otimes I_{3i}+$ $\lambda I_4\otimes e_2\otimes I_{3i}$	0	0	$I_4\otimes I_2\otimes T_{3i}^0$	0
(o,r,o_i)	$I_1\otimes S_2^0\alpha_2^T\otimes I_{3i}$	0	$T_1\otimes I_5\otimes I_{3i}+$ $I_1\otimes S_2\otimes I_{3i}+$ $I_1\otimes I_5\otimes T_{3i}+$ $-\lambda I_1\otimes I_5\otimes I_{3i}$	0	$T_1^0\otimes I_5\otimes I_{3i}+$ $\lambda e_1\otimes I_5\otimes I_{3i}$	0	0	$I_1\otimes I_5\otimes T_{3i}^0$
(r,w,m_i)	0	0	$S_1^0\alpha_1^T\otimes\beta_2^T\otimes I_{3i}^T$	$S_1\otimes I_{3i}$	0	0	0	0
(w,r,m_i)	0	$\beta_1^T\otimes S_2^0\alpha_2^T\otimes I_{3i}^T$	0	0	$S_2\otimes I_{3i}$	0	0	0
(m,m,r_i)	0	0	0	0	0	$I_1\otimes I_2\otimes S_{3i}$	0	0
(r,m,w_i)	0	0	0	0	0	$S_1^0\alpha_1^T\otimes I_2\otimes\beta_{3i}^T$	$S_1\otimes I_2$	0
(m,r,w_i)	0	0	0	0	0	$I_1\otimes S_2^0\alpha_2^T\otimes\beta_{3i}^T$	0	$I_1\otimes S_2$

Fig. 7.5: The matrix Q_{ij}

The effects of state dependent maintenance or overhauls can also be modelled. When, for example, the life time of a component is modelled by a (generalised) Erlang distribution and the phases of the Erlang distribution have a physical meaning, the effect of maintenance can be represented by allowing transitions from degradation stage i to i–1. Mine *et al.* (1974) consider such a state dependent maintenance policy under Markovian deterioration for a two–unit parallel system.

Summarising this chapter, it has been shown how the models in the previous chapters can be extended to model ageing on system level. Recurrence relations have been derived for the (interval) reliability and (joint) availability, exploiting the existence of time epochs which show a lack–of–memory property. Finally, the life and repair time distributions were assumed to be of phase type, which rendered the system state description process Markovian. In this case the randomisation technique can be applied to compute the transient performance measures, which are of particular interest when the system deteriorates.

Epilogue

Techniques for the computation of transient and stationary performance measures of a two-unit repairable dependent parallel system are developed in this monograph, using phase type distributions, the theory of regenerative stochastic processes and the discrete supplementary variable technique. Further, an opportunistic replacement policy of the control limit type is studied and it is shown how to obtain the system's operating characteristics under the replacement policy. Finally, a method is studied for the modelling of deterioration at system level and expressions are derived for the transient performance measures of the deteriorating system.

The conclusion is that in *practical applications*

i. the transient performance measures can be computed relatively simply for bivariate phase type distributed life times and phase type distributed repair times, by using the randomisation technique,

ii. the stationary performance measures can be obtained for bivariate phase type distributed life times and arbitrarily distributed repair times by solving sets of linear equations of Laplace transforms.

Throughout this study it is supposed that the life and repair time distributions are *known*. Moreover, the analysis is simplified considerably when the generators of the phase type distributions are diagonalisable. Hence, further research is recommended into the possibilities of approximating a distribution function with a phase type distribution with the property that its generator is diagonalisable. Subsequently, a sensitivity analysis is needed to investigate the numerical stability of the methods and techniques.
With respect to the replacement policy, further research is needed to improve the computation of the stationary joint pdf and to find out if assumptions about the costs functions can be exploited to minimise the total expected costs per unit of time.
On the other hand, it seems that the *theoretical results* can be extended without difficulties to other system configurations, such as cold or warm standby, intermittently used systems, priority systems or series systems with a small number of components, using the methods and techniques described in this study.

APPENDIX A

When the common cause component in section 5.4 has a PH distribution with representation (α_3,T_3) and dimension n_3, the system state description process $\{X(t),t\geq 0\}$ is the vector valued process $X(t)=(X_1(t),X_2(t),X_3(t))$, where $X_k(t)$ denotes the state of component k (k=1,2,3) at time t. Obviously, $X_k(t)\in\{o,1,..,n_k,r,w\}$, k=1,2 and $X_3(t)\in\{o,1,..,n_3,w\}$, where in the latter case the state 'w' denotes that C_3 is waiting to restart operating, instead of being queued for repair. Let the events $E_{(x_1,x_2,x_3)}$ be defined similarly to the events $E_{(x_1,x_2)}$ in section 5.1 and let $A_{ijk}(t)$ denote the system availability at time t, given an event $E_{(i,j,k)}$ at t=0. Then, considering the mutually exclusive and exhaustive cases described in section 5.4, it is easily seen that, for $(i,j,k)\in\{o,1,..,n_1\}x\{o,1,..,n_2\}x\{o,1,..,n_3\}$

$$A_{ijk}(t) = \bar{F}_{1,i}(t)\ \bar{F}_{3,k}(t) \tag{A.1}$$

$$+\sum_{m=1}^{n_3}\sum_{l=1}^{n_2}\int_0^t f_{1,i}(u)\ P_{2,jl}(u)\ \alpha_{3,k}^T\ \exp(T_3u)_{.m}\ A_{rlm}(t-u)\ du$$

$$+\int_0^t\int_0^u\int_0^v f_{1,i}(v)\ P_{2,jr}(v,w)\ \bar{F}_{3,k}(v)\ \frac{g_2(u-v+w)}{\bar{G}_2(w)}\ A_{roo}(t-u)\ dw\ dv\ du$$

$$+\sum_{l=1}^{n_2}\int_0^t \bar{F}_{1,i}(u)\ P_{2,jl}(u)\ f_{3,k}(u)\ \big(p\ A_{rww}(t-u)\ +(1-p)\ A_{wrw}(t-u)\big)\ du$$

$$+\int_0^t\int_0^u\int_0^v \bar{F}_{1,i}(v)\ P_{2,jr}(v,w)\ f_{3,k}(v)\ \frac{g_2(u-v+w)}{\bar{G}_2(w)}\ A_{roo}(t-u)\ dw\ dv\ du,$$

where

$$A_{wrw}(t) = \int_0^t g_2(u)\ A_{roo}(t-u)\ du,$$

and

$$A_{rww}(t) = \int_0^t g_1(u)\ A_{oro}(t-u)\ du.$$

An expression for the availability function, conditioned on an $E_{(i,r)}$ event at the time origin is, for $(i,k)\in\{o,1,..,n_1\}\times\{o,1,..,n_3\}$

$$A_{irk}(t) = \bar{F}_{1,i}(t)\ \bar{F}_{3,k}(t)\ \bar{G}_2(t) \tag{A.2}$$

$$+\sum_{m=1}^{n_3}\sum_{l=1}^{n_1}\int_0^t g_2(u)\ \alpha_{1,i}^T \exp(T_1u)_{.l}\ \alpha_{3,k}^T \exp(T_3u)_{.m}\ A_{lom}(t-u)\ du$$

$$+\int_0^t\int_0^u f_{1,i}(v)\ \bar{F}_{3,k}(v)\ dv\ g_2(u)\ A_{roo}(t-u)\ du$$

$$+\int_0^t\int_0^u f_{3,k}(v)\ \bar{F}_{1,i}(v)\ dv\ g_2(u)\ A_{roo}(t-u)\ du.$$

Interchanging the roles of C_1 and C_2 yields an expression for $A_{rjk}(t)$, $(j,k)\in\{o,1,..,n_2\}\times\{o,1,..,n_3\}$.

To obtain the limiting availability, the Laplace transforms of (A.1) and (A.2) are computed. The result is a set of linear equations in the Laplace transforms of the availability functions $A_{ijk}(t)$:

$$A_{ijk}^*(s) = \mathfrak{L}\Big[\bar{F}_{1,i}(t)\ \bar{F}_{3,k}(t)\Big] \tag{A.3}$$

$$+\sum_{m=1}^{n_3}\sum_{l=1}^{n_2}\mathfrak{L}\Big[f_{1,i}(t)\ P_{2,jl}(t)\ \alpha_{3,k}^T \exp(T_3t)_{.m}\Big]\ A_{rlm}^*(s)$$

$$+\mathfrak{L}\Big[\int_0^t\int_0^v f_{1,i}(v)\ P_{2,jr}(v,w)\ \bar{F}_{3,k}(v)\ \frac{g_2(t-v+w)}{\bar{G}_2(w)}\ dw\ dv\Big]\ A_{roo}^*(s)$$

$$+\sum_{l=1}^{n_2}\mathfrak{L}\Big[\bar{F}_{1,i}(t)\ P_{2,jl}(t)\ f_{3,k}(u)\Big]\ \big[pA_{rww}^*(s) + (1-p)A_{wrw}^*(s)\big]$$

$$+\mathfrak{L}\Big[\int_0^t\int_0^v \bar{F}_{1,i}(v)\ P_{2,jr}(v,w)\ f_{3,k}(v)\ \frac{g_2(t-v+w)}{\bar{G}_2(w)}\ dw\ dv\Big]\ A_{roo}^*(s),$$

where

$$A_{wrw}^*(s) = g_2^*(s)\ A_{roo}^*(s)$$

and

$$A_{rww}^*(s) = g_1^*(s)\ A_{oro}^*(s).$$

Subsequently,

$$A^*_{irk}(s) = \mathfrak{L}\left[\bar{F}_{1,i}(t)\ \bar{F}_{3,k}(t)\ \bar{G}_2(t)\right] \tag{A.4}$$

$$+ \sum_{m=1}^{n_3} \sum_{l=1}^{n_1} \mathfrak{L}\left[g_2(t)\ \alpha^T_{1,i}\ \exp(T_1t)_{.l}\ \alpha^T_{3,k}\ \exp(T_3u)_{.m}\right]\ A^*_{lom}(s)$$

$$+ \mathfrak{L}\left[g_2(t) \int_{0^-}^{t} f_{1,i}(u)\ \bar{F}_{3,k}(u)\ du\right]\ A^*_{roo}(s)$$

$$+ \mathfrak{L}\left[g_2(t) \int_{0^-}^{t} f_{3,k}(u)\ \bar{F}_{1,i}(u)\ du\right]\ A^*_{roo}(s).$$

As in section 5.2, the coefficients in (A.3) and (A.4) can be determined explicitly when the matrices T_k (k=1,2,3) are diagonalisable. Computing the Laplace transforms, the following property is used (Neuts, 1981, theorem 2.2.9): let $\bar{F}_{1,i}(t) = \alpha_1^T \exp(T_1t)\ e_1$ and $\bar{F}_{3,k}(t) = \alpha_3^T \exp(T_3t)\ e_3$, then

$$\bar{F}_{1,i}(t)\ \bar{F}_{3,k}(t) = \alpha^T \exp(Tt)\ e, \tag{A.5}$$

where $\alpha=\alpha_1\otimes\alpha_3$, $T=T_1\otimes I_3+I_1\otimes T_3$ and $e=e_1\otimes e_3$. Neuts also proves that T^{-1} exists.

Let α, T and e be defined as in (A.5) and suppose that T is diagonalisable, say $T=S^{-1}DS$. Further, let I be an identity matrix of dimension n_1n_3 and note that $\alpha^T_{3,k} \exp(T_3t)_{.m} = \alpha^T_{3,k} \exp(T_3t)\ \alpha_{3,m}$. Then it is easily verified that the Laplace transforms of the coefficients in the right hand side of (A.3) are given by

$$\mathfrak{L}\left[\bar{F}_{1,i}(t)\ \bar{F}_{3,k}(t)\right] = \alpha^T\ (sI-T)^{-1}\ e,$$

$$\mathfrak{L}\left[f_{1,i}(t)\ P_{2,jl}(t)\ \alpha^T_{3,k}\ \exp(T_3t)_{.m}\right] = \alpha^T\ S^{-1}\ P^*_{2,jl}(sI-D)\ S\ T_1^0\otimes\alpha_{3,m},$$

$$\mathfrak{L}\left[\bar{F}_{1,i}(t)\ P_{2,jl}(t)\ f_{3,k}(t)\right] = \alpha^T\ S^{-1}\ P^*_{2,jl}(sI-D)\ S\ e_1\otimes T_3^0,$$

$$\mathfrak{L}\left[\int_0^t \int_0^v f_{1,i}(v)\ P_{2,jr}(v,w)\ \bar{F}_{3,k}(v)\ \frac{g_2(t-v+w)}{\bar{G}_2(w)}\ dw\ dv\right]$$

$$= \alpha^T\ T^{-1}\ S^{-1}\ \mathfrak{P}^*_{2,jr}(sI-D)\ \left[g_2^*(sI-D)\ -\ g_2^*(sI)\right]\ S\ T_1^0 \otimes e_3$$

and

$$\mathfrak{L}\left[\int_0^t \int_0^v \bar{F}_{1,i}(v)\ P_{2,jr}(v,w)\ f_{3,k}(v)\ \frac{g_2(t-v+w)}{\bar{G}_2(w)}\ dw\ dv\right]$$

$$= \alpha^T\ T^{-1}\ S^{-1}\ \mathfrak{P}^*_{2,jr}(sI-D)\ \left[g_2^*(sI-D)\ -\ g_2^*(sI)\right]\ S\ e_1 \otimes T_3^0.$$

The Laplace transforms of the coefficients in the right hand side of (A.4) are

$$\mathfrak{L}\left[\bar{F}_{1,i}(t)\ \bar{F}_{3,k}(t)\ \bar{G}_2(t)\right]\ =\ \alpha^T\ S^{-1}\ \bar{G}_2^*(sI-D)\ S\ e,$$

$$\mathfrak{L}\left[g_2(t)\ \alpha_{1,i}^T\ \exp(T_1t)_{.l}\ \alpha_{3,k}^T\ \exp(T_3t)_{.m}\right]\ =\ \alpha^T\ S^{-1}\ g_2^*(sI-D)\ S\ \alpha_{1,l}^T \otimes \alpha_{3,m},$$

$$\mathfrak{L}\left[g_2(t)\int_0^t f_{1,i}(u)\ \bar{F}_{3,k}(u)\ du\right]\ =\ \alpha^T\ T^{-1}\ S^{-1}\ \left[g_2^*(sI-D)\ -\ g_2^*(sI)\right]\ S\ T_1^0 \otimes e_3$$

and

$$\mathfrak{L}\left[g_2(t)\int_0^t f_{3,k}(u)\ \bar{F}_{1,i}(u)\ du\right]\ =\ \alpha^T\ T^{-1}\ S^{-1}\ \left[g_2^*(sI-D)\ -\ g_2^*(sI)\right]\ S\ e_1 \otimes T_3^0.$$

Finally, the limiting availability can be obtained as in section 5.4, applying the Abelian theorem for Laplace transforms (Cohen, 1982).

APPENDIX B

To show that $M^*(s)$ in (5.28) is irreducible for $s \geq 0$, note that it follows from (5.29) that for real $s \geq 0$

$$\sum_{j=1}^{n} M^*_{ij}(s) \leq 1,$$

since either $M^*_{ij}(s)$ is strictly decreasing in s or $M^*_{ij}(s)=0$ for all $s \geq 0$. Hence $M^*(s)$ is a (sub)stochastic matrix for $s \geq 0$ (since $\forall_{s\geq 0} \forall_{i,j\in\{1,..,n\}} M^*_{ij}(s) \geq 0$) and can be considered as a matrix of transition probabilities. Obviously, the entries $M^*_{ij}(s)$ are positive if and only if the corresponding transition from state i to state j can occur. However, the assumed irreducibility of the representations (α_i, T_i) of the phase type life time distributions of the components C_i (i=1,2) guarantees that all n states communicate. Hence, for all $i,j \in \{1,..,n\}$ there exists an integer k_{ij} such that

$$\left[M^*(s)\right]^{k_{ij}}_{ij} > 0.$$

In other words: $M^*(s)$ is irreducible. ■

To prove that $\mu_1(s)$, the eigenvalue with maximum real part of $M^*(s)$, is differentiable in s, define the function $f(\mu,s)$ as

$$f(\mu,s) := \det(\mu I - M^*(s)).$$

Let $\tilde{s} \geq 0$, then $f(\mu,\tilde{s})$ is an analytic function in μ, since $f(\mu,\tilde{s})$ is a polynomial in μ. Now, let C be a closed contour in the complex μ-plane, such that

i. $f(\mu,\tilde{s})$ has no roots *on* the contour C and

ii. there is exactly one root of $f(\mu,\tilde{s})$ *inside* the contour C, *viz.* the root with maximum real part of $f(\mu,\tilde{s})$.

Then the root inside C is $\mu_1(\tilde{s})$ and it follows from the Perron-Frobenius theorem that $\mu_1(\tilde{s})$ is a simple root. Moreover, since $\mu_1(\tilde{s})$ is simple, Titchmarsh (1952, p.116) shows that

$$\frac{1}{2\pi i} \oint_C \frac{\mu\, f_\mu(\mu,\tilde{s})}{f(\mu,\tilde{s})}\, d\mu = \mu_1(\tilde{s}), \tag{B.1}$$

where $f_\mu(\mu,\tilde{s}) = \partial f(\mu,\tilde{s})/\partial\mu$.

To show that the left hand side of (B.1) is differentiable, note that the entries of the matrix $M^*(s)$ are all differentiable functions. Hence, $f(\mu,s)$ and $f_\mu(\mu,s)$ are differentiable as a function of s. Finally, since $f(\mu,s)\neq 0$ on the contour C, it follows that the integrand in the left hand side of (B.1) is differentiable as a function of s. Consequently, $\mu_1(s)$ is differentiable. ■

APPENDIX C

To prove that $\mathfrak{M}^*(s)$ in (5.67) is irreducible for all real $s>Re(\sigma)$, where σ is the pole with mrp of $\mathfrak{M}^*(s)$, note that

i. $\forall_{s>Re(\sigma)}\left[\mathfrak{M}^*_{12}(s) > 0\right]$ as the assumed irreducibility of the representation

(α_2, T_2) guarantees that $\forall_{t>0}\ \forall_{i,j\in\{1,..,n_2\}}\left[f_{1,0}(t)\ P_{2,ij}(t) > 0\right]$,

ii. $\forall_{s>Re(\sigma)}\left[\mathfrak{M}^*_{21}(s) \geq 0\right]$ and

iii. $\forall_{s>Re(\sigma)}\ \forall_{i\in\{1..n_2\}}\left[(\mathfrak{M}^*_{21}(s))_{ii}>0\right]$, as $\forall_{t\geq 0}\ \forall_{i\in\{1..n_2\}}\left[g_1(t)\exp(T_2t)_{ii}>0\right]$.

Hence $\forall_{s>Re(\sigma)}\left[\mathfrak{M}^*_{12}(s)\ \mathfrak{M}^*_{21}(s) > 0\right]$ and $\forall_{s>Re(\sigma)}\left[\mathfrak{M}^*_{21}(s)\ \mathfrak{M}^*_{12}(s) > 0\right]$.

Further, remark that, for $k\in\mathbb{N}$

$$\mathfrak{M}^*(s)^{2k} = \left[\begin{array}{c|c} (\mathfrak{M}^*_{12}(s)\mathfrak{M}^*_{21}(s))^k & O \\ \hline O & (\mathfrak{M}^*_{21}(s)\mathfrak{M}^*_{12}(s))^k \end{array}\right]$$

and

$$\mathfrak{M}^*(s)^{2k+1} = \left[\begin{array}{c|c} O & \mathfrak{M}^*_{12}(s)\ (\mathfrak{M}^*_{21}(s)\mathfrak{M}^*_{12}(s))^k \\ \hline \mathfrak{M}^*_{21}(s)\ (\mathfrak{M}^*_{12}(s)\mathfrak{M}^*_{21}(s))^k & O \end{array}\right]$$

and therefore $\mathfrak{M}^*(s)$ is irreducible for all $s>Re(\sigma)$. ∎

The following lemma is used to prove that the pole δ with mrp of $r^*(s)$ in (5.67) is a pole of $(I-\mathfrak{M}^*(s))^{-1}$.

Lemma:

i. If $\lambda(s)$ is eigenvalue of $\mathfrak{M}^*(s)$, then $\lambda^2(s)$ is eigenvalue of $\mathfrak{M}^*_{12}(s)\mathfrak{M}^*_{21}(s)$.

ii. If $\lambda(s)$ is eigenvalue of $\mathfrak{M}^*(s)$, then $-\lambda(s)$ is also eigenvalue of $\mathfrak{M}^*(s)$.

Proof: *i.* Suppose $\mathfrak{M}^*(s)\ x(s) = \lambda(s)\ x(s)$. Partition $\mathfrak{M}^*(s)$ and $x(s)$ and write (for ease of use the argument s is deleted)

$$\left[\begin{array}{c|c} \lambda\ I_{n_2} & -\mathfrak{M}^*_{12} \\ \hline -\mathfrak{M}^*_{21} & \lambda\ I_{n_2} \end{array}\right] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 0.$$

Thus $\mathfrak{M}^*_{12}x_2 = \lambda x_1$ and $\mathfrak{M}^*_{21}x_1 = \lambda x_2$, hence $\mathfrak{M}^*_{12}\mathfrak{M}^*_{21}x_1 = \lambda\mathfrak{M}^*_{12}x_2 = \lambda^2 x_1$ and it follows that $\det(\lambda^2 I_{n_2} - \mathfrak{M}^*_{12}\mathfrak{M}^*_{21}) = 0$.

ii. For a matrix M,

$$M = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix},$$

with invertible submatrix M_{11}, the Schur-complement (Fiedler, 1986) of M in M_{11} is defined as $M_{22}-M_{21}M_{11}^{-1}M_{12}$ and $\det(M) = \det(M_{11})\ \det(M_{22}-M_{21}M_{11}^{-1}M_{12})$.
Hence, $\det(\lambda I_{2n_2}-\mathfrak{M}^*) = \lambda^{n_2}\ \det(\lambda I_{n_2}-\lambda^{-1}\mathfrak{M}^*_{12}\mathfrak{M}^*_{21}) = \det(\lambda^2 I_{n_2}-\mathfrak{M}^*_{12}\mathfrak{M}^*_{21})$. In other words: $\det(\lambda I_{2n_2}-\mathfrak{M}^*)$ is an even function in λ and consequently $\det(-\lambda I_{2n_2}-\mathfrak{M}^*)=0$ whenever $\det(\lambda I_{2n_2}-\mathfrak{M}^*)=0$. ■

To show that the pole δ with mrp of $r^*(s)$ in (5.67) is a pole of $(I-\mathfrak{M}^*(s))^{-1}$, consider the candidates for δ. As $r^*(s) = (I-\mathfrak{M}^*(s))^{-1}b^*_r(s)$, these candidates are the poles of $\mathfrak{M}^*(s)$, $b^*_r(s)$ and $(I-\mathfrak{M}^*(s))^{-1}$.
With respect to the repair time distribution it is assumed that σ_i is the pole with mrp of $g^*_i(s)$. Further, suppose that the matrices T_i, i=1,2, are diagonalisable: let $T_i=S_i^{-1}D_iS_i$, where $D_i=\mathrm{diag}(d_{i,1},..,d_{i,n_i})$ and $d_{i,1}$ is the eigenvalue with mrp.
Then $\mathfrak{M}^*_{21}(s) = S_2^{-1}\ g^*_1(sI_2-D_2)\ S_2$ and it follows that the pole with mrp of $\mathfrak{M}^*_{21}(s)$ is $s=\sigma_1+d_{2,1}$.
Subsequently,

$$(\mathfrak{M}^*_{12}(s))_{ij} = \alpha_1\ S_1^{-1}\ P^*_{2,ij}(sI_1-D_1)\ S_1\ T_1^0$$

and

$$P^*_{2,ij}(s) = ((sI_2-T_2)^{-1})_{ij} + f^*_{2,i}(s)\ g^*_2(s)\ P^*_{2,oj}(s).$$

Careful inspection of $P^*_{2,ij}(s)$ learns, after substitution of $f^*_{2,i}(s)$, $g^*_2(s)$ and $P^*_{2,oj}(s)$, that the pole with mrp of $\mathfrak{M}^*_{12}(s)$ is $s=d_{1,1}$. As a result $s_{max}=\max\{d_{1,1},d_{2,1}+\sigma_1\}$ is the pole with mrp of $\mathfrak{M}^*(s)$.

Secondly, as

$$b_r^*(s) = [\mathfrak{L}(\bar{F}_{1,0}(t)) \ldots \mathfrak{L}(\bar{F}_{1,0}(t))\ \mathfrak{L}(\bar{F}_{2,1}(t)\ \bar{G}_1(t)) \ldots \mathfrak{L}(\bar{F}_{2,n_2}(t)\ \bar{G}_1(t))]^T,$$

$$\mathfrak{L}(\bar{F}_{1,0}(t)) = \alpha_1^T\ (sI_1-T_1)^{-1}\ e_1$$

and

$$\mathfrak{L}(\bar{F}_{2,1}(t)\ \bar{G}_1(t)) = \alpha_{2,j}^T\ S_2^{-1}\ \bar{G}_1^*(sI_2-D_2)\ S_2\ e_2,$$

it follows that s_{max} is also pole with mrp of $b_r^*(s)$.

Finally, in order to prove that $(I-\mathfrak{M}^*(s))^{-1}$ has a pole s_{mrp} which satisfies $Re(s_{mrp})>s_{max}$, the eigenvalues $\mathfrak{m}_i(s)$ $(i=1,..,2n_2)$ of $\mathfrak{M}^*(s)$ are considered. Let $\mathfrak{m}_1(s)$ be the eigenvalue of $\mathfrak{M}^*(s)$ with mrp. It will be shown that $\exists_{s\in(s_{max},0]}\ [\mathfrak{m}_1(s)=1]$, *i.e.* $\exists_{s\in(s_{max},0]}\ [\det(I-\mathfrak{M}^*(s))=0]$.

To prove that $\exists_{s\in(s_{max},0]}\ [\mathfrak{m}_1(s)=1]$, note that

$$\forall_{i,j\in\{1,..,n_2\}} \left[\lim_{s\downarrow s_{max}} (\mathfrak{M}_{12}^*(s)\mathfrak{M}_{21}^*(s))_{ij} = \infty\right].$$

Now, using the theorem that the spur of a matrix equals the sum of its eigenvalues (Marcus *et al.*, 1964, p.23), it follows from the above lemma that

$$\lim_{s\downarrow s_{max}} \mathfrak{m}_1(s) = \infty.$$

On the other hand, it is shown in section 5.2 that $\mathfrak{M}^*(s)$ is (sub)stochastic for $s\geq 0$ and hence $\mathfrak{m}_1(s)\leq 1$ for $s\geq 0$. However, replacing $M^*(s)$ by $\mathfrak{M}^*(s)$ in appendix B, it is immediate that $\mathfrak{m}_1(s)$ is continuous in s and hence $\exists_{s\in(s_{max},0]}\ [\mathfrak{m}_1(s)=1]$.

As a result, the pole δ with mrp of $r^*(s)$ is a pole of $(I-\mathfrak{M}^*(s))^{-1}$, $\delta\in(\max\{d_{1,1},d_{2,1}+\sigma_1\},0]$ and obviously, $\delta = \max_{s\in\mathbb{R}} \{\mathfrak{m}_1(s)=1\}$. ∎

Appendix D

The numerical examples in section 6.8 illustrate the OFRP for a system with identical units and one repair facility. In this case, it follows from (6.52) and (6.53) that, for $0<x<L_1$

$$H_1(x) = \bar{F}_3(x/2) \int_0^{x^-} g_2(y)\, H_2(x-y)\, dy + \bar{F}_3(x/2)\, g_2(x)\, H_2$$

and

$$H_2(x) = \int_0^{x^-} f_2(y)\, \bar{F}_3((y-x)/2)\, H_1(x-y)\, dy + \int_{x^+}^{x+L_1} f_2(y)\, \bar{F}_3((y-x)/2)\, H_1(y-x)\, dy.$$

For given H_2, the functions $H_1(x)$ and $H_2(x)$ can be obtained numerically by discretisation of the above equations. To illustrate this, let $n\in\mathbb{N}$, $h=L_1/n$, $x_i=ih$ $(i=1,..,n)$ and $x_0=0^+$. Then discretisation of the integrals yields, for $i=1,..,n$

$$H_1(x_i) = \bar{F}_3(x_i/2)\, h\left\{ \tfrac{1}{2} g_2(x_0) H_2(x_i) + \sum_{j=1}^{i-1} g_2(x_j) H_2(x_i-x_j) + \tfrac{1}{2} g_2(x_i) H_2(x_0)\right\} \qquad \text{(D.1)}$$

$$+ \bar{F}_3(x_i/2)\, g_2(x_i)\, H_2$$

and

$$H_2(x_i) = \qquad \text{(D.2)}$$

$$h\left\{ \tfrac{1}{2} f_2(x_0)\, \bar{F}_3(-x_i/2)\, H_1(x_i) + \sum_{j=1}^{i-1} f_2(x_j)\, \bar{F}_3((x_j-x_i)/2)\, H_1(x_i-x_j) + \tfrac{1}{2} f_2(x_i)\, H_1(x_0)\right\}$$

$$+ h\left\{ \tfrac{1}{2} f_2(x_i)\, H_1(x_0) + \sum_{j=1}^{n-1} f_2(x_i+x_j)\, \bar{F}_3(x_j/2)\, H_1(x_j) + \tfrac{1}{2} f_2(x_i+x_n)\, \bar{F}_3(x_n/2)\, H_1(x_n)\right\}$$

Further, it is easily seen that $H_1(x_0) = g_2(x_0)\, H_2$ and that $H_2(x_0)$ equals the second part of the right hand side of (D.2). Hence, discretisation results in a set of $2(n+1)$ linear equations in $2(n+1)$ unknown variables, *viz.* $H_1(x_i)$ and $H_2(x_i)$, $i=0,..,n$.

In the numerical examples in section 6.8 n equals 50. With respect to the infinite integrals in (6.54)–(6.56), the following choices are made:

i. The first integral in (6.55) is truncated at $\mu_R+5\sigma_R$, where μ_R and σ_R are as in section 6.8.

ii. The third and fourth integral in (6.55) are truncated at $\max\{3L_2,\mu_L+3\sigma_L\}$, where μ_L and σ_L are as in section 6.8.

iii. The first integral in (6.54) is truncated at $\max\{2x,\mu_L+3\sigma_L\}$.

iv. The integral in (6.56) is truncated at $\mu_L+3\sigma_L$.

The integrals are computed by Simpson's $\frac{1}{3}$–rule. From numerical experiments with identically exponentially distributed life times, respectively repair times, it appeared that the error in the stationary joint pdf was less than 0.5 % with the above choices.

REFERENCES

Aranda-Ordaz, F.J. (1980). *Transformations to additivity for binary data.* Unpublished Ph.D. thesis, Imperial College, London.

Arsham, H., Balana, A.R. & Gross, D. (1983). Numerical methods for transient solutions of machine repair problems. *Computers and Industrial Engineering*, **7**, pp. 149–157.

Ascher, H.E. (1983). Discussion of Lawless. *Technometrics*, **25**, 1983, pp. 320–326.

Ascher, H.E. & Feingold H. (1978). Is there repair after failure? *Proceedings* 1978 *annual reliability and maintainability symposium*, pp. 190–197.

Ascher, H.E. & Feingold H. (1984). *Repairable systems reliability: Modelling, inference, misconceptions and their causes.* Marcel Dekker, New York.

Bansard, J.P., Descamps, J.L., Maarek, G. & Morihain, G. (1969). Study of stochastic renewal process applied to a trigger-off policy for replacing whole sets of components. *Proceedings of the 5th International Conference on Operational Research*, 1969, ed. L. Lawrence, pp. 235–264 (Tavistock, London 1970).

Barlow, R.E. & Proschan, F. (1975). *Statistical theory of reliability and life testing.* Holt, Reinhart and Winston, New York.

Berg, M. (1978), General trigger-off replacement procedures for two-unit systems. *Naval Research Logistics Quarterly*, **25**, pp. 15–29.

Birolini, A. (1985). *On the use of stochastic processes in modeling reliability problems.* Springer, Berlin.

Blumenthal, S.B., Greenwood, J.A. & Herbach, L.H. (1971). Superimposed nonstationary renewal processes. *Journal of Applied Probability*, **8**, pp. 184–192.

Blumenthal, S.B., Greenwood, J.A. & Herbach, L.H. (1973). The transient reliability behavior of series systems of superimposed renewal processes. *Technometrics*, **15**, pp. 255–269.

Box, G.E. & Jenkins, G.M. (1970). *Time series analysis: forecasting and control.* Holden-Day, San Francisco.

Bux, W. & Herzog, U. (1977). The phase concept: Approximation of measured data and performance analysis. *Computer performance: proceedings of the international symposium on computer performance, modelling, measurement and evaluation*, Yorktown Heights, New York, August 16–18, 1977.

Cavender, J.A. (1978). Quasi-stationary distributions of birth-and-death processes. *Advances in Applied Probability*, **10**, pp. 570–586.

Çinlar, E. (1975). *Introduction to stochastic processes.* Prentice-Hall,London.

Cohen, J.W. (1982). *The single server queue* (revised edition). North-Holland Publishing Company, Amsterdam.

Cox, D.R. (1955[a]). The analysis of non-Markovian stochastic processes by the inclusion of supplementary variables. *Proceedings of the Cambridge Philosophical Society*, **51**, pp. 433–441.

Cox, D.R. (1955[b]). A use of complex probabilities in the theory of stochastic processes. *Proceedings of the Cambridge Philosophical Society*, **51**, pp. 313–319.

Cox, D.R. (1972). Regression models and life tables (with discussion). *Journal of the Royal Statistical Society, series* B, **34**, pp. 187–220.

Cox, D.R. & Miller, H.D. (1965). *The theory of stochastic processes.* Chapman and Hall, London.

Cox, D.R. & Lewis, P.A.W. (1966). *The statistical analysis of series of events.* Methuen, London.

Cox, D.R. & Isham, V. (1980). *Point processes.* Chapman & Hall, London.

Cox, D.R. & Oakes, D. (1984). *Analysis of survival data.* Chapman and Hall, London.

Darroch, J.N. & Seneta, E. (1965). On quasi-stationary distributions in absorbing discrete-time finite Markov chains. *Journal of Applied Probability*, **2**, pp. 88–100.

Darroch, J.N. & Seneta, E. (1967). On quasi-stationary distributions in absorbing continuous-time finite Markov chains. *Journal of Applied Probability*, **4**, pp. 192–196.

David, H.A. & Moeschberger M.L. (1978). *The theory of competing risks.* Griffin, London.

Drenick, R.F. (1960). The failure law of complex equipment. *Journal of the Society for Industrial and Applied Mathematics*, **8**, pp. 680–690.

Duyn-Schouten, F.A. van der & Ronner, T. (1989). Calculation of the availability of a two-unit parallel system with cold standby. *Probability in the Engineering and Informational Sciences*, **3**, pp. 341–353.

Duyn-Schouten, F.A. van der & Vanneste, S.G. (1990). Analysis and computation of (n,N)-strategies for maintenance of a two-component system. *European Journal of Operational Research*, **48**, pp. 260–274.

Elandt-Johnson, R.C. (1980). Some prior and posterior distributions in survival analysis : A critical insight on relationships derived from cross-sectional data. *Journal of the Royal Statistical Society, series* B, **42**, pp. 96–106.

Feller, W. (1966). *An introduction to probability theory and its applications.* Vol. II. John Wiley, New York.

Fiedler, M. (1986). *Special matrices and their applications in numerical mathematics.* Martinus Nijhoff Publishers, Dordrecht, The Netherlands.

Finkbeiner, D.T. II (1978). *Introduction to matrices and linear transformations* (3rd edition). Freeman, San Francisco.

Gaver, D.P. (1963). Time to failure and availability of parallel systems with repair. *IEEE Transactions on Reliability*, **R-12**, pp. 30–38.

Gaver, D.P. (1964). Failure time for a redundant repairable system of two dissimilar elements. *IEEE Transactions on Reliability*, **R-13**, pp. 14–22.

Gerald, C.F. (1978). *Applied numerical analysis (Second edition).* Addison-Wesley, Reading, Massachusetts, U.S.A.

Grassmann, W.K. (1977). Transient solutions in Markovian queueing systems. *Computers & Operations Research*, **4**, pp. 47–53.

Grigelionis, B.I. (1964). Limit theorems for sums of renewal processes. *Cybernetics in the Service of Communism, Vol.* **2**, *Reliability Theory and Queueing Theory*. Editors: A.I. Berg, N.G. Bruevich & B.V. Gnedenko, "Energy" Publishing House, Moscow–Leningrad, pp. 246–266 (in Russian).

Gross, D. & Miller, D.R. (1984). The randomization technique as a modeling tool and solution procedure for transient Markov processes. *Operations Research*, **32**, pp. 343–361.

Harris, R. (1968). Reliability applications of a bivariate exponential distribution. *Operations Research*, **16**, pp. 18–27.

Johnson, M.A. & Taaffe, M.R. (1989). Matching moments to phase distributions: Mixtures of Erlang distributions of common order. *Communications in Statistics – Stochastic Models*, **5**, pp. 711–743.

Johnson, M.A. & Taaffe, M.R. (1990a). Matching moments to phase distributions: Nonlinear programming approaches. *Communications in Statistics–Stochastic Models*, **6**, pp. 259–281.

Johnson, M.A. & Taaffe, M.R. (1990b) Matching moments to phase distributions: Density function shapes. *Communications in Statistics–Stochastic Models*, **6**, pp. 283–306.

Kalpakam S. & Shahul Hameed M.A. (1983). Quasi–stationary distribution of a two–unit warm standby redundant system. *Journal of Applied Probability*, **20**, pp. 429–435.

Kao, E.P.C. (1988). Computing the phase–type renewal and related functions. *Technometrics*, **30**, pp. 87–93.

Kodama M. & Deguchi H. (1974). Reliability considerations for a two–unit redundant system with Erlang failure and repair distributions. *IEEE Transactions on Reliability*, **R–23**, pp. 75–81.

Lawless, J.F. (1982). *Statistical models and methods for lifetime data*. John Wiley, New York.

Lewis, P.A.W. (1964). A branching Poisson process model for the analysis of computer failure patterns. *Journal of the Royal Statistical Society, series* B, **26**, pp. 398–456.

Lewis, E.E. (1987). *Introduction to reliability engineering*. John Wiley, New York.

Liebowitz, B.H. (1966). Reliability considerations for a two–element redundant system with generalised repair times. *Operations Research*, 14, pp. 233–241.

Linton, J.G. (1976). Some advancements in the analysis of a two–unit parallel redundant system. *Microelectronics & Reliability*, 16, pp. 39–46.

Marcus, M. & Minc, H. (1964). *A survey of matrix theory and matrix inequalities*. Allyn and Bacon, Boston.

Marshall, A.W. & Olkin, I. (1967). A multivariate exponential distribution. *Journal of the American Statistical Association*, **62**, pp. 30–44.

McCall,J.J. (1965). Maintenance policies for stochastically failing equipment: A Survey. *Management Science*, **11**, pp. 493–524.

Mine, H. (1988). Studies on reliability and maintainability. Ed. by H. Mine, Kansai University.

Mine, H. & Kawai, H. (1974). An optimal maintenance policy for a 2–unit parallel system with degraded states. *IEEE Transactions on Reliability*, **R–23**, pp. 81–86.

Neuts, M.F. (1981). *Matrix–geometric solutions in stochastic models*. John Hopkins University Press.

Neuts, M.F. & Meier, K.S. (1981). On the use of phase type distributions in reliability modelling of systems with two components. *OR Spectrum*, **2**, pp. 227–234.

Ohashi, M., Huang, J.M. & Nishida, T. (1980). Stochastic behaviour of a two–unit parallel system. *Microelectronics & Reliability*, **20**, pp. 471–476.

Ohashi, M. & Nishida, T. (1981). Optimal replacement policy for two–unit system, *Journal of the Operations Research Society of Japan*, **24**, pp. 283–294.

Osaki, S. (1980). A two–unit parallel redundant system with bivariate exponential lifetimes. *Microelectronics & Reliability*, **20**, pp. 521–523.

Osaki, S. (1985). *Stochastic system reliability modelling*. World Scientific Publishing Company, Singapore.

Osaki, S. & Nakagawa, T. (1976). Bibliography of reliability and availability of stochastic systems. *IEEE Transactions on Reliability*, **R–25**, pp. 284–287.

Parzen, E. (1962). *Stochastic processes*. Holden–Day, London.

Pierskalla, W.P. & Voelker, J.A. (1976). A survey of maintenance models: The control and surveillance of deteriorating systems. *Naval Research Logistics Quarterly*, **23**, pp. 353–388 (1976).

Pijnenburg, M.J.P. (1991). Additive hazards models in repairable systems reliability. *Reliability Engineering and System Safety*, 31, pp. 369–390.

Pijnenburg, M.J.P, Ravichandran, N. (1991). Quasi–stationary distributions of a two–unit dependent parallel system. *Opsearch*, **28**, pp. 264–281.

Pijnenburg, M.J.P., Ravichandran, N. & Regterschot, G.J.K. Stochastic analysis of a dependent parallel system. Accepted by *European Journal of Operational Research*.

Pollak, M. & Siegmund, D.(1986). Convergence of quasi–stationary to stationary distributions for stochastic monotone Markov Processes. *Journal of Applied Probability*, **23**, pp. 215–220.

Prentice, R.L., Williams, B.J. & Peterson, A.V. (1981). On the regression analysis of multivariate failure time data. *Biometrika*, **68**, pp. 373–379.

Ravichandran, N. (1981). Analysis of a two–unit parallel redundant system with phase type failure and general repair. *Microelectronics & Reliability*, **21**, pp. 567–572.

Ravichandran, N. (1990). *Stochastic methods in reliability theory*. Wiley Eastern Limited.

Ross, S.M. (1970). *Applied probability models with optimization applications.* Holden-Day, San Francisco.

Scales, L.E. (1985). *Introduction to non-linear optimization.* Macmillan, London.

Seneta, E. (1973). *Non-negative matrices: an introduction to theory and applications.* Allen & Unwin, London.

Seneta, E. & Vere-Jones, D. (1967) On quasi-stationary distributions in discrete time Markov chains with a denumerable infinity of states. *Journal of Applied Probability*, **3**, pp. 403–434.

Seneta, E. & Tweedie, R.L.(1985). Moments for stationary and quasi-stationary distributions of Markov chains. *Journal of Applied Probability*, **22**, pp. 148–155.

Sethi, D.P.S. (1977). Opportunistic replacement policies. In: *The theory and applications of reliability*, ed. by C.P. Tsokos and I.N. Shimi.

Sherif, Y.S. (1982). Reliability analysis: Optimum inspection and maintenance schedules of failing systems. *Microelectronics & Reliability*, **22**, pp. 59–115.

Sherif, Y.S. & Smith, M.L. (1981). Optimum maintenance models for systems subject to failure – A review. *Naval Research Logistics Quarterly*, **28**, pp. 47–74.

Singpurwalla, N.D. (1978). Estimating reliability growth (or deterioration) using time series analysis. *Naval Research Logistics Quarterly*, **25**, pp. 1–14

Smith, W.L. (1955). Regenerative stochastic processes. *Proceedings of the Royal Society* A, **232**, pp. 6–31.

Smith, W.L. (1958). Renewal theory and its ramifications. *Journal of the Royal Statistical Society*, Series B, **20**, pp. 243–302.

Souza e Silva, E. de & Gail, H.R. (1986). Calculating cumulative operational time distributions of repairable computer systems. *IEEE Transactions on Computers*, **C-35**, pp. 322–332.

Srinivasan, S.K. (1974). *Stochastic Point Processes.* Griffin, London.

Subramanian, R. & Ravichandran, N. (1979). Stochastic behavior of a 2-unit parallel redundant system. *IEEE Transactions on Reliability*, **R-28**, pp. 419–420.

Thomas, L.C. (1986). A survey of maintenance and replacement models for maintainability and reliability of multi-item systems. *Reliability Engineering*, **16**, pp. 297–309.

Thompson Jr., W.A. (1981). On the foundations of reliability. *Technometrics*, **23**, pp. 1–13.

Thompson Jr., W.A. (1988). *Point process models with applications to safety and reliability.* Chapman & Hall, London.

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

Titchmarsh, E.C. (1952). *The theory of functions* (second edition). Oxford University Press, London.

Tweedie, R.L. (1974). Quasi-stationary distributions for Markov chains on a general state space. *Journal of Applied Probability*, **11**, pp. 726–741.

Valdez-Flores, C. & Feldman, R.M. (1989). A survey of preventive maintenance models for stochastically deteriorating single-unit systems. *Naval Research Logistics Quarterly*, **36**, pp. 419–446.

Walls, L.A. & Bendell, A. (1985). The structure and exploration of reliability field data; what to look for and how to analyse it. *Reliability '85, Proceedings of the National Reliability Conference*, July 1985, UKAEA, pp. 5B/5/1–5B/5/18.

Widder, D.V. (1946). *The Laplace transform.* Princeton University Press, Princeton.

Yearout, R.D., Prabhaker Reddy & Grosh, D.L. (1986). Standby redundancy in reliability: a review. *IEEE Transactions on Reliability*, **R-35**, pp. 285–292.

Ziedins, I. (1987). Quasi-stationary distributions and one-dimensional circuit switched networks. *Journal of Applied Probability*, **24**, pp. 965–977.

Summary

A two-unit parallel system (with a single repair facility), a basic redundancy configuration in the reliability field, is subject of this thesis. The units of the system are statistically dependent, caused by the occurrence of common cause failures, which destroy both units simultaneously. Apart from a brief review basic stochastic models and processes, used to model the failure behaviour of repairable systems, the study concentrates on the derivation of expressions for the system's performance measures. Expressions are obtained for the system reliability, the availability, interval reliability, joint availability, the system state probabilities and the stationary counterparts of these measures. Further, the quasi-stationary distribution of the residual life time and the quasi-stationary system state probabilities are investigated.

The performance measures are characterised in terms of the system's state description process $\{X(t),t\geq0\}$: stochastic analysis of the process $X(t)$ yields explicit expressions for the above quantities. The state space of the process $X(t)$ is determined according to the kind of information wanted about the system and by the type of probability distribution which is used to model the life and repair times of the units. The use of phase type (PH) distributions (as defined by Neuts, 1981) and the discrete supplementary variable technique play a key role in the analysis of the system. It is shown that under bivariate phase type (BVPH) distributed life times and generally distributed repair times, including the phases of the BVPH distribution in the state space of $X(t)$ gives a regenerative state description process. The regeneration points of the process $X(t)$ are used to derive recurrence relations for the system's performance measures, conditioned on an entry into a particular regenerative state. Since the relations are in terms of sets of convolution integrals, taking the Laplace transforms yields sets of linear equations in the Laplace transforms of the performance measures. Subsequently, applying the Abelian theorem for Laplace transforms, the system's *stationary* performance measures are computed. On the other hand, the system's *transient* behaviour is investigated under BVPH distributed life times and PH distributed repair times. It is shown that including the phases of both life and repair time distributions renders the state description process $X(t)$ Markovian. Although the state space of the Markov process under consideration can become large, the transient performance measures of the system can be computed with minimum

round off error by using the randomisation method.

Apart from studying the transient and stationary performance measures, an extension is made in the direction of maintenance of a two-unit dependent parallel system: a continuous time opportunistic replacement policy of the control limit type is analysed, under generally distributed life and repair times. A partial differential equation is derived for the stationary joint pdf of the system's state description process $\{X(t),t\geq0\}$, which describes the operating repair and waiting time for both units at time t, under the replacement policy. Expressions for a number of long run operating characteristics are derived (such as the expected number of failures per unit of time, the expected number of preventive replacements, etc.) and a cost function is constructed to determine the optimal value of the control limit, which minimises the expected costs per unit of time in the long run. Numerical examples illustrate the techniques.

Finally, the ageing of a repairable system is discussed and it is shown how deterioration on system level can be modelled by connecting the parallel system in series with an additional component, called the system body, for which the life times are not necessarily identically distributed. Since the system deteriorates and hence the state description process $X(t)$ is non-stationary, the transient behaviour of the system is of particular interest. Time epochs which show a lack-of-memory property are exploited to derive expressions for the system's (interval) reliability and (joint) availability, under generally distributed life and repair times of the system body. However, the analysis is simplified considerably by the introduction of phase type distributions, which render the state description process Markovian. In the latter case the randomisation technique can be used to compute the system's transient performance measures.

Samenvatting

Onderwerp van dit proefschrift is een repareerbaar systeem met twee parallel geschakelde units die statistisch afhankelijk zijn en één reparateur. De afhankelijkheid wordt veroorzaakt door het optreden van zogenaamde common cause failures, die biede units tegelijkertijd doen falen. Na een kort overzicht van de basismodellen die in de literatuur gebruikt worden om het faalgedrag van een repareerbaar systeem te modelleren, concentreert het onderzoek zich op het afleiden van uitdrukkingen voor de prestatiematen van het systeem. Formules worden afgeleid voor de betrouwbaarheid van het systeem, de beschikbaarheid, interval betrouwbaarheid, simultane beschikbaarheid, de toestandskansen en de stationaire versies van genoemde grootheden. Bovendien wordt een tweetal quasi-stationaire verdelingen onderzocht, te weten de restlevensduurverdeling van het systeem en de verdeling van de stationaire toestandskansen, beide geconditioneerd op overleven van tijdstip t, waarbij $t \to \infty$.

De prestatiematen worden beschreven aan de hand van een stochastisch proces $\{X(t), t \geq 0\}$, dat de toestand beschrijft van het systeem op tijdstip t. Analyse van het proces $X(t)$ levert expliciete uitdrukkingen voor genoemde prestatiematen. De toestandsruimte van $X(t)$ wordt bepaald door de soort informatie die men wenst over het systeem en het type kansverdelingen dat gebruikt wordt om de levens- en reparatieduren van de units te modelleren. Het gebruik van fasetype (FT) verdelingen (Neuts, 1981) en de discrete supplementaire variabele techniek spelen hierbij een sleutelrol: het blijkt namelijk dat wanneer men de levensduren modelleert met een bivariate fasetype (BVFT) verdeling en de reparatieduren met een willekeurige verdeling, het proces $X(t)$ regeneratief wordt indien de identiteit van de fase waarin de BVFT verdeling zich bevindt, opgenomen wordt in de toestandsruimte van $X(t)$. Regeneratieve tijdstippen worden benut om recurrente betrekkingen af te leiden voor de prestatiematen van het systeem, geconditioneerd op de binnenkomst in een toestand die regeneratief van aard is. Daar deze betrekkingen bestaan uit convolutie-integralen leidt het nemen van de Laplacegetransformeerde tot een stelsel vergelijkingen in de Laplacegetransformeerde van de prestatiematen. Door het toepassen van de Abelstelling voor Laplacegetransformeerden kunnen vervolgens de *stationaire* prestatiematen eenvoudig bepaald worden. Anderzijds wordt het transiënte gedrag van het systeem onderzocht onder de veronderstelling van BVFT verdeelde levensduren en willekeurig verdeelde reparatieduren. Indien nu

niet alleen de identiteit van de fase waarin de BVFT verdeling zich bevindt, wordt opgenomen in de toestandsruimte van het proces X(t), maar ook de identiteit van de fase waarin de FT verdeling zich bevindt, is het resultaat een Markovproces. Alhoewel de toestandsruimte van dit Markovproces geweldig groot kan worden, kunnen de transiënte prestatiematen toch op een numeriek stabiele wijze berekend worden indien een methode toegepast wordt, die in de literatuur bekend staat onder de naam 'randomisation technique'.

Behalve het transiënte en stationaire gedrag wordt een onderhoudsstrategie bestudeerd, toegepitst op een systeem met twee parallel geschakelde, statistisch afhankelijke units. De beschouwde strategie is een oppotunistische vervangingsstrategie in continue tijd van het 'control limit' type. Onder de veronderstelling van willekeurig verdeelde levens- en reparatieduren, wordt een partiële differentiaalvergelijking afgeleid voor stationaire simultane kansdichtheid van het proces X(t), dat in dit geval beschrijft hoe lang een unit reeds werkt, in reparatie is of wacht op reparatie. Uitdrukkingen worden afgeleid voor een aantal stationaire systeemgrootheden, zoals het verwachte aantal storingen per tijdseenheid, het verwachte aantal preventieve vervangingen, etc. Bovendien wordt een kostenfunctie geconstrueerd om de optimale waarde van de control limit te bepalen, die de verwachte kosten per tijdseenheid minimaliseert. Numerieke voorbeelden illustreren de techniek.

Tenslotte wordt aandacht besteed aan een methode om geleidelijke veroudering van een repareerbaar systeem te modelleren. Hiertoe wordt het parallel systeem in serie geschakeld met een component wiens levenesduren niet (noodzakelijk) identiek verdeeld zijn. Voor een verouderend systeem is slechts het transiënte gedrag van het proces X(t), dat de systeemtoestand op tijdstip t beschrijft, van belang. De transiënte prestatiematen die afgeleid worden zijn de (interval) betrouwbaarheid en de (simultane) beschikbaarheid. De functie van regeneratieve tijdstippen wordt hierbij vervuld door tijdstippen die een geheugenloosheidseigenschap manifesteren. Echter, de analyse wordt aanzienlijk vereenvoudigd door de introductie van FT verdelingen, die van het proces X(t) een Markovproces maken. In dat geval kan de randomisation technique toegepast worden om de transiënte prestatiematen van het systeem te berekenen.

Acknowledgement

Ondanks de meerdere malen gebezigde uitspraak van de auteur een éénmans bedrijfje te leiden binnen de TUE, moge het duidelijk zijn dat het schrijven van een proefschrift geen éénmans bezigheid is. Om dit zwart op wit te bevestigen, zou ik dan ook willen bedanken: Peter Sander, de eerste promotor, voor de uiterst plezierige samenwwrking gedurende vier jaar, waarin alleen het tijdstip van de besprekingen niet altijd meeviel, Jaap Wessels, de tweede promotor, voor het nuttige commentaar, zijn aanwijzingen en schaafwerk, Martin Newby, voor zijn raad en hulp op velerlei gebied, zowel wiskundig als taalkundig, Gert Regterschot, voor het beschikbaar stellen van zijn encyclopedische kennis, niet aflatende behulpzaamheid en begeleiding, Jolanda Nelissen, voor de ruim 1000 koppen thee en de spreekwoordelijke, maar niet vanzelfsprekende secretariële ondersteuning (zoals dat zo mooi heet), collega's en ex-collega's van de vakgroep Operationele Research en Statistiek, voor de prettige sfeer waarbinnen 4 jaar lang dagelijks met plezier gewerkt kon worden, N. Ravichandran (Indian Institute of Management, Ahmedabad), die het onderzoek op een beslissend moment over een dood punt heen tilde, Jos Jansen van de vakgroep Numerieke Wiskunde, voor zijn bereidwilligheid en deskundige hulp bij het numeriek oplossen van de vele integraalvergelijkingen en anderen die op enigerlei wijze een bijdrage aan dit proefschrift geleverd hebben.

Curriculum vitae

De auteur van dit proefschrift werd geboren op 28 juni 1964 te Tilburg. Van 1976 tot 1982 heeft hij het gymnasium β doorlopen aan het Odulphuslyceum te Tilburg. Hierop aansluitend heeft hij wiskunde gestudeerd aan de Technische Universiteit Eindhoven, alwaar in 1988 het ingenieursdiploma behaald werd. Het afstudeerproject vond plaats bij de vakgroep Besliskunde & Stochastiek van de faculteit Wiskunde. Het project werd echter uitgevoerd in opdracht van prof. dr. P.C. Sander van de vakgroep Operationele Research & Statistiek van de faculteit Technische Bedrijfskunde en betrof de analyse van de gemiddelde niet-beschikbaarheid van repareerbare 1-uit-1-systemen. In september 1988 trad de auteur als assistent in opleiding in dienst bij de faculteit Technische Bedrijfskunde en begon onder leiding van prof. dr. P.C. Sander zijn promotie-onderzoek. De resultaten van het onderzoek, betreffende de stochastische analyse van een repareerbaar parallel systeem met statistisch afhankelijke componenten, zijn vastgelegd in dit proefschrift.

Stellingen

behorende bij het proefschrift

RELIABILITY ANALYSIS OF A REPAIRABLE DEPENDENT PARALLEL SYSTEM

I

De relatieve relevantie van reliability onderzoek wordt perfect getypeerd door de volgende uitspraak (Ascher *et al.*, 1984):

... We have indications of all three oil pressures on all three engines down to zero. We believe it to be faulty indications since the chances of all three engines having zero oil pressure and zero quantity is almost nil ...

An airline pilot

Ascher, H.E. & Feingold, H. (1984). *Repairable systems reliability: Modelling, inference, misconceptions and their causes.* Marcel Dekker, New York.

II

Naar derden toe verdient het de voorkeur bij de beschrijving van fasetype verdelingen (zoals gedefinieerd door Neuts, 1981) de subklasse bestaande uit eindige mengsels of sommen van Erlangverdelingen als uitgangspunt te nemen.

Neuts, M.F. (1981). *Matrix-geometric solutions in stochastic models.* John Hopkins University Press.

III

Dat enige voorzichtigheid in acht genomen dient te worden bij het klakkeloos berekenen van de stationaire performance measures van een systeem, getuige de volgende uitspraak van de econoom J.M. Keynes:

'In the long run we are all dead.'

Zie hoofdstuk 3 tot en met 7 van dit proefschrift.

IV

Gezien de enorme kloof die gaapt tussen theorie en praktijk, is een gesprek met mensen uit de onderhoudswereld vaak uitermate onderhoudend.

V

De naam 'Survival Analysis' getuigt van een optimistische kijk op het leven, gezien het feit dat (Elandt-Johnson, 1980):

... There is only a semantic difference between living and dying. From birth onwards we are approaching death ...

Elandt-Johnson, R.C. (1980). Some prior and posterior distributions in survival analysis: a critical insight on relationships derived from cross-sectional data. *Journal of the Royal Statistical Society, series* B, **42**, pp. 96-106.

VI

Uit publiciteitsoverwegingen had de naam 'Stochastiek' vervangen moeten worden door 'Chaostheorie'.

VII

Het frustrerende van een promotieonderzoek is dat N(k), het aantal personen dat een oordeel uitspreekt over een proefschrift aan de hand van k gelezen pagina's, dalend is in k.

VIII

Het dubbel tellen van 'uit' gescoorde doelpunten in een europacup treffen bij een gelijke eindstand na twee wedstrijden is flauwekul, daar het team met het grootste aantal 'uit' gescoorde goals eenvoudigweg winnaar is.

IX

Over het aanvaarden van een betrekking als AIO denkt men niet na, immers (Cato): 'Fronte capillata, post est occasio calva.'

Publius Valerius Cato. Disticha Catonis.

X

Partiële afschaffing van de coëducatie is de succesvolste methode om Thea techniek te laten studeren.

XI

Het voeren van knipperlicht overdag zou een typisch Nederlands compromis geweest zijn tussen geen licht en dimlicht.

M.J.P. Pijnenburg, augustus 1992.