QUEUEING AND COMMUNICATION NETWORKS GOVERNED BY GENERALISED LINDLEY-LOYNES EQUATIONS

by
DAVID MICHAEL ROSE

Submitted in partial fulfilment of the
requirements for the degree of
Master of Science,
in the
Department of Mathematical Statistics,
University of Natal
1993

Durban
1993

PREFACE

The research described in this thesis was carried out in the Department of Mathematical Statistics (Faculty of Science) at the University of Natal, Durban, during the period January 1993–December 1993, under the supervision of Dr S.A. Berezner.

These studies represent original work by the author and have not been submitted in any form to another university. Where use was made of the work of others it has been duly acknowledged in the text.

ACKNOWLEDGEMENTS

A debt of gratitude is owed to a number of individuals and organisations for their respective contributions to this thesis.

First and foremost, I wish to thank my supervisor, Dr Berezner, most sincerely for his time and effort, as well as his patience and guidance throughout the year. His colleague, Dr Y.M. Suhov, from the University of Cambridge must be thanked for his valuable contributions, as must Mr Alexander Harin.

I am grateful to Mrs Jackie de Gaye for her technical assistance and advice on the LaTeX typing, and to Professor L. Troskie for allowing me to use the facilities, especially the computers, in his department.

I am also greatly indebted to the University of Natal, the Foundation for Research Development and the Standard Bank. Without their financial support through scholarships and bursaries, it is unlikely that this research would have been possible.

This list of acknowledgements would be incomplete if I did not include the members of my family whom I thank for their interest and encouragement.

ABSTRACT

Several decades after A.K. Erlang originated the theory of queues and queueing networks, D.V. Lindley added impetus to the development of this field by determining a recursive relation for waiting times.

Part I of this thesis provides a theoretical treatment of single-server and multiserver queues described by the basic Lindley relation and its extensions, which are referred to collectively as Lindley-Loynes equations. The concepts of stability, and minimal and maximal solutions are investigated.

The interdependence of theory and practice becomes evident in Part II, where the results of recent and current research are highlighted. While the main aim of the first part of the thesis is to provide a firm theoretical framework for the sequel, the objective in Part II is to derive generalised forms of the Lindley-Loynes equations from different network protocols. Such protocols are regulated by different switching rules and synchronization constraints.

Parts I and II of the thesis are preceded by Chapter 0 in which several fundamental ideas (including those on notation and probability) are described. It is in this chapter too that a more detailed overview of the concept of the thesis is provided.

CONTENTS

LIST OF FIGURES AND TABLES

FIGURES

TABLES

CHAPTER 0

FOUNDATIONS

0.1 NOTATION AND CONVENTIONS

The first few pages of this chapter are devoted to the notation, terminology and concepts to be used throughout the thesis.

At the outset it is important to recognise that as a general rule, random variables and their realisations/values usually will not be distinguished notationally. Often the intended meaning will be evident from the context (e.g., $\mathrm{E}(x_n)$) or, in some cases, either interpretation may be valid. In addition, situations may arise where a capital letter such as W represents a random variable (or its realisation), distinct from some other random variable w.

General Conventions

Vectors and matrices are distinguished from scalars by **boldface type**.
An empty sum is 0 and an empty product is 1.
x is *positive* means $x > 0$, in contrast to x is *nonnegative* $\Rightarrow x \geq 0$. Similarly with *negative* and *nonpositive.*
If $x \in \mathbb{R}^*$ (f an extended real-valued function), then the *positive part* of x, $x^+ = \max(0, x)$ $(f^+(t) = \max(0, f(t)))$. Similarly, the *negative part* of x, $x^- = \max(0, -x) = -\min(0, x)$. Consequently, $x = x^+ - x^-$ and $|x| = x^+ + x^-$. $\lfloor x \rfloor$ denotes the *floor* or *greatest integer function* (satisfying $\lfloor x \rfloor \leq x < \lfloor x \rfloor + 1$).

STYLE OF PROOFS

Often in a proof a sequence of equalities and inequalities will be presented thus:

$$\begin{array}{rclll} A & \geq & B & (=) & \text{Lemma 6} \\ & \geq & C & & \forall\, n \end{array}$$

which must be interpreted as:

"$A \geq B$ because of Lemma 6 and $B = C$ for every n (so that $A \geq C$)".

The conclusion of a proof is indicated by a box, thus $\Box$.

ABBREVIATIONS

iff	if and only if
i.t.o.	in terms of
s.t.	such that
w.l.g.	without loss of generality
w.r.t.	with respect to

REFERENCES

p.	(*pp.*)	page(s)
§	(§§)	section(s)

SETS AND SEQUENCES

The distinction shall be made between a set (class or collection) and a sequence by using braces ($\{\}$) for a set and angle brackets ($\langle\rangle$) for sequences. The symbol for a proper subset shall be written as $\subset$, which should be distinguished from $\subseteq$. An analogous convention is adopted in the case of supersets (i.e., $\supset$ in contrast to $\supseteq$).

The cardinality of a (countable) set is delimited by $||$ (which will also be used to denote *absolute value* or *modulus*). *Equivalent* (*equipotent*) sets are indicated by the symbol $\simeq$.

Number Systems

$\mathbb{C}$	set of complex numbers	
$\mathbb{N}$	set of natural numbers	$\{1, 2, 3, \dots\}$
$\mathbb{Q}$	set of rational numbers	
$\mathbb{R}$	set of real numbers	
$\mathbb{Z}$	set of integers	$\{\dots, -2, -1, 0, 1, 2, \dots\}$

Let $X \subseteq \mathbb{R}$. Then the following notation is adopted:

$$\begin{array}{lcllcll}
X_- & = & \{x \in X : x < 0\} & X_-^* & = & X_- \cup \{-\infty\} & \text{if } X_- \neq \emptyset \\
X_\ominus & = & \{x \in X : x \leq 0\} & X_\ominus^* & = & X_\ominus \cup \{-\infty\} & \text{if } X_- \neq \emptyset \\
X_\oplus & = & \{x \in X : x \geq 0\} & X_\oplus^* & = & X_\oplus \cup \{\infty\} & \text{if } X_+ \neq \emptyset \\
X_+ & = & \{x \in X : x > 0\} & X_+^* & = & X_+ \cup \{\infty\} & \text{if } X_+ \neq \emptyset
\end{array}$$

Thus $\mathbb{Z}_- = -\mathbb{N}$, $\mathbb{Z}_\oplus = \mathbb{N} \cup \{0\}$ and $\mathbb{Z}_+ = \mathbb{N}$.

Moreover, the extended real number system $\mathbb{R}^* = \mathbb{R} \cup \{-\infty, \infty\}$, and $\mathbb{R}_\oplus^* = [0, \infty]$.

Next we let

$$[a, b]_X = [a, b] \cap X \quad \text{where } X \subseteq \mathbb{R} \text{ and } a, b \in \mathbb{R}$$

The extension to open and semi-open intervals is self-evident, so that $[-1, 4)_{\mathbb{Z}} = \{-1, 0, 1, 2, 3\}$.

A special symbol is reserved for the frequently encountered set $[1, n]_{\mathbb{N}}$. Write

$$\begin{array}{rcll}
\mathbb{N}_n & = & \{1, \dots, n\} & n \in \mathbb{N} \\
\text{and} \quad \overline{\mathbb{N}}_n = \mathbb{N} \setminus \mathbb{N}_n & = & \{n+1, n+2, \dots\} &
\end{array}$$

Let X be as above. Then denote the m-fold cartesian product of X as follows:

$$X^m = \bigtimes_{i=1}^{m} X$$

Probability Concepts

The following symbols and abbreviations will be used:

a.e.	almost every(where)
a.s.	almost sure(ly)
$\xrightarrow{d}$	convergence in distribution
$=_d$, $\geq_d$	Refer to Appendix C (Stochastic Ordering)
E	expectation (operator)
I	indicator function
IID	independent and identically distributed
·ID	identically distributed (but not necessarily independent)
I·D	independent (but not necessarily identically distributed)
i.o.	infinitely often
L	law (distribution)
P	probability (measure)
$\xrightarrow{p}$	convergence in probability
r.v.(s)	random variable(s)
$=_{st}$, $\geq_{st}$	Refer to Appendix C (Stochastic Ordering)
w.p.1	with probability one
$\sim$	distributed as

$$X \sim F \quad \Rightarrow \quad \begin{cases} L(X) = F \\ \overline{F}(x) = 1 - F(x), \text{ the tail (distribution) of } X \end{cases}$$

$$Y \sim \mathrm{Exp}(\lambda^{-1}) \quad \Rightarrow \quad \begin{cases} f_Y(y) = \lambda e^{-\lambda y} \text{ and } F_Y(y) = 1 - e^{-\lambda y} & y \geq 0 \\ \mathrm{E}Y = \lambda^{-1} & \lambda > 0 \end{cases}$$

$$Z_n \sim Z \quad \Rightarrow \quad Z_n =_d Z$$

0.2 PROBABILITY THEORY AND RELATED FIELDS

It will be instructive to review some concepts from probability theory and related fields. Assume an underlying probability space $(\Omega, \mathcal{F}, P)$ where necessary.

Definition 0.2.1 (Proper and Improper Random Variables)
Let X be a random variable taking values in $\mathbb{R}^*$ and having a distribution function F. Write

$$F(\infty) = \lim_{x\to\infty} F(x) \text{ and } F(-\infty) = \lim_{x\to-\infty} F(x)$$

X (or its distribution function F) is said to be *proper* or *honest*

iff $\quad F(-\infty) = 0$ and $F(\infty) = 1$
i.e. iff $\quad X$ is a.e. finite

On the other hand, X (or F) is referred to as being *improper, dishonest* or *defective*

iff $\quad F(-\infty) > 0$ or $F(\infty) < 1$
i.e. iff $\quad P[X = -\infty] > 0$ or $P[X = \infty] > 0$

Definition 0.2.2 (Strict Stationarity)
$\langle \xi_n, n \in \mathbb{N} \rangle$ is a *strictly, strongly* or *completely stationary* sequence iff

$$P\left[\bigcap_{j=1}^{N} \left[\xi_{n_j} \le x_j\right]\right] = P\left[\bigcap_{j=1}^{N} \left[\xi_{n_j+k} \le x_j\right]\right] \quad \forall k, N \in \mathbb{N} \text{ and } (n_1, \ldots, n_N) \in \mathbb{N}^N \tag{0.2.1}$$

Remark 0.2.1

In some instances, it may be desirable to generalise the situation to allow $k, N \in \mathbb{Z}$ but, of course, $\mathbb{N}$ and $\mathbb{Z}$ are equipotent. In this regard, one should also refer to the footnote in Borovkov [12, *p.* 8].

Furthermore, if one is considering the stationary *process* $(\xi_t, t \in \mathcal{T} \subseteq \mathbb{R}^n)$, (0.2.1) becomes

$$P\left[\bigcap_{j=1}^{N}\left[\xi_{t_j} \le x_j\right]\right] = P\left[\bigcap_{j=1}^{N}\left[\xi_{t_j+t} \le x_j\right]\right] \quad \forall N \in \mathbb{N}, t \in \mathcal{T} \text{ and } (t_1, \ldots, t_N) \in \mathcal{T}^N \tag{0.2.1$'$}$$

Remark 0.2.2

A necessary condition for the strict stationarity of $\langle \xi_n \rangle$ is that $\langle \xi_n \rangle$ is an identically distributed sequence. (Set $N = 1$ in (0.2.1).)

A sufficient condition for the strict stationarity of $\langle \xi_n \rangle$ is that $\langle \xi_n \rangle$ is an IID sequence. (Again this is evident from (0.2.1).)

Remark 0.2.3 (Equivalent Definition of Strict Stationarity)

Often the following alternative equation for the strict stationarity of $\langle \xi_n \rangle$ will prove useful:

$$P\left[(\xi_1, \xi_2, \ldots) \in B\right] = P\left[(\xi_{k+1}, \xi_{k+2}, \ldots) \in B\right] \ \forall B \in \mathcal{B}\left(\mathbb{R}^\infty\right)$$

where $\mathbb{R}^\infty$ is the space of (ordered) real sequences and $\mathcal{B}\left(\mathbb{R}^\infty\right)$, the σ-algebra generated by that space.

Before considering the idea of ergodicity, we recall some additional useful concepts.

Definition 0.2.3 (Measurable & Measure-Preserving Transformations)

A transformation $T : \Omega \to \Omega$ is *measurable* iff

$$T^{-1}A = \{\omega : T\omega \in A\} \in \mathcal{F} \quad \forall\, A \in \mathcal{F}$$

Further, a measurable transformation T is *measure-preserving* iff

$$P\left(T^{-1}A\right) = P(A) \quad \forall\, A \in \mathcal{F}$$

Remark 0.2.4 (Relationship between Measure-Preserving Transformations and Strictly Stationary Sequences)

(a) First we apply Definition 0.2.2 to demonstrate how stationary sequences may be constructed from a measure-preserving transformation T. Inductively we deduce that $P(T^{-n}A) = P(A)\ \forall\, n \in \mathbb{N}$. Bearing this in mind, let $\xi_k = \xi_k(\omega) = \xi_1(T^{k-1}\omega)$ $(k \in \overline{\mathbb{N}}_1)$ where $\xi_1 = \xi_1(\omega)$ is some random variable. Define $A_k = \left\{\omega : \bigcap_{j=1}^{N} \left[\xi_{n_j+k}(\omega) \le x_j\right]\right\} = \left\{\omega : \bigcap_{j=1}^{N} \left[\xi_{n_j}\left(T^k\omega\right) \le x_j\right]\right\}$ so that $A_0 = \left\{\omega : \bigcap_{j=1}^{N} \left[\xi_{n_j}(\omega) \le x_j\right]\right\}$. Consequently $\omega \in A_k \Leftrightarrow T^k\omega \in A_0$ or $\omega \in A_k \Leftrightarrow A_k = T^{-k}A_0$ so $P(A_0) = P(T^{-k}A_0) = P(A_k)$, where we use the fact that T is measure-preserving. It follows (by the arbitrary nature of N and k) that $\langle \xi_n \rangle$ is strictly stationary.

(b) (a) has a converse of sorts. That converse will be explained in terms of the definition given in Remark 0.2.3. We demonstrate that for every stationary sequence $\langle \xi_n \rangle$ considered on $(\Omega, \mathcal{F}, P)$, a new probability space, random variable and measure-preserving transformation $\left(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P}\right)$, $\tilde{\xi}_1(\tilde{\omega})$ and $\tilde{T}$ respectively can be constructed in such a way that the distributions of $\langle \xi_n \rangle$ and $\langle \tilde{\xi}_1(\tilde{\omega}), \tilde{\xi}_1(\tilde{T}\tilde{\omega}), \ldots \rangle$ coincide.

Let

$$\tilde{\Omega} = \mathbb{R}^\infty, \ \ \tilde{\mathcal{F}} = \mathcal{B}\left(\mathbb{R}^\infty\right)$$

$$\tilde{P}(B) = P\left\{\omega \in \Omega : \boldsymbol{\xi} = (\xi_1, \xi_2, \ldots) \in B\right\} \ \forall B \in \mathcal{B}\left(\mathbb{R}^\infty\right)$$

$$\tilde{T}\left(x_1, x_2, \ldots\right) = \left(x_2, x_3, \ldots\right) \ \forall\, \tilde{\omega} = \left(x_1, x_2, \ldots\right) \in \tilde{\Omega}$$

$$\tilde{\xi}_1\left(\tilde{\omega}\right) = x_1, \ \ \tilde{\xi}_k\left(\tilde{\omega}\right) = \tilde{\xi}_1(T^{k-1}\tilde{\omega}) = x_k \ \ \forall\, k \in \overline{\mathbb{N}}_1$$

Next, for $C \in \mathcal{B}(\mathbb{R}^k)$, write

$$\begin{aligned} A &= \left\{\tilde{\omega} : (x_1, \ldots, x_k) \in C\right\} \\ \text{and } \ \tilde{T}^{-1}A &= \left\{\tilde{\omega} : (x_2, \ldots, x_{k+1}) \in C\right\} \end{aligned}$$

It follows that

$$\begin{aligned} \tilde{P}(A) &= P\left\{\omega : \boldsymbol{\xi} \in A\right\} = P\left\{\omega : \boldsymbol{\xi} \in \left\{\tilde{\omega} : (x_1, \ldots, x_k) \in C\right\}\right\} \\ &= P\left\{\omega : (\xi_1, \ldots, \xi_k) \in C\right\} \end{aligned}$$

Similarly $\tilde{P}(\tilde{T}^{-1}A) = P\left\{\omega : (\xi_2, \ldots, \xi_{k+1}) \in C\right\}$ but strict stationarity of $\langle \xi_n \rangle \Rightarrow \tilde{P}(A) = \tilde{P}(\tilde{T}^{-1}A) \Rightarrow \tilde{T}$ is measure-preserving. In addition

$$\tilde{P}\left\{\tilde{\omega} : \left(\tilde{\xi}_1, \ldots, \tilde{\xi}_k\right) \in C\right\} = \tilde{P}(A) = P\left\{\omega : (\xi_1, \ldots, \xi_k) \in C\right\} \ \forall\, k \in \mathbb{N}$$

$$\Rightarrow \langle \xi_n \rangle =_d \langle \tilde{\xi}_n \rangle$$

Definition 0.2.4 (Invariance)
Let T be a measure-preserving transformation on $(\Omega, \mathcal{F}, P)$. Then $A \in \mathcal{F}$ is *invariant* (w.r.t. T) iff $T^{-1}A = A$ except for a possible exceptional set of probability zero.[1] Denote by $\mathcal{I}$ the class of invariant sets, which is also a σ-algebra.
$A \in \mathcal{F}$ is *invariant w.r.t. the sequence* $\langle \xi_n \rangle$ iff $\exists\, B \in \mathcal{B}(\mathbb{R}^\infty)$ such that $A = \{\omega : (\xi_n, \xi_{n+1}, \ldots) \in B\}\ \forall\, n \in \mathbb{N}$. Write $\mathcal{I}_\xi$ for the σ-algebra of such invariant sets.
A random variable $X = X(\omega)$ is *invariant* iff $P\{\omega : X(\omega) = X(T\omega)\} = 1$ iff it is $\mathcal{I}$-measurable.

Definition 0.2.5 (Ergodicity/Metric Transitivity)
A measure-preserving transformation T is *ergodic* or *metrically-transitive*

$$\text{iff} \qquad P(A) = 0 \text{ or } 1\ \forall \text{ invariant sets } A \in \mathcal{F}$$

$$\text{or equivalently} \quad \lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} P(A \cap T^{-j}B) = P(A)P(B)\ \forall\, A, B \in \mathcal{F} \tag{0.2.2}$$

Similarly, a stationary sequence $\langle \xi_n \rangle$ is *ergodic* iff the measure of every invariant set is 0 or 1.

[1]Some authors prefer to call such an event *almost invariant*, reserving the term *invariant* for the case $T^{-1}A = A$.

Remark 0.2.5

It is not difficult to see why (0.2.2) implies that $P(A) = 0$ or 1 for an invariant set $A(= B)$. For:

$$[P(A)]^2 = \lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} P(A \cap T^{-j}A) = \lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} P(A) = P(A)$$

Definition 0.2.6 (Mixing)

A measure-preserving transformation T is *(strongly) mixing* iff

$$\lim_{n\to\infty} P(A \cap T^{-n}B) = P(A)P(B)\ \forall\ A, B \in \mathcal{F}$$

Remark 0.2.6 (Relationship between Mixing and Ergodicity)

Suppose T is mixing. Then for an invariant set $A = B$ (in Definition 0.2.6)

$$[P(A)]^2 = \lim_{n\to\infty} P(A \cap T^{-n}A) = P(A)$$

as in Remark 0.2.5. In other words T mixing $\Rightarrow$ T ergodic.

Remark 0.2.7 (One-to-One Transformations)

If T is injective, then in the previous definitions and remarks $T^{-1}A$ may be substituted with TA, $T^{-j}B$ with $T^{j}B$, and so on.

Example 0.2.1

Let $(\Omega, \mathcal{F}, P) = ([0,1), \mathcal{B}([0,1)), \lambda)$ where λ is the Lebesgue measure, and consider the transformation

$$T\omega = (\omega + \theta) \bmod 1 \quad \omega \in \Omega,\ \theta \in [0,1)$$

Since Ω is the semiclosed interval $[0,1)$, T is injective so that Remark 0.2.7 may be applied.

Although $\omega + \theta \in [0,2)$, $(\omega+\theta) \bmod 1 \in [0,1)$, from which it is fairly obvious that $\forall\, A \in \mathcal{F}\ \ TA \in \mathcal{F}$, or T is measurable.

Write

$$\begin{array}{rclrcl} A_1 & = & \{\omega \in A : \omega + \theta < 1\} & TA_1 & = & \{\omega + \theta : \omega \in A_1\} \\ A_2 & = & \{\omega \in A : 1 \le \omega + \theta < 2\} & TA_2 & = & \{\omega + \theta - 1 : \omega \in A_2\} \end{array}$$

Then $A_1 \cup A_2 = A$ and $TA_1 \cup TA_2 = TA$.

For continuous intervals I_1 and I_2, and $a_i, b_i \in [0,1)$ $(i = 1,2)$, we obtain

$$\begin{array}{rclclcrcl} I_1 & \subseteq & A_1 & \Rightarrow & I_1 & = & [a_1, b_1) & \Rightarrow & \lambda(I_1) = b_1 - a_1 \\ & & & & & & \Updownarrow & & \\ J_1 & \subseteq & TA_1 & \Rightarrow & J_1 & = & [a_1 + \theta, b_1 + \theta) & \Rightarrow & \lambda(J_1) = \lambda(I_1) \end{array}$$

$$\begin{array}{rclclcrcl} I_2 & \subseteq & A_2 & \Rightarrow & I_2 & = & [a_2, b_2) & \Rightarrow & \lambda(I_2) = b_2 - a_2 \\ & & & & & & \Updownarrow & & \\ J_2 & \subseteq & TA_2 & \Rightarrow & J_2 & = & [a_2 + \theta - 1, b_2 + \theta - 1) & \Rightarrow & \lambda(J_2) = \lambda(I_2) \end{array}$$

Consequently $\lambda(A) = \lambda(TA)\ \forall\, A \in \mathcal{F}$ so T is measure-preserving.

If $\theta = \frac{m}{n} \in [0,1)_{\mathbb{Q}}$ $(m \in \mathbb{Z}_{\oplus}, n \in \mathbb{N})$, then T is nonergodic.

Write $A = \bigcup_{j=0}^{n-1} A_j$ where $A_j = \left\{\omega : \frac{j}{n} \le \omega \le \frac{2j+1}{2n}\right\}$. Then

$\lambda(A) = \sum_{j=0}^{n-1} \lambda(A_j) = \sum_{j=0}^{n-1} \frac{1}{2n} = \frac{1}{2}$ and $TA = \bigcup_{j=0}^{n-1} TA_j$ where

$$\begin{aligned} TA_j &= \{T\omega : \omega \in A_j\} &= \left\{(\omega + \theta) \bmod 1 : \tfrac{j}{n} \le \omega \le \tfrac{2j+1}{2n}\right\} \\ && = \left\{\omega' : \left(\tfrac{j+m}{n}\right) \bmod 1 \le \omega' \le \left(\tfrac{j+m}{n} + \tfrac{1}{2n}\right) \bmod 1\right\} \end{aligned}$$

Since the case $m = 0$ is trivial, suppose $m \in \mathbb{N}_{n-1}$. It follows that $\exists k \in \mathbb{N}_{n-1}$ s.t. $k + m = n$

$$\Rightarrow \quad \left(\tfrac{j+m}{n}\right) \bmod 1 = \begin{cases} \frac{j+m}{n} & j \in [0, k-1]_{\mathbb{Z}} \\ \frac{j+m-n}{n} & j \in [k, n-1]_{\mathbb{Z}} \end{cases}$$

$$\Rightarrow \quad TA_j = \begin{cases} A_{j+m} & j \in [0, k-1]_{\mathbb{Z}} \\ A_{j+m-n} & j \in [k, n-1]_{\mathbb{Z}} \end{cases}$$

$$\begin{aligned} &\Rightarrow \quad TA = A \\ &\Rightarrow \quad A \text{ is invariant} \quad \text{but } \lambda(A) \notin \{0, 1\} \\ &\Rightarrow \quad T \text{ is nonergodic} \end{aligned}$$

If, however, $\theta \in [0, 1)_{\mathbb{R}\setminus\mathbb{Q}}$ then T is ergodic. The interested reader is referred to Rosenblatt [51, *p.* 32] or Shiryayev [56, *p.* 380].

When θ is irrational so that T is ergodic, T is not mixing. For instance, if $\theta = \frac{1}{\sqrt{2}}$ and, in Definition 0.2.6, we let $B = [0, \frac{1}{\sqrt{2}})$ and $A = T^{n+1}B$, we find that $\lambda(T^{n+1}B \cap T^n B) = \sqrt{2} - 1 \; \forall n \in \mathbb{N} \Rightarrow \lim_{n\to\infty} P(A \cap T^n B) = \sqrt{2} - 1$ but $P(A)P(B) = \lambda(T^{n+1}B)\lambda(T^n B) = [\lambda(B)]^2 = \frac{1}{2} \ne \sqrt{2} - 1$ where we use the fact that T is measure-preserving. Similar results can be deduced for more general θ.

We are now in a position to state some theorems of particular interest in the forthcoming sections. The proofs will be found in most good books on probability.

Theorem 0.2.1 (Strong Law of Large Numbers for IID Sequences)
Let $\langle X_n \rangle$ be an IID sequence with $X_n \sim X \; \forall \, n \in \mathbb{N}$. If $\mathrm{E}|X| < \infty$, then

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} X_j = \mathrm{E}(X) \text{ a.s.}$$

Theorem 0.2.2 (Ergodic Theorem or **Strong Law of Large Numbers for Strictly Stationary Sequences)**
Let $\langle X_n \rangle$ be a strictly stationary sequence with $X_n \sim X \; \forall \, n \in \mathbb{N}$. If $\mathrm{E}|X| < \infty$, then

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} X_j = \mathrm{E}(X \mid \mathcal{I}_X) \text{ a.s.}$$

If, in addition, $\langle X_n \rangle$ is an ergodic sequence, then

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} X_j = \mathrm{E}(X) \text{ a.s.}$$

Remark 0.2.8
Versions of the previous two theorems, especially Theorem 0.2.1, exist in which it is required that only $\mathrm{E}(X^2) < \infty$ or even $\mathrm{E}(X) < \infty$. Indeed it is also possible to deduce a strong law for the case where $\mathrm{E}(X)$ exists but is

not necessarily finite.

We conclude this section with some lemmata from other areas of mathematics. These lemmata are rather simple, nonetheless, they will prove to be valuable in the thesis.

Lemma 0.2.1

For sequences $\langle a_r \rangle$ and $\langle b_r \rangle$

$$\sup_r a_r + \inf_r b_r \leq \sup_r (a_r + b_r)$$

<u>PROOF</u>

Let $\beta = \inf_r b_r$ and $\gamma = \sup_r (a_r + b_r)$.

Then, by definition

$$\begin{array}{llcll} & a_r + b_r & \leq & \gamma & \forall\, r \\ \text{but} & \beta & \leq & b_r & \forall\, r \\ \Rightarrow & a_r + \beta & \leq & a_r + b_r & \\ & & \leq & \gamma & \forall\, r \\ \Rightarrow & \gamma - \beta & \text{is} & \text{an upper bound for } \langle a_r \rangle & \\ \Rightarrow & \sup_r a_r & \leq & \gamma - \beta & \end{array}$$

from which the result follows.

$\square$

Lemma 0.2.2

Suppose $x_{n+1} = x_n + y_n \; \forall \, n \geq N \; (N \in \mathbb{Z})$. Then

$$x_{n+1} = x_N + \sum_{j=N}^{n} y_j$$

PROOF

We may write

$$\begin{aligned} x_{j+1} - x_j &= y_j \qquad \forall \, j \geq N \\ \Rightarrow \sum_{j=N}^{n} (x_{j+1} - x_j) &= \sum_{j=N}^{n} y_j \\ \Rightarrow \quad x_{n+1} - x_N &= \sum_{j=N}^{n} y_j \end{aligned}$$

□

0.3 AN OVERVIEW OF THE THESIS

As it was indicated in the Abstract, the major section of the thesis is divided into two parts, the first of which is based on the classic papers of Loynes [47] and Brandt [15]. Loynes's article deals mainly with the single-server queue, whereas Brandt [15] concentrates on multiserver queues. Despite the fact that more than two decades separates the publication of the two articles, Loynes [47] and Brandt [15] have much in common. An appropriate waiting-time equation, the idea of which was originated by Lindley [45], provides an important theoretical foundation for investigating the existence and uniqueness (or otherwise) of a stationary regime (for the queues under consideration). Chapters 1 and 2 are concerned with an analysis — similar to that of Loynes [47] and Brandt [15] — of single-server and multiserver queues.

The issue of the existence and uniqueness of a stationary regime for complex networks is investigated in Part II. This second part of the thesis is concerned with both the practical and theoretical aspects of queueing networks.

The necessary terminology on network protocols, switching rules and synchronization constraints is explained in Chapter 3, where examples from computer and communication technology are provided.

In the next chapter these general principles are applied to the specific case of a starlike network which is a convenient mechanism for deriving the theory and application (in telecommunications) of generalised Lindley-Loynes equations. When this network operates according to the $FAFS$ discipline (explained in Chapter 4), its waiting-times are governed by an extended or second-order Lindley equation, depending on the underlying switching rule. The extended Lindley equation, associated with a message-switched starlike network, is discussed in Chapter 5 which is based on Baccelli *et al.* [4]. The subsequent chapter is concerned with Lindley equations of second or higher order. Higher-order Lindley relations (Karpelevich *et al.* [35]) are determined by a suitable generalisation of the starlike network.

A more intricate network, with an appropriate synchronization rule, is considered in Chapter 7 where the paper, Berezner & Malyshev [8], is reviewed. The final, brief chapter of the thesis is concerned with the application of the $FAAFS$ protocol, introduced in Chapter 4, to the starlike network.

One of the differences between the basic queueing systems of Part I and the more complex queueing networks of Part II is that the non-uniqueness of a stationary regime in the latter case is a far more natural phenomenon. This is especially true of infinite networks.

The article, Kelbert *et al.* [36], provides some insight into the nature of non-uniqueness in countable Jackson networks where the set of service counters is denumerable. While this paper may not relate to starlike networks, the search for uniqueness and non-uniqueness conditions for the invariant measure of the underlying Markov process certainly has relevance.

More detailed introductions are given in Parts I & II and their respective chapters. Notation which is not used globally is also introduced in the appropriate chapter. Appendices are provided to avoid interrupting the flow of the central argument of the thesis.

PART I
BASIC QUEUEING SYSTEMS

At a discussion on a paper by David Kendall [40] in 1951, D.V. Lindley proposed a recursive-type equation for the waiting-times in a single-server queue. Soon thereafter he formalised his ideas in an article (Lindley [45]). Subsequently much interest was shown in formulating results for single- and many-server queues on the basis of waiting-time relations we call Lindley-Loynes equations. This is precisely the topic of Part I of the thesis.

Throughout, a working knowledge of basic concepts from the theory of queues is tacitly assumed. The reader is referred to a number of monographs, including Cohen [23], Cooper [24], Franken *et al.* [31], Gross & Harris [32], Kleinrock [42] (& [43]), Saaty [54] and Wolff [59]. It is appropriate, nonetheless, to review a few concepts and establish some notation to be used in the first part of the thesis.

The arrival-time sequence is denoted by $\langle t_n \rangle$, the service times by $\langle s_n \rangle$ and the interarrival-time sequence by $\langle \tau_n = t_{n+1} - t_n \rangle$, where $\tau_n \in \mathbb{R}_\oplus$ and $s_n \in \mathbb{R}_\oplus$. Typically $n \in \mathbb{N}$ or $n \in \mathbb{Z}$ where, particularly under the former assumption, n may often be interpreted as an index for the nth customer.[1] $\langle u_n \rangle$ will represent a suitable difference between the service and interarrival sequences. Although it is not always necessary to do so, *throughout Parts I and II we restrict our attention to realisations without batch arrivals.* In the language of point processes (Appendix B) we will disregard RMPPs which are not simple.

[1] Refer also to Doob [30, Chapter 10].

Multiserver queues will be assumed to consist of $r \in \overline{\mathbb{N}}_1$ servers. When $r = 1$ we are, of course, considering a single-server queue.

We will make one general assumption about the queues we consider in Part I:

[I.1] $\langle(\tau_n, s_n)\rangle$ is a (strictly) stationary ergodic sequence

for which the Kendall notation $G/G/r$ is used.[2] Additional restrictions may be imposed:

[I.2] $\langle\tau_n\rangle$ and $\langle s_n\rangle$ are independent sequences

[I.3] $\langle\tau_n\rangle$ is an IID sequence

[I.4] $\langle s_n\rangle$ is an IID sequence

When [I.1] and [I.2] are satisfied simultaneously with [I.3], then the queue is of type $GI/G/r$; the notation $G/GI/r$ is used if [I.4], rather than [I.3], holds.[3] The queue for which all four conditions are satisfied is denoted by $GI/GI/r$. In all cases a first-come, first-served ($FCFS$) (also referred to as $FIFO$: first-in, first-out) discipline is in operation. The situation may be summarised informally as:

$$GI/GI/r \;\subseteq\; \left\{ \begin{array}{c} GI/G/r \\ G/GI/r \end{array} \right\} \;\subseteq\; G/G/r$$

where we also note that if $\langle\tau_n\rangle$ (or $\langle s_n\rangle$) is an IID sequence, then it is stationary (Remark 0.2.2) and ergodic. (Apply stationarity and independence

[2] although $G/G/r$ need only mean a general input distribution and a general service distribution

[3] GI signifies a general independent distribution.

in Definition 0.2.5 or 0.2.6.) Furthermore, the stationarity stipulated in [I.1] implies that $\tau_n \sim \tau_0$ and $s_n \sim s_0 \; \forall \, n$ (Remark 0.2.2). Thus the traffic intensity may be given as $\rho = \frac{\mathrm{E}s_0}{r\mathrm{E}\tau_0}$.

The results proved by Lindley [45] are those for a $GI/GI/1$ queue (a generalisation of the $M/GI/1$ case treated by Kendall [40]) in which each of $\mathrm{E}(\tau_0)$ and $\mathrm{E}(s_0)$ is finite. In particular, Lindley showed that a necessary and sufficient condition for convergence to a steady-state waiting-time distribution is that either $\rho < 1$ or $s_0 = \tau_0$ a.s.

Both Kendall and Lindley seemed intimidated by the apparent difficulties associated with an r-server ($r > 1$) queue. In their paper, Kiefer & Wolfowitz [41] not only formulated a Lindley-Loynes equation for multiserver queues but also proved that, for a $GI/GI/r$ queue, $\rho < 1$ (or $\rho = 1$ and $s_0 = r\tau_0$ a.s.) is necessary and sufficient for convergence of the distribution of the waiting-times to a unique stationary (proper) distribution. The prestationary waiting-time sequence $\langle \boldsymbol{w}_n \rangle$ forms a Markov chain (because of the independence assumptions and Lindley-Loynes equation (2.2.6)). Uniqueness is then in the class of all stationary Markov state-processes.

A few years later Loynes [47] adapted the ideas of Kiefer & Wolfowitz, and applied several of his own, to a $G/G/r$ queue. The relaxation of assumptions on $\langle \tau_n \rangle$ and $\langle s_n \rangle$ inevitably leads to weaker conclusions. $\rho < 1$ is shown to be sufficient (but not necessary) for convergence to a stationary (proper) waiting-time distribution, the uniqueness of which may be guaranteed for $r = 1$.

Over two decades later Brandt [15] provided greater clarity on the issue of uniqueness of the waiting-time distribution for $G/G/r$ queues, and he was able to demonstrate (Brandt [16]) that $\rho < 1$ is sufficient for uniqueness of the stationary state process of a $G/GI/r$ queue. Example 2.5.1 (of this thesis) provides a trivial instance of a $GI/G/r$ queue which fails to satisfy this uniqueness property.

Part I comprises a review of Loynes [47] and Brandt [15], with Chapter 1 devoted to single-server queues and Chapter 2 to multiserver queues. An attempt is made at unifying the notation: that is, unless this is a hindrance rather than a help. A number of additional useful lemmata and theorems have been provided, especially in §2.3. Most proofs not given in the original papers *are* given by the author; even when proofs have been supplied in the article concerned, they are expanded upon — with any necessary corrections — in the thesis. Expanded proofs are especially required in response to Loynes's concise demonstration of results. While remarks clarify particular issues, more extensive details are also provided in the examples, some of which have been generalised.

We are now in a position to address specific problems and investigate results in depth.

CHAPTER 1

THE SINGLE-SERVER QUEUE

1.1 INTRODUCTION

Most of the preliminaries for this chapter were outlined in the introduction to Part I, so *this* introduction will be fairly short.

The sequence $\langle u_n \rangle$ consists of terms of the form $s_n - \tau_n$ ($\in \mathbb{R}$), which implies that the aforementioned sequence is stationary ergodic. $\mathrm{E}(u_0)$ is assumed to exist, but need not be finite: nonetheless, this means that at least one of $\mathrm{E}(s_0)$ and $\mathrm{E}(\tau_0)$ must be finite. τ_n is used for Loynes's T_n. In contrast to his notation, capital letters will *not* be used to symbolise strict stationarity. If stationarity of a random variable needs to be emphasised, then the r.v. will be underlined.

The (prestationary) Lindley equation for a single-server $FCFS$ queue is fairly easily derived (see Lemma 3.5.1):

$$w_{n+1} \quad = \quad (w_n + u_n)^+ \qquad \texttt{(1.1.1)}$$

The independence assumptions of Lindley [45] permit the transformation of (1.1.1) and similar relationships into integral equations (and convolutions), a topic covered in Kleinrock [42, Chapter 8].

Although we do not necessarily have the advantage of independence, a single-server queue is still more easily analysed than its multiserver counterpart.

Even then, some techniques used by Loynes (and others) to prove results for the $G/G/1$ case may be generalised and adapted to the $G/G/r$ queue.

In the next section the notion of stability is introduced using terminology a little different from Loynes. In particular, his *substability* is equivalent to our *stability*. The concept of a minimal stationary solution to (1.1.1) — also investigated in §1.2 — proves to be of considerable importance.

The third and final section of the chapter uses the results of §1.2 to provide a thorough investigation of $G/G/1$ queues for the three distinct domains $\{\rho > 1\}, \{\rho < 1\}$ and $\{\rho = 1\}$ of the traffic intensity.

1.2 STABILITY AND THE MINIMAL SOLUTION

Definition 1.2.1 (Stability)

Let $\langle x_n, n \in \mathbb{N}\rangle$ be a sequence of a.e. finite random variables such that x_n has a (proper) distribution function F_n $\forall\, n \in \mathbb{N}$. Then $\langle x_n \rangle$ shall be said to be

(a) *convergent stable* iff F_n converges weakly to a (unique) distribution function F

(b) *stable* iff every (infinite) subsequence contains a convergent stable sub-subsequence

(c) *divergent stable* iff $\langle x_n \rangle$ is stable but not convergent stable

(d) *unstable* iff it is not stable

Lemma 1.2.1 (Equivalent Condition for Stability)
$\langle x_n \rangle$ is a stable sequence iff there exist proper distribution functions G_1 and G_2 such that

$$G_1(x) \leq F_n(x) \leq G_2(x) \;\; \forall\, x \in \mathbb{R},\; \forall\, n \in \mathbb{N} \tag{1.2.1}$$

where F_n is as in Definition 1.2.1.

<u>PROOF</u>

Necessity.
Suppose $\langle x_n \rangle$ is a stable sequence but suppose to the contrary that no proper distribution $G_1(x)$ exists which satisfies (1.2.1). $\forall\, x$ let $G(x) = \inf_n F_n(x)$, which is known to exist. $G(x)$ cannot be a proper distribution function (has a defect at ∞) otherwise it would satisfy the requirements for $G_1(x)$. Consequently, since $\lim_{x\to\infty} G(x) < 1$, we have

$$\exists\, \delta \in (0,1) \text{ s.t. } \forall\, x \in \mathbb{R},\; G(x) < 1 - \delta \tag{1.2.2}$$

For the purposes of the proof of this lemma, it suffices to consider sequences $\langle F_n(m); m, n \in \mathbb{N} \rangle$. Then, by definition of the infimum,
$\forall\, \epsilon_1 > 0\; \exists\, n'$ s.t. $F_{n'}(n) < G(n) + \epsilon_1$ for each $n \in \mathbb{N}$. Choose such a $n' = k_n$ (which often will not be unique) that satisfies

$$F_{k_n}(n) < G(n) + \tfrac{\delta}{2} \tag{1.2.3}$$

and construct the subsequence $\langle F_{k_n}(x), n \in \mathbb{N} \rangle$ of $\langle F_n(x) \rangle$.

By Definition 1.2.1(b), there exists sub-subsequence $\left\langle F_{(m_k)_n}(x), n \in \mathbb{N}\right\rangle$ such that $F_{(m_k)_n}(x)$ converges to a (unique) proper distribution function $\widetilde{F}(x)$ for all points of continuity x of $\widetilde{F}$. Without loss of generality, and to avoid unnecessary complications in notation, we will assume that the natural numbers n are points of continuity of $\widetilde{F}$. For if they were not, we could consider a corresponding infinite sequence $\langle c_n \rangle$ of continuity points, where c_n is suitably defined.

Using the fact that $\widetilde{F}$ is proper, we have that $\forall\, \epsilon_2 > 0\ \exists\, X_0$ s.t. $x > X_0 \Rightarrow \widetilde{F} > 1 - \epsilon_2$ so, in particular

$$\exists\, X \text{ s.t. } x > X \Rightarrow \widetilde{F}(x) > 1 - \tfrac{\delta}{4} \tag{1.2.4}$$

Next, we may write that (for each point of continuity x of $\widetilde{F}$)
$\forall\, \epsilon_3 > 0\ \exists\, N$ s.t. $m_{k_n} \geq N \Rightarrow |F_{(m_k)_n}(x) - \widetilde{F}(x)| < \epsilon_3$. In particular

$$\exists\, N_1 \text{ s.t. } m_{k_n} \geq N_1 \Rightarrow \widetilde{F}(t_0) < F_{(m_k)_n}(t_0) + \tfrac{\delta}{4} \tag{1.2.5}$$

because of the convergence at continuity point t_0, where we may assume $t_0 > X$. Choose $N_2 \geq \max(N_1, t_0)$. Since $N_3 = k_{N_2} \geq N_2$

$$\begin{aligned}
\widetilde{F}(t_0) &< F_{N_3}(t_0) + \tfrac{\delta}{4} && (\leq) \text{ by } (1.2.5) \\
&< F_{N_3}(N_2) + \tfrac{\delta}{4} && F \text{ nondecreasing} \\
&< [G(N_2) + \tfrac{\delta}{2}] + \tfrac{\delta}{4} && \text{by } (1.2.3) \\
&< [(1-\delta) + \tfrac{\delta}{2}] + \tfrac{\delta}{4} && (=) \text{ by } (1.2.2) \\
&< 1 - \tfrac{\delta}{4}
\end{aligned}$$

which contradicts (1.2.4).

Thus the required G_1 *does* exist. A similar proof will verify the existence of G_2. So (1.2.1) holds.

Sufficiency.

Suppose (1.2.1) is satisfied. Let $\langle x_{k_n} \rangle$ be a subsequence of $\langle x_n \rangle$. By Helly's Theorem (Billingsley [11, Appendix II, *pp.* 226–227] or Shiryayev [56, *p.* 316]), $\exists$ subsequence $\langle x_{(m_k)_n} \rangle$ such that $F_{(m_k)_n}(x) \xrightarrow{d} \tilde{F}(x)$ but $G_1(x) \leq F_{(m_k)_n}(x) \leq G_2(x)$ $\forall\, n \in \mathbb{N}$ $(m_{k_n} \in \mathbb{N})$, $x \in \mathbb{R} \Rightarrow G_1(x) \leq \tilde{F}(x) \leq G_2(x)$ $\forall$ points of continuity x. Thus

$$
\begin{array}{rcccccccc}
1 & \geq & \lim\limits_{x \to \infty} \tilde{F}(x) & \geq & \lim\limits_{x \to \infty} G_1(x) & = & 1 \Rightarrow \lim\limits_{x \to \infty} \tilde{F}(x) & = & 1 \\
\text{and } 0 & \leq & \lim\limits_{x \to -\infty} \tilde{F}(x) & \leq & \lim\limits_{x \to -\infty} G_2(x) & = & 0 \Rightarrow \lim\limits_{x \to -\infty} \tilde{F}(x) & = & 0
\end{array}
$$

Consequently $\tilde{F}$ is a proper distribution and $\langle x_n \rangle$ is a stable sequence.

□

Corollary 1.2.1*a*

Sequence $\langle x_n \rangle$ is stable iff

$$\lim_{x \to -\infty} F_n(x) = 0 \text{ and } \lim_{x \to \infty} F_n(x) = 1 \text{ uniformly in } n$$

Corollary 1.2.1*b* (Sums of Stable Sequences)

Suppose $\langle x_n \rangle$ and $\langle y_n \rangle$ $(n \in \mathbb{N})$ are two stable sequences. Then $\langle x_n + y_n, n \in \mathbb{N} \rangle$ is also a stable sequence.

<u>PROOF</u>

$\langle x_n \rangle$ stable $\Rightarrow \lim_{t\to\infty} P[x_n \le t] = 1$ uniformly in $n \Rightarrow$

$$\forall\, \epsilon > 0\ \exists\, T_1 = T_1(\epsilon) > 0 \text{ s.t. } \forall\, n\ \ t \ge T_1 \Rightarrow P[x_n \le \tfrac{1}{2}t] > 1 - \tfrac{\epsilon}{2}$$

Similarly, the stability of $\langle y_n \rangle \Rightarrow$

$$\forall\, \epsilon > 0\ \exists\, T_2 = T_2(\epsilon) > 0 \text{ s.t. } \forall\, n\ \ t \ge T_2 \Rightarrow P[y_n \le \tfrac{1}{2}t] > 1 - \tfrac{\epsilon}{2}$$

Let $T = T(\epsilon) = T_1 + T_2$.[1] Then $\forall\, \epsilon > 0$ and $t \ge T$, uniformly in n

$$\begin{aligned} P[x_n + y_n \le t] &\ge P[\{x_n \le \tfrac{1}{2}t\} \cap \{y_n \le \tfrac{1}{2}t\}] \qquad (=) \\ &\ge P[x_n \le \tfrac{1}{2}t] + P[y_n \le \tfrac{1}{2}t] \\ &\quad -P[\{x_n \le \tfrac{1}{2}t\} \cup \{y_n \le \tfrac{1}{2}t\}] \\ &> 1 - \tfrac{\epsilon}{2} + 1 - \tfrac{\epsilon}{2} - 1 \qquad (=) \\ &> 1 - \epsilon \end{aligned}$$

$$\Rightarrow \lim_{t\to\infty} P[x_n + y_n \le t] = 1 \text{ uniformly in } n$$

That $\lim_{t\to-\infty} P[x_n + y_n \le t] = 0$ uniformly in n follows by observing that

[1] In fact, $T = \frac{1}{2}(T_1 + T_2)$ will do.

$\lim_{t \to -\infty} P[x_n + y_n > t] = 1$ uniformly in n, after applying an argument similar to that above. Hence $\langle x_n + y_n \rangle$ is stable, as claimed.

□

Remark 1.2.1

If $\langle x_n \rangle$ is a sequence of *nonnegative* random variables, then the distribution function $G_2(x)$ of Lemma 1.2.1 and the stipulation $\lim_{x \to -\infty} F_n(x) = 0$ of Corollary 1.2.1a become redundant as a condition for stability. The reason is obvious: $F_n(x) = 0 \; \forall \, x < 0$.

Lemma 1.2.2

Let the random variables $\langle w_n, n \in \mathbb{N} \rangle$ be determined by the relationship[2]

$$w_{n+1} = f(w_n, u_n) \tag{1.2.6}$$

where

initial condition w_1 is stipulated,
$\langle u_n, n \in \mathbb{Z} \rangle$ is a stationary sequence, and
$f(x, y)$ is a nonnegative function which is nondecreasing and continuous from below (even at ∞) in its first argument.

Fix $m \in \mathbb{Z}$, and consider the sequence $\langle w_n^m, n \in \mathbb{Z} \rangle$ given by

$$w_n^m = \begin{cases} 0 & n \in (-\infty, 1-m]_{\mathbb{Z}} \\ f(w_{n-1}^m, u_{n-1}) & n \in [2-m, \infty)_{\mathbb{Z}} \end{cases} \tag{1.2.7}$$

Then

(a) for fixed n, w_n^m is nondecreasing in m

[2]In this lemma w_n and u_n need not represent waiting times or the difference $(s_n - \tau_n)$ respectively, but can be more general than this.

(b) the sequence $\left\langle w_n^{l-n}, n \in \mathbb{Z} \right\rangle$ is

(i) strictly stationary for fixed $l \in \mathbb{Z}$, and

(ii) determined $\forall\, n \in \mathbb{Z}$ by

$$w_n^{l-n} = \left[\sup_{k \in (n-l,n)_{\mathbb{Z}}} \sum_{j=k}^{n-1} u_j \right]^+ \qquad l \in \mathbb{Z} \tag{1.2.8}$$

when

$$f(x,y) = (x+y)^+$$

<u>PROOF</u>

(a) Assume initially that $n \in (-\infty, 1-m)_{\mathbb{Z}}$. Then

$$w_n^{m+1} = 0 = w_n^m$$

If $n = 1-m$, then

$$\begin{aligned} w_n^{m+1} &= f\left(w_{n-1}^{m+1}, u_{n-1}\right) \\ &\geq 0 \qquad\qquad (=)\ \ f \text{ nonnegative} \\ &\geq w_n^m \end{aligned}$$

Suppose now that $w_n^{m+1} \geq w_n^m$ for some $n \in [1-m, \infty)_{\mathbb{Z}}$. Then

$$\begin{aligned} w_{n+1}^{m+1} &= f\left(w_n^{m+1}, u_n\right) \\ &\geq f\left(w_n^m, u_n\right) \qquad (=) \\ &\geq w_{n+1}^m \end{aligned}$$

We have thus established, by the inductive principle, that

$$\begin{aligned} \forall\, n \in \mathbb{Z} \quad w_n^{m+1} &\geq w_n^m \\ \text{i.e.} \qquad w_n^{\overline{m}} &\geq w_n^m \quad \text{for } \overline{m} = m+1 \end{aligned}$$

Suppose that $\overline{m} > m \Rightarrow w_n^{\overline{m}} \geq w_n^m \;\forall\, n \in \mathbb{Z}$. Then

Case 1 : $n \leq 1 - (\overline{m}+1) < 1 - m$

$$w_n^{\overline{m}+1} = 0 = w_n^m$$

Case 2 : $1 - (\overline{m}+1) < n \leq 1 - m$

$$\begin{aligned} w_n^{\overline{m}+1} &= f\left(w_{n-1}^{\overline{m}+1}, u_{n-1}\right) \\ &\geq 0 && (=)\quad f \text{ nonnegative} \\ &\geq w_n^m \end{aligned}$$

Case 3 : $n > 1 - m > 1 - (\overline{m}+1)$

$$\begin{aligned} w_n^{\overline{m}+1} &= f\left(w_{n-1}^{\overline{m}+1}, u_{n-1}\right) \\ &\geq f\left(w_{n-1}^{\overline{m}}, u_{n-1}\right) && \text{from original hypothesis} \\ &\geq f\left(w_{n-1}^{m}, u_{n-1}\right) \quad (=) && \text{by the inductive hypothesis} \\ &\geq w_n^m \end{aligned}$$

So finally

$$w_n^{\overline{m}} \geq w_n^m \qquad \forall\, n \in \mathbb{Z} \text{ and } \forall\, \overline{m} > m$$

(b) In accordance with the definition of w_n^m, on setting $m = l - n$, we obtain

$$w_n^{l-n} = \begin{cases} 0 & l \in \mathbb{Z}_\ominus \cup \{1\} \\ f(w_{n-1}^{l-n}, u_{n-1}) & l \in \overline{\mathbb{N}}_1 \end{cases}$$

(i) Define a sequence $\left\langle f_*^{[l-1]}, l \in \mathbb{Z} \right\rangle$ with terms

$$\begin{aligned} f_*^{[l-1]} &= 0 \;\forall\, l \in \mathbb{Z}_\ominus \cup \{1\}; \\ f_*^{[1]} &= f(0, u_{n-1}); \\ f_*^{[2]} &= f[f(0, u_{n-2}), u_{n-1}]; \\ f_*^{[3]} &= f\{f[f(0, u_{n-3}), u_{n-2}], u_{n-1}\}; \ldots \end{aligned}$$

It is not difficult to show that

$$w_n^{l-n} = f_*^{[l-1]} = \begin{cases} 0 & l \in \mathbb{Z}_\ominus \cup \{1\} \\ f_*^{[l-1]}(u_{n-(l-1)}, \ldots, u_{n-1}) & l \in \overline{\mathbb{N}}_1 \end{cases}$$

In particular, for a fixed $l \in \mathbb{Z}_\ominus \cup \{1\}$, $\langle w_n^{l-n} \rangle$ is merely a constant sequence of zeroes and is, therefore, obviously stationary. On the other hand, if l is some fixed integer greater than one, then w_n^{l-n} is a function $(f_*^{[l-1]})$ of $l-1$ consecutive terms of the stationary sequence $\langle u_n \rangle$ which implies that $\langle w_n^{l-n} \rangle$ is again stationary.

(ii) Initially let $l \in \mathbb{Z}_\ominus \cup \{1\}$. Then

$$\begin{aligned} & (n-l, n)_{\mathbb{Z}} &=& \emptyset \\ \Rightarrow \quad & \sum_{j=k}^{n-1} u_j &=& 0 \qquad \forall\, k \in (n-l, n)_{\mathbb{Z}} \\ \Rightarrow \quad & \left[\sup_{k \in (n-l,n)_{\mathbb{Z}}} \sum_{j=k}^{n-1} u_j \right]^+ &=& 0 \\ & &=& w_n^{l-n} \quad \text{as required} \end{aligned}$$

Now consider $l \in \overline{\mathbb{N}}_1$. We proceed by induction on l.

$$
\begin{aligned}
w_n^{2-n} &= \left(w_{n-1}^{2-n} + u_{n-1}\right)^+ \\
&= \left(w_{n-1}^{1-(n-1)} + u_{n-1}\right)^+ \\
&= (0 + u_{n-1})^+ \\
&= \left[\sup \sum_{j=n-1}^{n-1} u_j\right]^+ \\
&= \left[\sup_{k\in(n-2,n)_{\mathbb{Z}}} \sum_{j=k}^{n-1} u_j\right]^+
\end{aligned}
$$

Suppose (1.2.8) holds for some $l \in \overline{\mathbb{N}}_1$.

$$
\begin{aligned}
w_n^{l+1-n} &= \left(w_{n-1}^{l+1-n} + u_{n-1}\right)^+ \\
&= \left(w_{n-1}^{l-(n-1)} + u_{n-1}\right)^+ \\
&= \left(\left[\sup_{k\in(n-1-l,n-1)_{\mathbb{Z}}} \sum_{j=k}^{n-2} u_j\right]^+ + u_{n-1}\right)^+ \\
&= \left[\max\left\{\left(\sup_{k\in(n-1-l,n-1)_{\mathbb{Z}}} \sum_{j=k}^{n-1} u_j\right), u_{n-1}\right\}\right]^+ \\
&= \left[\sup_{k\in(n-(l+1),n)_{\mathbb{Z}}} \sum_{j=k}^{n-1} u_j\right]^+
\end{aligned}
$$

which is nothing more than a special case of $f_*^{[l-1]}$.

$\square$

Remark 1.2.2

For Lemma 1.2.2, recently proved, and Theorem 1.2.1, which follows shortly, it is helpful to consider a grid of $\mathbb{Z} \times \mathbb{Z}$, with m and n as axes.

Remark 1.2.3 (Initial Waiting-Time)

When $\langle w_n \rangle$ of equation (1.1.1) (and (1.2.6)) represents a sequence of waiting-times, then w_1, the waiting-time of the first customer to the queue, is treated as an initial value. If $w_1 = \alpha$ (either a nonnegative real number or a non-negative random variable), then we shall write $\langle w_n(\alpha) \rangle$ to emphasise that we are considering the waiting-time sequence with initial value α (or $\alpha(\omega)$).

From (1.2.7), we observe that

$$w_n^0 = \begin{cases} 0 & n \leq 1 \\ f\left(w_{n-1}^0, u_{n-1}\right) & n > 1 \end{cases}$$

or $w_n(0) = w_n^0$ on $\mathbb{N}$

Finally, we note that (1.1.1), although physically meaningful only when $n \in \mathbb{N}$, may always be interpreted mathematically $\forall\, n \in \mathbb{Z}$.

Theorem 1.2.1

Return to Lemma 1.2.2. Under the conditions associated with (1.2.6), there exists a stationary sequence of (possibly improper) random variables, $\langle m_n, n \in \mathbb{Z}\rangle$, such that

(a) $m_{n+1} = f(m_n, u_n)$

(b) $L\left(w_n^0\right) \downarrow L(m_0)$

(c) (i) $w_n(0)$ $(= w_n^0)$ is the minimal sequence satisfying (1.2.6) $\forall\, n \in \mathbb{N}$

(ii) $\langle m_n\rangle$ is the minimal (stationary) sequence satisfying (1.2.6) $\forall\, n \in \mathbb{Z}$

(iii) $w_n^0 \leq m_n \;\forall\, n \in \mathbb{Z}$

(d)

$$m_n = \left[\sup_{k\in\mathbb{N}} \sum_{j=1}^{k} u_{n-j}\right]^+ \tag{1.2.9}$$

PROOF

(a) Recall the nondecreasing (more especially, the monotonic) sequence $\langle w_n^m, m \in \mathbb{Z}\rangle$ of Lemma 1.2.2. It follows that $\exists\ \langle m_n, n \in \mathbb{Z}\rangle$ such that $w_n^m \uparrow m_n$, which will be finite if $\langle w_n^m, m \in \mathbb{Z}\rangle$ is bounded above.

Next

$$\begin{aligned} m_{n+1} &= \lim_{m\to\infty} w_{n+1}^m \\ &= \lim_{m\to\infty} f\left(w_n^m, u_n\right) \quad \text{from (1.2.7)} \\ &= f\left(\lim_{m\to\infty} w_n^m, u_n\right) \quad \text{by the left continuity of } f \\ &= f(m_n, u_n) \end{aligned}$$

This is valid $\forall\, n \in \mathbb{Z}$ since $1 - m \underset{\infty}{\overset{m}{\rightarrow}} -\infty$.

The strictly stationary sequence $\langle w_n^{m-n}, n \in \mathbb{Z}\rangle \uparrow \langle m_n, n \in \mathbb{Z}\rangle$ which implies that $\langle m_n, n \in \mathbb{Z}\rangle$ is strictly stationary, so we have proved (a).

(b) Recall Remark 1.2.3.

We may write w_n^0 as w_n^{n-n} and w_0^n as w_0^{n-0} so each of w_n^0 and w_0^n is a term in the strictly stationary sequence $\left\langle w_k^{n-k}, k \in \mathbb{Z}\right\rangle$ (Lemma 1.2.2(b)(i)) which implies that $w_n^0 =_d w_0^n \uparrow m_0$, where we use Remark 0.2.2 and (a) of the current theorem. Then $[w_0^n \leq x] \downarrow [m_0 \leq x] \Rightarrow P\left[w_n^0 \leq x\right] = P\left[w_0^n \leq x\right] \downarrow P\left[m_0 \leq x\right]$. Consequently $L\left(w_n^0\right) \downarrow L\left(m_0\right)$.

(c) If $\langle x_n\rangle$ is a nonnegative solution sequence of some equation and $\langle y_n\rangle$ is the minimal such sequence, then $x_n \geq y_n \;\forall\, n$ under consideration.

(i) Suppose $w_{n+1}^0 = f\left(w_n^0, u_n\right) \;\forall\, n \in \mathbb{N}$ and $\langle x_n\rangle$ is also a solution sequence to (1.2.6). Then $x_1 \geq 0 = w_1^0$.

Suppose now that $x_n \geq w_n^0$ for some $n \in \mathbb{N}$. It follows that

$$\begin{aligned} x_{n+1} &= f(x_n, u_n) \\ &\geq f(w_n^0, u_n) \quad (=) \quad f \text{ nondecreasing} \\ &\geq w_{n+1}^0 \end{aligned}$$

hence the result (by induction).

(ii) We have $m_{n+1} = f(m_n, u_n)\ \forall\, n \in \mathbb{Z}$; assume that $\langle z_n \rangle$ also satisfies (1.2.6) $\forall\, n \in \mathbb{Z}$.
Initially we demonstrate by induction that $z_n \geq w_n^m\ \forall\, n \in \mathbb{Z}$.
If $n \leq 1 - m$, then $w_n^m = 0 \Rightarrow z_n \geq 0 = w_n^m$.
Next suppose that the result holds for some $n \geq 1 - m$. Then

$$\begin{aligned} z_{n+1} &= f(z_n, u_n) \\ &\geq f(w_n^m, u_n) \quad (=) \quad f \text{ nondecreasing} \\ &\geq w_{n+1}^m \end{aligned}$$

Thus $z_n = \lim\limits_{m\to\infty} z_n \geq \lim\limits_{m\to\infty} w_n^m = m_n \quad \forall\, n \in \mathbb{Z}$.

(iii) $m_n \geq 0 = w_n^0 \quad \forall\, n \in \mathbb{Z}_\ominus \cup \{1\}$
Suppose $m_n \geq w_n^0$ for some $n \in \mathbb{N}$. Then

$$\begin{aligned} m_{n+1} &= f(m_n, u_n) \\ &\geq f(w_n^0, u_n) \quad (=) \quad f \text{ nondecreasing} \\ &\geq w_{n+1}^0 \end{aligned}$$

It follows that $m_n \geq w_n^0\ \forall\, n \in \mathbb{Z}$.

(d) Recall that it was shown in (1.2.8) that

$$w_n^{l-n} = \left[\sup_{k \in (n-l,n)_{\mathbb{Z}}} \sum_{j=k}^{n-1} u_j \right]^+$$

Then

$$\begin{aligned} w_n^m &= w_n^{(m+n)-n} \\ &= \left[\sup_{k' \in (-m,n)_{\mathbb{Z}}} \sum_{j'=k'}^{n-1} u_{j'} \right]^+ \; m \uparrow \infty \; \left[\sup_{k' \in (-\infty,n-1]_{\mathbb{Z}}} \sum_{j'=k'}^{n-1} u_{j'} \right]^+ \\ \text{or} \quad m_n &= \left[\sup_{k' \in (-\infty,n-1]_{\mathbb{Z}}} \sum_{j=1}^{n-k'} u_{n-j} \right]^+ \\ &= \left[\sup_{k \in \mathbb{N}} \sum_{j=1}^{k} u_{n-j} \right]^+ \end{aligned}$$

which is (1.2.9).

□

Remark 1.2.4 (Minimal Solutions to (1.2.6))
The significance of the result contained in part (c) of Theorem 1.2.1 was emphasised in the introduction to Part I. (It is now also evident why Loynes [47] chose the notation M_n (and I, m_n): m_n is the minimal stationary solution to the equation under consideration.

Since $w_n^0 = 0 \; \forall \, n \leq 1$, the sequence $\langle w_n^0 \rangle$ generally will not satisfy (1.2.6) on $\mathbb{Z}_\ominus$. Consequently, it cannot be viewed as a minimal solution to that equation on $\mathbb{Z}$. As demonstrated in Theorem 1.2.1(c)(iii), however, $\langle w_n^0 \rangle$ and $\langle m_n \rangle$ *may* be compared on $\mathbb{Z}$, but $\langle w_n(0) \rangle$ and $\langle m_n \rangle$ only on $\mathbb{N}$. The comparison between $\langle w_n(\alpha) \rangle$ and $\langle m_n \rangle$ for a more general α extends this idea $\forall\, n \in \mathbb{N}$:

$$\begin{aligned} w_n(\alpha) &\geq m_n \quad \text{on} \quad A = \{\omega : \alpha(\omega) \geq m_1(\omega)\} \\ w_n(\alpha) &\leq m_n \quad \text{on} \quad \overline{A} \end{aligned}$$

Lemma 1.2.2 and Theorem 1.2.1 are concerned with a fairly general function. Return now to the more specific function $f(x, y) = (x + y)^+$, which is obviously nonnegative. With respect to its first argument[3] it is continuous from the left everywhere (in fact, continuous everywhere) and nondecreasing since $x_1 > x_2 \Rightarrow f(x_1, y) \begin{cases} = & f(x_2, y) \quad x_1 \leq -y \\ > & f(x_2, y) \quad x_1 > -y \end{cases}$

Theorem 1.2.2

(a) (i) A queue is stable if $\limsup_{k \to \infty} \sum_{j=1}^{k} u_{-j} < \infty$ w.p.1 and is unstable otherwise.

(ii) For an initial condition $w_1 = 0$, however, the queue cannot be divergent stable.

(b) In the case of an unstable queue, the waiting-time distribution function tends to zero.

[3] and its second argument, for that matter

PROOF

(a) From previous theory, it is immediately clear that $\left\langle m_n = \lim_{m\to\infty} w_n^m \right\rangle$ is a stationary solution to the equation $w_{n+1} = (w_n + u_n)^+$ for $n \in \mathbb{Z}$. Also, m_n is the minimal stationary solution for $n \in \mathbb{Z}$ and $L(w_n^0) \downarrow L(m_0)$.

(i) It follows that, if the queue is to be stable, m_0 must be a proper random variable; m_0 improper implies an unstable queue. From (1.2.9)

$$m_0 = \left[\sup_{k\in\mathbb{N}} \sum_{j=1}^{k} u_{-j}\right]^+$$

$$\Longrightarrow \quad m_0 < \infty \text{ a.e.} \quad \Leftrightarrow \quad \sup_{k\in\mathbb{N}} \sum_{j=1}^{k} u_{-j} < \infty \text{ a.e.}$$

$$\Leftrightarrow \quad \limsup_{k} \sum_{j=1}^{k} u_{-j} < \infty \text{ a.e.}$$

It is clear that the event $\left[\limsup_{k} \sum_{j=1}^{k} u_{-j} < \infty\right]$ is invariant w.r.t. the sequence $\langle u_n \rangle$ (Definition 0.2.4) (certainly it is a tail event) and since $\langle u_n \rangle$ is ergodic, Definition 0.2.5 $\Rightarrow$

$$P\left[\limsup_{k} \sum_{j=1}^{k} u_{-j} < \infty\right] = \begin{cases} 0 & \Rightarrow \quad m_0 \text{ a.e. infinite: queue unstable} \\ & \text{or} \\ 1 & \Rightarrow \quad m_0 \text{ a.e. finite: queue stable} \end{cases}$$

(ii) $L(w_n^0) \downarrow L(m_0)$ implies convergence to a (unique) distribution which may be proper or improper but, in either case, the queue is not divergent stable.

(b) Since it was shown in (a) that the unstable queue corresponds to the case of m_0 a.e. infinite (i.e., $P[m_0 = \infty] = 1$), it follows that $P[m_0 \leq x] = 0 \; \forall \, x \in \mathbb{R}$.

□

Remark 1.2.5

If m_0 is dishonest, then Theorem 1.2.2(a) implies that m_0 is a.e. infinite but, of course, the stationarity of m_n, in accordance with Remark 0.2.2 implies that $\langle m_n \rangle$ is a sequence of a.e. infinite random variables. An analogous conclusion can be drawn in the case where m_0 is honest.

1.3 REGIMES OF STABILITY AND STATIONARITY

Theorem 1.3.1 (Supercritical Queue)

Consider a queue with $\mathrm{E}(u_0) > 0$, or $\rho > 1$. Then

(a) the queue is unstable

(b) initial conditions do not affect this instability

(c) there exists no stationary sequence of a.e. finite random variables satisfying (1.1.1) $\forall \, n \in \mathbb{Z}$.

(d) For any $w_1 = \alpha \geq 0$

$$\frac{w_n(\alpha)}{n} \underset{\infty}{\overset{n}{\rightarrow}} \mathrm{E}(u_0) \text{ a.s.}$$

PROOF

$$\lim_{k\to\infty}\sum_{j=1}^{k} u_{-j} = \lim_{k\to\infty} k\left(\frac{1}{k}\sum_{j=1}^{k} u_{-j}\right) = \mathrm{E}(u_0)\lim_{k\to\infty} k = \infty \text{ a.s.}$$

where the strong law for stationary sequences (Theorem 0.2.2) is applied in the second equality.

(a) Since $\lim_{n\to\infty} x_n = \infty \Leftrightarrow \limsup_{n\to\infty} x_n = \infty$, we obtain $\limsup_{k\to\infty}\sum_{j=1}^{k} u_{-j} = \infty$ (w.p.1).

On the basis of Theorem 1.2.2(a), we deduce that the queue is unstable, when the initial condition is given by $w_1 = 0$.

(b) Suppose now that the waiting-time of the first customer is some $\alpha \geq 0$. The case $\alpha = 0$ has already been considered, so let $\alpha > 0$.

Recall from Theorem 1.2.1(c) that $w_n(0)$ is the minimal solution of (1.1.1) for $n \in \mathbb{N}$. As a result, $w_n(\alpha) \geq w_n^0 \; \forall\, n \in \mathbb{N}$

$$\begin{aligned}
\Rightarrow\quad 0 &\leq \lim_{n\to\infty} P[w_n(\alpha) \leq x] \\
&\leq \lim_{n\to\infty} P[w_n^0 \leq x] \\
&= P[m_0 \leq x] && \text{by Theorem 1.2.1(b)} \\
&= 0 && \forall\, x \in \mathbb{R}
\end{aligned}$$

$$\Rightarrow\quad \lim_{n\to\infty} P[w_n(\alpha) \leq x] = 0 \quad \forall\, x \in \mathbb{R}$$

$$\Rightarrow\quad \lim_{n\to\infty} P[w_n(\alpha) = \infty] = 1 \quad \text{since } w_n(\alpha) \geq 0$$

(c) From (a), Theorem 1.2.2(a) and Remark 1.2.5, we deduce that $\langle m_n \rangle$ is an a.e. infinite sequence; from Theorem 1.2.1(c), $\langle m_n, n \in \mathbb{Z} \rangle$ is the minimal stationary sequence satisfying (1.1.1). Any other stationary sequence would, therefore, also be a.e. infinite. Hence the result.

(d) Since $\lim_{k\to\infty} \sum_{j=1}^{k} u_{-j} = \infty$ a.e. and $w_{n+1}(0) = (w_n(0) + u_n)^+ \geq w_n(0)+u_n$, a point is reached in almost every realization when $w_M(0)$ is positive for some $M \in \mathbb{N}$. Ultimately then, the queue builds up and never empties. Thereafter every customer must wait for service, and $w_{n+1}(0)$ reduces to $w_n(0) + u_n \; \forall\, n \geq N \geq M$, say.

Using Lemma 0.2.2 and Theorem 0.2.2

$$
\begin{aligned}
w_n(0) &= w_N(0) + \sum_{j=N}^{n-1} u_j \\
\Rightarrow \quad \frac{w_n(0)}{n} &= \frac{w_N(0)}{n} + \left(\tfrac{n-N}{n}\right) \tfrac{1}{n-N} \sum_{j=N}^{n-1} u_j \\
\Rightarrow \quad \lim_{n\to\infty} \frac{w_n(0)}{n} &= 0 + \lim_{n\to\infty} \tfrac{1}{n-N} \sum_{j=N}^{n-1} u_j \qquad \text{a.s.} \\
&= \mathrm{E}(u_0) \qquad \text{a.s.}
\end{aligned}
$$

A similar proof can be formulated for $\frac{w_n(\alpha)}{n}$, $\alpha > 0$. Hence the result.

□

Theorem 1.3.2 (Subcritical Queue)

Consider a queue with $\mathrm{E}(u_0) < 0$, or $\rho < 1$. Then

(a) the queue is convergent stable

(b) initial conditions do not affect this convergent stability nor do they alter the limiting distribution

(c) there exists precisely one a.e. finite stationary solution sequence $\langle \underline{m}_n \rangle$ of (1.1.1)

(d) $P[m_0 = 0] > 0$

(e) For any $w_1 = \alpha \geq 0$

$$\frac{w_n(\alpha)}{n} \underset{\infty}{\overset{n}{\to}} 0 \text{ a.s.}$$

<u>Proof</u>

$$\lim_{k\to\infty} \sum_{j=1}^{k} u_{-j} = \lim_{k\to\infty} k\left(\frac{1}{k}\sum_{j=1}^{k} u_{-j}\right) = \mathrm{E}(u_0) \lim_{k\to\infty} k = -\infty$$

where the strong law for stationary sequences (Theorem 0.2.2) is applied, as in Theorem 1.3.1.

(a) Since $\lim_{n\to\infty} x_n < \infty \Leftrightarrow \limsup_{n\to\infty} x_n < \infty$, we obtain $\limsup_{k\to\infty} \sum_{j=1}^{k} u_{-j} < \infty$ (w.p.1).

On the basis of Theorem 1.2.2(a), we deduce that the queue is convergent stable, when the initial condition is given by $w_1 = 0$.

(b) Consider any sequence $\langle w_n, n \in \mathbb{N} \rangle$ of waiting-times satisfying (1.1.1), and with w_1 a.e. finite. Suppose now that $\forall n \in \mathbb{N}\ w_n > 0$ (i.e., $w_n \neq 0$) with positive probability. Then (with positive probability)

$$\begin{array}{rrcll} & 0 < w_{n+1} & = & (w_n + u_n)^+ & \\ \Rightarrow & w_{n+1} & = & w_n + u_n & \\ \Rightarrow & w_{n+1} & = & w_1 + \sum_{j=1}^{n} u_j & \forall\, n, \text{ using Lemma 0.2.2} \\ \Rightarrow & \liminf_n w_{n+1} & = & w_1 + \liminf_n \sum_{j=1}^{n} u_j & \\ & & = & w_1 - \infty & \\ & & = & -\infty & \text{since } w_1 \text{ is a.e. finite} \\ \text{but} & \liminf_n w_{n+1} & \geq & 0 & \end{array}$$

Thus $\exists\ N_1 \in \mathbb{N}$ s.t. $w_{N_1} = 0$ w.p.1 (because $P[w_1 = \infty] = 0$ and $P\left[\liminf_n \sum_{j=1}^{n} u_j > -\infty\right] = 0$).

If, again, one were to suppose that $\forall\, n > N_1\ w_n \neq 0$ (with positive probability), a contradiction similar to that above would arise. Repetition of this argument leads to the conclusion that $P[w_n = 0 \text{ i.o.}] = 1$.

From the minimality of $\langle w_n^0 \rangle$ deduced in Theorem 1.2.1(c)

$$\begin{array}{rlcccccl} & w_n & \geq & & w_n^0 & \geq & 0 & \forall\, n \in \mathbb{N} \\ \Rightarrow & w_{N_1} & = & 0 \ \Rightarrow & w_{N_1}^0 & = & 0 & \\ \Rightarrow & w_{N_1+1} & = & & u_{N_1}^+ & = & w_{N_1+1}^0 & \text{substituting in (1.1.1)} \end{array}$$

Suppose $w_n = w_n^0$ for some $n \geq N_1$. Then

$$\begin{aligned} w_{n+1} &= (w_n + u_n)^+ \\ &= (w_n^0 + u_n)^+ \\ &= w_{n+1}^0 \\ \Rightarrow \quad w_n &= w_n^0 \qquad\qquad \forall\, n \geq N_1 \text{ (by induction)} \end{aligned}$$

In the above, if the sequence $\langle w_n \rangle$ has initial condition α, then $w_n(\alpha) = w_n^0$ $\forall\, n \geq N_1$, where N_1 is sufficiently large. The agreement of these sequences after some finite N_1 implies that neither the convergent stability nor the limiting waiting-time distribution can be affected by initial conditions.

(c) Application of the argument given in (b) implies that $\exists N_2 \in \mathbb{N}$ s.t. $\underline{m}_{N_2} = 0$ w.p.1 where, for convenience, $N_2 = \min\{n : \underline{m}_n = 0 \text{ w.p.1}\}$ so that $\underline{m}_n = w_n^0\ \forall\, n \geq N_2$.

Let $N_0 = \max(N_1, N_2)$. Then $\underline{m}_n = w_n^0 = w_n\ \forall n \geq N_0$. To understand this, suppose w.l.g. that $N_0 = N_1$. It follows that $\underline{m}_{N_1} = w_{N_1}^0$ (because $N_1 \geq N_2$) and the minimality of w_n^0 (Theorem 1.2.1(c))$\Rightarrow$

$$\begin{array}{ccccccccc} 0 & = & w_{N_1} & \geq & w_{N_1}^0 & = & \underline{m}_{N_1} & \geq & 0 \\ & \Rightarrow & w_{N_1} & = & w_{N_1}^0 & = & \underline{m}_{N_1} & & \\ & & & & \| & & & & \\ & & & & 0 & & & & \end{array}$$

From N_1 on, then, all three quantities will coincide for the same reason as in (b).

Let $\langle \underline{w}_n, n \in \mathbb{Z}\rangle$ be some other stationary sequence of a.e. finite random variables satisfying (1.1.1). The above discussion allows us to conclude that $\underline{w}_n$ and $\underline{m}_n$ eventually coincide. (Before they do coincide, we know that $\underline{w}_n \geq \underline{m}_n$ because of the minimality of $\underline{m}_n$ (Theorem 1.2.1(c))).

So, $L(\underline{m}_n) \overset{n}{\underset{\infty}{\to}} F$ (say) at points of continuity of $F \Rightarrow L(\underline{w}_n) \overset{n}{\underset{\infty}{\to}} F$. Now, using the fact that $\langle \underline{m}_n \rangle$ is an identically-distributed sequence (because it is stationary), we obtain $P[\underline{m}_n \leq x] = P[\underline{m}_0 \leq x] = \lim_{n\to\infty} P[\underline{m}_0 \leq x] = \lim_{n\to\infty} P[\underline{m}_n \leq x] = F \; \forall \, n \in \mathbb{Z}$. Similarly $P[\underline{w}_n \leq x] = F \; \forall \, n \in \mathbb{Z}$ so that $L(\underline{m}_n) = F = L(\underline{w}_n) \; \forall \, n \in \mathbb{Z}$. We have, however, noted that $\underline{m}_n \geq \underline{w}_n \Rightarrow \underline{m}_n = \underline{w}_n$ a.e. (Shiryayev [56, *p.* 183, Property H]) $\Rightarrow$ a unique stationary sequence satisfying (1.1.1).

(d) Suppose, to the contrary, that $P[m_0 = 0] = 0 = P[m_n = 0]$. Then $0 < m_{n+1} = (m_n + u_n)^+ \Rightarrow m_{n+1} = m_n + u_n$ and if we use Lemma 0.2.2, we obtain

$$\begin{array}{ccccc} \frac{1}{n}\sum_{j=1}^{n} u_j & = & \frac{m_{n+1}}{n} - \frac{m_1}{n} & & \\ \downarrow & & \downarrow & & \\ \mathrm{E}(u_0) & = & 0 & & \text{a.e.} \end{array}$$

by the strong law (Theorem 0.2.2). Thus the assumption that $\mathrm{E}(u_0) < 0$ has been contradicted. So we must have that $P[m_0 = 0] > 0$.

To show that $\frac{m_{n+1}}{n} \overset{n}{\underset{\infty}{\to}} 0$, in fact, requires knowledge of the result in (e). From (a), Theorem 1.2.2(a) and Remark 1.2.5, we deduce that $\langle m_n, n \in \mathbb{Z}\rangle$ is a sequence of honest random variables. Thus there exists some finite α (or a.e. finite $\alpha(\omega)$) such that $m_1 \leq \alpha = w_1(\alpha)$ w.p.1.

As before, we conclude that $0 \leq m_n \leq w_n(\alpha)\ \forall\, n \in \mathbb{N}$. From (e) $0 \leq \frac{m_n}{n} \leq \frac{w_n(\alpha)}{n} \overset{n}{\underset{\infty}{\to}} 0$ a.s. $\Rightarrow \frac{m_n}{n} \overset{n}{\underset{\infty}{\to}} 0$ a.s.

(e) The proof that $\frac{w_n(\alpha)}{n} \overset{n}{\underset{\infty}{\to}} 0$ used here is actually based on the analogous result of Theorem 1.3.1(d) and is valid both for $\mathrm{E}(u_0) < 0$ and $\mathrm{E}(u_0) = 0$ (the latter case receiving greater attention in Theorem 1.3.3).

We proceed as follows. First, let $\langle u'_n \rangle$ be that sequence determined by $u'_n = u_n I_{[u_n \geq -\beta]} \geq u_n$, where $\beta > 0$. The truncation, if necessary, is applied to make certain that $\mathrm{E}(u'_n) > -\infty$. Then $w'_{n+1}(\alpha) = (w'_n(\alpha) + u'_n)^+ \geq (w_n(\alpha) + u_n)^+ = w_{n+1}(\alpha)\ \forall\, n$.

Next define $u''_n = u'_n + c > u'_n \geq u_n$ for some $c > 0$, in order to guarantee that $0 < \mathrm{E}\,(u''_n) \leq \varepsilon\ \forall\, \varepsilon > 0$. Obviously $w''_n(\alpha) \geq w_n(\alpha)\ \forall\, n$. From Theorem 1.3.1(d), we deduce that

$$0 \leq \liminf_n \frac{w_n(\alpha)}{n} \leq \limsup_n \frac{w_n(\alpha)}{n} \leq \mathrm{E}(\varepsilon) = \varepsilon\ \forall \varepsilon > 0$$

because the same is true for $\langle w''_n(\alpha) \rangle$. Thus $\lim\limits_{n \to \infty} \dfrac{w_n(\alpha)}{n} = 0$ a.e.

□

Theorem 1.3.3 (Critical Queue)

Consider a queue for which $\mathrm{E}(u_0) = 0$, or $\rho = 1$.

(a) (i) $m_{n+1} = m_n + u_n$ (1.3.1)

(ii) $\lim\limits_{n\to\infty} \frac{w_n(\alpha)}{n} = 0$

(iii) $$m_0 = \limsup_k \sum_{j=1}^{k} u_{-j} = \sup_{k\in\mathbb{N}} \sum_{j=1}^{k} u_{-j} = -\inf_{k\in\mathbb{Z}_\oplus} \sum_{j=0}^{k} u_j = -\liminf_k \sum_{j=0}^{k} u_j \tag{1.3.2a}$$

$$0 \geq \liminf_k \sum_{j=1}^{k} u_{-j} = \inf_{k\in\mathbb{N}} \sum_{j=1}^{k} u_{-j} = -\sup_{k\in\mathbb{Z}_\oplus} \sum_{j=0}^{k} u_j = -\limsup_k \sum_{j=0}^{k} u_j \tag{1.3.2b}$$

(b) If we assume m_0 to be dishonest, then

(i) the queue is unstable

(ii) initial conditions do not affect this instability

(iii) there exists no a.e. finite *stationary* solution to (1.1.1)

(iv) any a.e. finite sequence $\langle w_n, n \in \mathbb{N}\rangle$ satisfying (1.1.1) has $P[w_n = 0 \text{ i.o.}]=1$

(c) If, finally, m_0 is supposed to be honest, then

(i) the queue is stable

(ii) initial conditions affect the asymptotic distributions but not the stability

(iii) there exists a continuum of a.e. finite stationary solutions to (1.1.1)

(iv) $\liminf_n m_n = 0 = \liminf_n w_n^0$ w.p.1

PROOF

(a) From Remark 1.2.5, we know that, if m_0 is dishonest, then $\langle m_n \rangle$ is a sequence of a.e. infinite random variables.

(i) If follows that $m_{n+1} = m_n + u_n$ is trivially valid when m_0 is dishonest.

This equation is, however, equally valid when m_0 is honest. To see why this is the case write

$$\begin{aligned} m_{n+1} &= (m_n + u_n)^+ \geq m_n + u_n \\ \Rightarrow \quad \tilde{u}_n &= m_{n+1} - m_n \geq u_n \end{aligned} \tag{1.3.3}$$

From (1.3.3), we note that $\langle \tilde{u}_n \rangle$ is a stationary sequence because it is a function of the stationary random variables, m_{n+1} and m_n. Next, $\mathrm{E}(u_n) = 0$ so that analysis i.t.o. u_n^+ and u_n^- leads to the conclusion that $E|u_n| < \infty$ (and exists) (Shiryayev [56, *p.* 180]) so the strong law (Theorem 0.2.2) is valid for $\langle u_n \rangle$. In addition (1.3.3) $\Rightarrow 0 \leq E(\tilde{u}_n^-) \leq E(u_n^-) < \infty$ which, of course, implies that $E(\tilde{u}_n)$ exists (Shiryayev [56, *p.* 180, Definition 2]). These facts, together with (1.3.3) and the strong law (again), including Remark 0.2.8, allow us to deduce the following

$$\begin{array}{lcccccc} & \frac{m_{n+1}}{n} - \frac{m_1}{n} & = & \frac{1}{n}\sum_{j=1}^{n} \tilde{u}_j & \geq & \frac{1}{n}\sum_{j=1}^{n} u_j & \\ \Rightarrow & 0 & = & \lim_n \frac{1}{n}\sum_{j=1}^{n} \tilde{u}_j & \geq & E(u_0) & = \; 0 \text{ (a.s.)} \\ \text{or} & E(\tilde{u}_0) & = & 0 & = & E(u_0) & \text{a.s} \end{array}$$

where we observed that $\frac{m_n}{n} \to 0$ a.s. for essentially the same reason that $\frac{m_n}{n} \to 0$ in Theorem 1.3.2. Then $\tilde{u}_n - u_n \geq 0$ together with the fact derived above that $E(\tilde{u}_n - u_n) = 0$ (a.s.)$\Rightarrow$ (Shiryayev [56, *p.* 183, Property H]) $\tilde{u}_n = u_n$ a.s. $(\forall\, n)$. As a result (1.3.3) reduces to $m_{n+1} = m_n + u_n$ in the honest case as well. (The positive part becomes redundant when $\rho = 1$.)

(ii) The fact that $\frac{w_n(\alpha)}{n} \to 0$ for the case $\rho = 1$ was shown together with the case $\rho < 1$ in Theorem 1.3.2(e).

(iii) Obviously the equation (1.2.9) derived in Theorem 1.2.1:

$$m_n = \left[\sup_{k\in\mathbb{N}} \sum_{j=1}^{k} u_{n-j}\right]^+$$

is equally valid here, but can be refined in view of (a)(i) of the current theorem. The calculations that led to (1.2.9) would, in our case, proceed without the positive part, resulting in

$$0 \leq m_n = \sup_{k \in \mathbb{N}} \sum_{j=1}^{k} u_{n-j} \tag{1.3.4}$$

when $\rho = 1$.

Using (1.3.1), we get

$$\begin{aligned} \sum_{j=0}^{k} (m_{j+1} - m_j) &= \sum_{j=0}^{k} u_j \\ \text{or} \quad m_{k+1} + \sum_{j=0}^{k} (-u_j) &= m_0 \end{aligned} \tag{1.3.5}$$

Then

$$\begin{aligned} \sup_{k \in \mathbb{N}} \sum_{j=1}^{k} u_{-j} &= m_0 && \text{from (1.3.4)} \\ &= \sup_{k \in \mathbb{Z}_\oplus} m_0 \\ &= \sup_{k \in \mathbb{Z}_\oplus} \left[m_{k+1} + \sum_{j=0}^{k} (-u_j) \right] && \text{from (1.3.5)} \\ &\geq \inf_{k \in \mathbb{N}} m_k + \sup_{k \in \mathbb{Z}_\oplus} \sum_{j=0}^{k} (-u_j) && \text{from Lemma 0.2.1} \\ &\geq \sup_{k \in \mathbb{Z}_\oplus} \sum_{j=0}^{k} (-u_j) && \text{since } \inf m_n \geq 0 \end{aligned}$$

$$\tag{1.3.6}$$

Consider next that queue with sequence $\langle -u_{-n-1}\rangle$ which is stationary because $\langle u_n\rangle$ is, and also has (common) mean zero. Consequently the above argument holds true for such a queue, that is

$$\begin{aligned}
\sup_{k\in\mathbb{Z}_\oplus}\sum_{j=0}^{k}(-u_j) &= \sup_{k\in\mathbb{N}}\sum_{j=1}^{k}(-u_{j-1}) \\
&\geq \sup_{k\in\mathbb{Z}_\oplus}\sum_{j=0}^{k}\left(u_{-(j+1)}\right) \quad (=) \quad \text{from (1.3.6)} \\
&\geq \sup_{k\in\mathbb{N}}\sum_{j=1}^{k}u_{-j} \\
&\geq m_0 \qquad \text{from (1.3.4)}
\end{aligned} \tag{1.3.7}$$

Combination of (1.3.6) and (1.3.7) means that

$$\sup_{k\in\mathbb{Z}_\oplus}\sum_{j=0}^{k}(-u_j) \overset{(1.3.7)}{\geq} m_0 \overset{(1.3.6)}{\geq} \sup_{k\in\mathbb{Z}_\oplus}\sum_{j=0}^{k}(-u_j)$$

so, in fact, the inequalities in (1.3.6) and (1.3.7) are equalities, and are summarised by

$$\begin{aligned}
\sup_{k\in\mathbb{Z}_\oplus}\sum_{j=0}^{k}(-u_j) &= \sup_{k\in\mathbb{N}}\sum_{j=1}^{k}u_{-j} && \text{from (1.3.7)} \\
&= m_0 && \text{from (1.3.4)} \\
&= \inf_{k\in\mathbb{N}} m_k + \sup_{k\in\mathbb{Z}_\oplus}\sum_{j=0}^{k}(-u_j) && \text{from (1.3.6)} \\
&= \inf_{k\in\mathbb{N}} m_k + m_0 && \text{from (1.3.7)}
\end{aligned} \tag{1.3.8}$$

which implies that $\inf_{k \in \mathbb{N}} m_k = 0$ if m_0 is honest. (At this stage, all we can say if m_0 is dishonest (so a.e. infinite) is that $\inf_{k \in \mathbb{N}} m_k$ is some nonnegative value.)

For this honest case, consideration of the sequence $\langle u_{n+N} \rangle$ in a similar way to that given in (1.3.6)–(1.3.8) allows us to conclude that

$$\begin{aligned} \inf_{k \geq N} m_k &= 0 \quad \forall N \in \mathbb{N} \\ \Rightarrow \quad \liminf_k m_k &= 0 \end{aligned} \tag{1.3.9}$$

Taking the lim inf throughout in (1.3.5) yields

$$m_0 + \liminf_k \sum_{j=0}^{k} u_j = 0 \tag{1.3.10}$$

Returning to (1.3.8), this means that

$$\begin{aligned} \limsup_k \sum_{j=1}^{k} u_{-j} &= -\liminf_k \sum_{j=0}^{k} u_j && \text{from (1.3.8)} \\ &= m_0 && \text{from (1.3.10)} \\ &= \sup_{k \in \mathbb{N}} \sum_{j=1}^{k} u_{-j} && \text{from (1.3.8)} \\ &= -\inf_{k \in \mathbb{N}} \sum_{j=1}^{k} (-u_{-j}) && \\ &= -\inf_{k \in \mathbb{Z}_{\oplus}} \sum_{j=0}^{k} u_j && \text{from (1.3.8)} \end{aligned} \tag{1.3.11}$$

with all the terms nonnegative because $m_0 \geq 0$. (1.3.2a) may then be extracted from (1.3.11) (initially for the case of m_0 honest).

To establish (1.3.2b) merely requires that (1.3.2a) be applied to the sequence $\langle -u_n \rangle$ (obviously having the properties we required for $\langle u_n \rangle$). This gives us the equality of the expressions on the left-hand side of the system of equations below.

$$\begin{aligned}
\limsup_k \sum_{j=1}^{k} (-u_{-j}) &= -\liminf_k \sum_{j=1}^{k} u_{-j} \\
\sup_{k\in\mathbb{N}} \sum_{j=1}^{k} (-u_{-j}) &= -\inf_{k\in\mathbb{N}} \sum_{j=1}^{k} u_{-j} \\
-\inf_{k\in\mathbb{Z}_\oplus} \sum_{j=0}^{k} (-u_j) &= \sup_{k\in\mathbb{Z}_\oplus} \sum_{j=0}^{k} u_j \\
-\liminf_k \sum_{j=0}^{k} (-u_j) &= \limsup_k \sum_{j=0}^{k} u_j
\end{aligned}$$

Thus, on multiplying the expressions on the right through by -1, we obtain (1.3.2b).

Although (1.3.9) is valid (only) in the case of m_0 honest, clearly (1.3.2a) and (1.3.2b) hold when m_0 is either honest or dishonest. In the latter case, (1.3.2a) represents a string of a.e. (positively) infinite quantities, and (1.3.2b), a.e. negatively infinite random variables.

Expressions for m_n analogous to (1.3.2a) and (1.3.2b) are easily obtainable by a suitable shift. For instance

$$m_n = \limsup_k \sum_{j=1}^{k} u_{n-j} = -\inf_{k \in \mathbb{Z}_\oplus} \sum_{j=0}^{k} u_{n+j}$$

(b) We now investigate properties of the queue when m_0 is dishonest.

(i) In Theorem 1.2.2(a), without considering the nature of ρ, we deduced that, if m_0 is dishonest, then the queue is unstable.

(ii) The reasoning for this is essentially the same as in Theorem 1.3.1.

(iii) Since $\langle m_n, n \in \mathbb{Z} \rangle$ is a minimal stationary solution to (1.3.1) (Theorem 1.2.1(c)), $x_n \geq m_n$ for some other stationary solution sequence $\langle x_n, n \in \mathbb{Z} \rangle$, but m_n is a.e. infinite so that x_n is also a.e. infinite and no a.e. finite *stationary* solution to (1.3.1) exists.

(iv) Although no a.e. finite *stationary* solution to (1.1.1) exists, Theorem 1.2.1(c) does not exclude the possibility of a.e. finite solutions.

Suppose, to the contrary, that $\forall\, n > N \; w_n > 0$. Then, if we use

Lemma 0.2.2, we obtain

$$\begin{aligned}
w_{n+1} &= w_N + \sum_{j=N}^{n} u_j && (n \geq N) \\
\Rightarrow \quad \liminf_n w_{n+1} &= w_N + \liminf_n \sum_{j=N}^{n} u_j \\
&= w_N + \liminf_n \sum_{j=0}^{n-N} u_{N+j} \\
&= w_N - m_N && \text{from (1.3.2a)} \\
&= -\infty
\end{aligned}$$

because $m_N = \infty$ a.e. and w_N is a.e. finite

but $\liminf_n w_{n+1} \geq 0$

Consequently, $\exists\, M > N$ s.t. $w_M = 0$. In a similar way to that in Theorem 1.3.2(b), we deduce that there are infinitely many n for which $w_n = 0$, or $P\left[w_n = 0 \text{ i.o.}\right] = 1$.

(c) Finally we examine the case of m_0 honest.

(i) The stability of the queue is deduced in much the same way as instability was deduced in (b)(i).

(ii) We know that $m_1 \geq 0 \Rightarrow m_1 + \alpha \geq \alpha = w_1(\alpha)$.
If $m_n + \alpha \geq w_n(\alpha)$ for some $n \in \mathbb{N}$, then

$$\begin{aligned}
m_{n+1} + \alpha &= m_n + u_n + \alpha && \text{from (1.3.1)} \\
&= (m_n + u_n + \alpha)^+ \\
&\geq (w_n(\alpha) + u_n)^+ && (=) \\
&\geq w_{n+1}(\alpha) \\
\Rightarrow \quad m_n + \alpha &\geq w_n(\alpha) && \forall\, n \in \mathbb{N} \text{ (by induction)}
\end{aligned}$$

$$
\begin{aligned}
\Rightarrow \qquad P[m_0 \leq x - \alpha] &= P[m_n \leq x - \alpha] \\
&\leq P[w_n(\alpha) \leq x] \\
\Rightarrow \qquad P[m_0 \leq x - \alpha] &\leq \lim_{n\to\infty} P[w_n(\alpha) \leq x] \\
\Rightarrow \qquad 1 &= \lim_{x\to\infty} P[m_0 \leq x - \alpha] \\
&\leq \lim_{x\to\infty}\lim_{n\to\infty} P[w_n(\alpha) \leq x] \\
\Rightarrow \quad \lim_{x\to\infty}\lim_{n\to\infty} P[w_n(\alpha) \leq x] &= 1
\end{aligned}
$$

$\Rightarrow$ the queue is still stable

We now demonstrate how initial conditions can affect the limiting distributions. Suppose $\alpha = w_1(\alpha) = m_1 + c$ $(c > 0)$ a.e. Then $w_n(\alpha) = m_n + c$ for some $n \in \mathbb{N} \Rightarrow$

$$
\begin{aligned}
w_{n+1}(\alpha) &= (w_n(\alpha) + u_n)^+ \\
&= (m_n + c + u_n)^+ \\
&= (m_{n+1} + c)^+ && \text{from } (1.3.1) \\
&= m_{n+1} + c && m_{n+1} + c > 0 \\
\Rightarrow \qquad w_n(\alpha) &= m_n + c && \forall\, n \in \mathbb{N} \text{ (induct.)} \\
\Rightarrow \quad \lim_{n\to\infty} P[w_n(\alpha) \leq x] &= \lim_{n\to\infty} P[m_n + c \leq x] \\
&= P[m_0 + c \leq x] \\
&= P[m_n \leq x - c]
\end{aligned}
$$

and we show next that the $\langle m_n + c\rangle$ is also a solution sequence.

(iii) From (1.3.1)

$$
\begin{aligned}
& m_{n+1} + c = m_n + c + u_n \\
\Rightarrow\; & \langle m_n + c \rangle \text{ also satisfies (1.3.1), certainly } \forall\, c \geq 0 \\
\Rightarrow\; & \text{a continuum of solutions, so that the uniqueness} \\
& \text{which held when } \rho < 1 \text{ is no longer valid}
\end{aligned}
$$

Conversely, if $\langle \underline{w}_n \rangle$ is some stationary sequence satisfying (1.1.1), then the same reasoning as before implies that $\langle \underline{w}_n \rangle$ satisfies (1.3.1) and

$$
\begin{array}{lrcl}
 & \underline{w}_{n+1} & = & \underline{w}_n + \underline{u}_n \\
\text{but} & \underline{m}_{n+1} & = & \underline{m}_n + \underline{u}_n \\
\Rightarrow & \underline{w}_{n+1} - \underline{m}_{n+1} & = & \underline{w}_n - \underline{m}_n \\
\Rightarrow & x \quad = x_n & = & \underline{w}_n - \underline{m}_n
\end{array}
$$

is a random variable independent of n;
it is also nonnegative because of the minimality of $\underline{m}_n$

Then $\underline{w}_n = \underline{m}_n + x$.

(iv) In (1.3.9), we noted that $\liminf_n m_n = 0$ (i.e., for a.e. ω). From Theorem 1.2.1(b), $0 \leq w_n^0 \leq m_n \Rightarrow \liminf_n w_n^0 = 0$ w.p.1.

□

Remark 1.3.1

If equation (1.3.1) were to describe a random walk on $\mathbb{R}_\oplus$, then the usual requirement would be that $\langle u_n \rangle$ is an IID sequence. Our stipulation is that $\langle u_n \rangle$ is stationary ergodic. Nonetheless, if we borrow from the language of random walks and Markov chains, we might say that (1.1.1) implies a reflecting barrier at 0, and (1.3.1) implies that no overshoot can occur in the stationary state.

In addition, an honest sequence $\langle m_n \rangle$ satisfying (1.3.1) physically would represent a position of neutral equilibrium, as opposed to the stable equilibrium characterising the subcritical queue of Theorem 1.3.2.

Remark 1.3.2 (The Effect of Nonergodicity/Metric Intransitivity) Up until now we have required that the sequence $\langle(s_n, \tau_n)\rangle$, and hence $\langle u_n \rangle$, be not only stationary but also ergodic. If we dispense with the latter requirement, then results, analogous to those already proved, may be deduced.

A fundamental difference, of course, lies in the strong law (Theorem 0.2.2) where $\mathrm{E}(u_0 \mid \mathcal{I}_u)$ replaces $\mathrm{E}(u_0)$. A consequence of this is that m_0 may be finite with probability $p \in (0,1)$.

Part (d) of Theorem 1.3.1 would be replaced by

On the set $A = \{\omega : \mathrm{E}(u_0 \mid \mathcal{I}_u)(\omega) > 0\}$, $\frac{w_n(\alpha)}{n} \to \mathrm{E}(u_0 \mid \mathcal{I}_u)$ a.s. $\forall \alpha \geq 0$.

As regards Theorem 1.3.2, one could conclude, as a necessary condition, that

If $P[\mathrm{E}(u_0 \mid \mathcal{I}_u) < 0] = 1$, then the queue is convergent stable.

Let $\mathcal{E} = \{\omega : \mathrm{E}(u_0 \mid \mathcal{I}_u)(\omega) < 0\}$. Now $\mathrm{E}(u_0 \mid \mathcal{I}_u)$ is $\mathcal{I}_u$-measurable (Loève [46, *p.* 7] or Shiryayev [56, *p.* 211]) thus $\{\omega : \mathrm{E}(u_0 \mid \mathcal{I}_u)(\omega) \leq x\} \in \mathcal{I}_u \; \forall x \in \mathbb{R}$ which implies that $\mathcal{E}$ is an invariant set so $\mathcal{E}$ is an invariant random variable (using Definition 0.2.4).

In order to deal with this situation, we define a new queue determined by

$\langle \hat{u}_n \rangle$, where

$$\begin{aligned} \hat{u}_n &= u_n I_{\mathcal{E}} - c(1 - I_{\mathcal{E}}), \quad c > 0 \\ \Rightarrow \quad \hat{u}_n(\omega) &= \begin{cases} u_n(\omega) & \omega \in \mathcal{E} \\ -c & \omega \in \bar{\mathcal{E}} \end{cases} \end{aligned}$$

$$\Rightarrow \quad \hat{u}_n \text{ and } u_n \text{ agree on } \mathcal{E}$$

For a measure-preserving transformation T

$$\hat{u}_n(T\omega) = u_n(T\omega) I_{\mathcal{E}}(T\omega) - c(1 - I_{\mathcal{E}}(T\omega)) = \hat{u}_n(\omega)$$

because of the stationarity of $\langle u_n \rangle$ and the invariance of $I_{\mathcal{E}}$. Thus $\langle \hat{u}_n \rangle$ is a stationary (hence an identically-distributed) sequence. Also $\mathrm{E}(\hat{u}_0) = \mathrm{E}\,(u_0 I_{\mathcal{E}}) - cP(\bar{\mathcal{E}})$ so this expectation obviously exists. Therefore, from the strong law (Theorem 0.2.2),

$$\begin{aligned} \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} \hat{u}_j &= \left(\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} u_j \right) I_{\mathcal{E}} - c(1 - I_{\mathcal{E}}) \\ \Rightarrow \qquad \mathrm{E}\,(\hat{u}_0 \mid \mathcal{I}_{\hat{u}}) &= \mathrm{E}\,(u_0 \mid \mathcal{I}_u)\, I_{\mathcal{E}} - c(1 - I_{\mathcal{E}}) \ \text{ a.s.} \\ \Rightarrow \quad P\,[\mathrm{E}\,(\hat{u}_0 \mid \mathcal{I}_{\hat{u}}) < 0] &= P(\mathcal{E}) + P(\bar{\mathcal{E}}) \\ &= 1 \end{aligned}$$

Hence $\langle \hat{u}_n \rangle$ has the desired properties, and results will apply to the queue involving $\langle u_n \rangle$ on $\mathcal{E}$.

In the case of Theorem 1.3.3 we have, for example, that equation (1.3.2a) is valid only on $\mathcal{E}' = \{\omega : \mathrm{E}(u_0 \mid \mathcal{I}_u) = 0\}$. The situation is dealt with similarly to that considered for Theorem 1.3.2 by defining $\hat{u}_n = u_n I_{\mathcal{E}'}$.

Remark 1.3.3 (Honesty/Dishonesty and Stability/Instability)
In Theorem 1.2.2, a criterion for stability/instability based on $\limsup \sum u_{-j}$ (and, in effect, the honesty/dishonesty of m_0) was provided. This condition was applied successfully in Theorem 1.3.1 to prove the instability of the supercritical queue and in Theorem 1.3.2 to deduce the (convergent) stability of the subcritical queue. Until now, though, conclusions for the critical queue (Theorem 1.3.3) were dependent upon the honesty or dishonesty of m_0. No satisfactory method was provided to determine which of these cases applied. A partial solution to this problem encountered with the critical queue is offered in the sequel.

Remark 1.3.4 (The critical queue: $\langle \boldsymbol{u}_n \rangle$ a sequence of independent random variables)
If the (critical) queue is to be stable we know (from Theorem 1.2.2) that m_0 must be honest. In that case, we may consider characteristic function $\phi_m(t) = \mathrm{E}\left(e^{itm_n}\right)$ ($\forall\, n$ because of stationarity) and $\phi_u(t) = \mathrm{E}\left(e^{itu_0}\right)$. Then because (1.3.1) implies that $m_1 = m_0 + u_0, \quad \forall\, t$

$$\phi_m(t) = \mathrm{E}\left(e^{itm_1}\right) = \phi_m(t)\phi_u(t)$$

where we use the independence of m_0 and u_0 implied by $m_0 = \sup_{k \in \mathbb{N}} \sum_{j=1}^{k} u_{-j}$ in (1.3.2a).

Hence $\mathrm{E}\left(e^{itu_0}\right) = 1 \; \forall t$ which must be consistent with the fact that $u_0 = 0$ a.e.

On the basis of Remark 1.3.4 and Lindley [45, *p.* 281], we are able to state the following theorem.

Theorem 1.3.4
Consider a $GI/GI/1$ queue for which $\rho = 1$. Then the queue is stable iff $u_0 = 0$ a.e.

After we have defined a useful concept, we will be in a position to generalise the situation somewhat.

Definition 1.3.1 (Conditional Independence)
Let $\mathcal{B}_i$ $(i = 1, 2)$ and $\mathcal{B}$ represent sigma-fields. Then $\mathcal{B}_1$ and $\mathcal{B}_2$ shall be said to be *conditionally independent given* $\mathcal{B}$ iff

$$P[B_1 \cap B_2 \mid \mathcal{B}] = P[B_1 \mid \mathcal{B}]P[B_2 \mid \mathcal{B}] \text{ a.s. } \forall\, B_i \in \mathcal{B}_i \;\; (i = 1, 2)$$

In particular, random sequence $\langle \xi_n, n \in \mathbb{Z} \rangle$ shall be said to have *conditional independence of its past and future* (CIPF) iff $\mathcal{B}_1^{(0)} = \sigma\left(\langle \xi_n, n \in \mathbb{Z}_{\oplus} \rangle\right)$ and $\mathcal{B}_2^{(0)} = \sigma\left(\langle \xi_n, n \in \mathbb{Z}_{-} \rangle\right)$ are conditionally independent given some sigma-field $\mathcal{B}^{(0)}$.

Remark 1.3.5 (Conditional Independence)

(a) The conditional independence of past and future given the present is effectively the Markov property. (See Loève [46, *p.* 290].)

(b) If in Definition 1.3.1 $\langle \xi_n \rangle$ is a stationary sequence, then CIPF is equivalent to the conditional independence of $\mathcal{B}_1^{(N)} = \sigma\left(\langle \xi_n, n \geq N \rangle\right)$ and $\mathcal{B}_2^{(N)} = \sigma\left(\langle \xi_n, n < N \rangle\right)$ given sigma-field $\mathcal{B}^{(N)}$ for some $N \in \mathbb{Z}$.

Theorem 1.3.5 (Stability Criterion)

Consider a critical $G/G/1$ queue (so $\rho = 1$), for which $\mathcal{B}_1^{(n)} = \sigma(u_n, u_{n+1}, \ldots)$ and $\mathcal{B}_2^{(n)} = \sigma(u_{n-1}, u_{n-2}, \ldots)$ are conditionally independent given $\mathcal{B}^{(n)}$. Then

(a) m_n is measurable w.r.t. $\mathcal{B}^{(n)}$, which is (almost surely) given by

$$\mathcal{B}^{(n)} = \begin{cases} \{\emptyset, \Omega, B_0, \overline{B}_0\} & \text{if } \exists\, x_0 \in \mathbb{R}_\oplus \text{ s.t. } B_0 = \{m_n \leq x_0\} \notin \{\emptyset, \Omega\} \\ \{\emptyset, \Omega\} & \text{otherwise} \end{cases}$$

(b) the queue is unstable unless u_n is measurable w.r.t. compound sigma-field $\left(\mathcal{B}^{(n)}, \mathcal{B}^{(n+1)}\right)$.

PROOF

(a) Using (1.3.2a), we obtain

$$m_n = \begin{cases} \sup\limits_{k \in \mathbb{N}} \sum\limits_{j=1}^{k} u_{n-j} \\ -\inf\limits_{k \in \mathbb{Z}_\oplus} \sum\limits_{j=0}^{k} u_{n+j} \end{cases}$$

Now the infimum and supremum terms are measurable w.r.t. $\mathcal{B}_1^{(n)}$ and $\mathcal{B}_2^{(n)}$ respectively. Hence m_n is measurable w.r.t. both $\mathcal{B}_1^{(n)}$ and $\mathcal{B}_2^{(n)}$ so that, using the CIPF property, we get (for each $x \in \mathbb{R}$)

$$\begin{aligned}
\left[P[\{m_n \le x\} \mid \mathcal{B}^{(n)}]\right]^2 &= P[\{m_n \le x\} \mid \mathcal{B}^{(n)}] \cdot P[\{m_n \le x\} \mid \mathcal{B}^{(n)}] \\
&= P[\{m_n \le x\} \cap \{m_n \le x\} \mid \mathcal{B}^{(n)}] \\
&= P[\{m_n \le x\} \mid \mathcal{B}^{(n)}] \\
\Rightarrow \qquad P[\{m_n \le x\} \mid \mathcal{B}^{(n)}] &= 0 \text{ or } 1 \\
\Rightarrow \qquad P[\{m_n \le x\} \mid B] &= 0 \text{ or } 1 \quad \forall\, B \in \mathcal{B}^{(n)} \\
\text{but} \qquad B \in \mathcal{B}^{(n)} &\Rightarrow \overline{B} \in \mathcal{B}^{(n)} \\
\Rightarrow \qquad P[\{m_n \le x\} \mid \overline{B}] &= 0 \text{ or } 1 \quad \forall\, B \in \mathcal{B}^{(n)}
\end{aligned}$$

Case 1 : $P[\{m_n \le x\} \mid B] = 0$ implying

$$B \subseteq \{\overline{m_n \le x}\} \quad \text{or} \quad \{m_n \le x\} \subseteq \overline{B} \qquad \text{(a.s.)}$$

Case 1.1 : $P[\{m_n \le x\} \mid \overline{B}] = 0$

$$\Rightarrow \overline{B} \subseteq \{\overline{m_n \le x}\} \quad \text{or} \quad \{m_n \le x\} \subseteq B$$

so $\{m_n \le x\} = \emptyset$ (a.s.)

Case 1.2 : $P[\{m_n \le x\} \mid \overline{B}] = 1$

$$\overline{B} \subseteq \{m_n \le x\} \quad \text{or} \quad \{\overline{m_n \le x}\} \subseteq B$$

so $\{m_n \le x\} = \overline{B}$ (a.s.)

$$Case \quad 2: \quad P[\{m_n \leq x\} \mid B] = 1$$

similarly leads to

$$\begin{aligned} Case\ 2.1: \quad \{m_n \leq x\} &= B \quad \text{(a.s.)} \\ Case\ 2.2: \quad \{m_n \leq x\} &= \Omega \quad \text{(a.s.)} \end{aligned}$$

The measurability of m_n w.r.t. $\mathcal{B}^{(n)}$ is therefore immediate.

If $\exists\, x_0 \in \mathbb{R}$ for which $\{m_n \leq x_0\} \notin \{\emptyset, \Omega\}$ (which means that $x_0 \in \mathbb{R}_\oplus$ because $\{m_n \leq x\} = \emptyset \ \forall\, x \in \mathbb{R}_-$), then obviously $\{m_n \leq x_0\} \in \mathcal{B}^{(n)}$ (hence $\overline{\{m_n \leq x_0\}} \in \mathcal{B}^{(n)}$) in addition to $\emptyset$ and Ω.

Suppose, moreover, that $\exists\, x_1 \neq x_0$ for which $\{m_n \leq x_1\} \notin \{\emptyset, \Omega\}$. This would exclude cases 1.1 and 2.2. Consequently

$$\{m_n \leq x_1\} = B \text{ or } \overline{B} \ \ \forall\, B\ (\notin \{\emptyset, \Omega\}) \ \in \mathcal{B}^{(n)}$$

and since $\{m_n \leq x_0\} \in \mathcal{B}^{(n)} \setminus \{\emptyset, \Omega\}$

$$\{m_n \leq x_1\} \quad = \quad \{m_n \leq x_0\} \text{ or } \overline{\{m_n \leq x_0\}}$$

In other words $\mathcal{B}^{(n)} = \{\emptyset, \Omega\}$ or $\mathcal{B}^{(n)} = \{\emptyset, \Omega, B_0, \overline{B}_0\}$ (a.s.), where $B_0 = \{m_n \leq x_0\} \notin \{\emptyset, \Omega\}$.

(b) The measurability of m_n w.r.t. $\mathcal{B}^{(n)}$ is not a restriction if m_n is dishonest. Otherwise (if m_n is honest), write u_n as $m_{n+1} - m_n$, in accordance with (1.3.1). From (a) we know that m_{n+1} is measurable w.r.t. $\mathcal{B}^{(n+1)}$ and m_n is measurable w.r.t. $\mathcal{B}^{(n)}$. Thus, if $\langle m_n \rangle$ is a sequence of proper random variables, then u_n necessarily must be measurable w.r.t. the compound sigma-field $\left(\mathcal{B}^{(n)}, \mathcal{B}^{(n+1)}\right)$.

$\square$

Remark 1.3.6

Theorem 1.3.5 provides a necessary, but not a sufficient, condition for stability/honesty of m_0. For instance, if $\langle m_n \rangle$ satisfies the given conditions, then so does $\langle -m_n \rangle$ which cannot do as a waiting time sequence.

Corollary 1.3.5*a*

Consider a queue with $\langle s_n \rangle$ and $\langle \tau_n \rangle$ independent sequences, (at least) one of which is formed of non-constant mutually independent random variables. Then the queue is stable iff $\rho < 1$.

PROOF

First suppose that $\langle \tau_n \rangle$ is a sequence of non-constant mutually independent random variables. We claim that a suitable candidate for the "given sigma-field" $\mathcal{B}^{(m)}$ is $\sigma\left(\langle s_n, n \in \mathbb{Z} \rangle\right)$, and verify this fact below.

$$\begin{aligned}
& L\left[\sigma\left(u_m, u_{m+1}, \ldots\right), \sigma\left(u_{m-1}, u_{m-2}, \ldots\right) \mid \sigma\left(\ldots, s_{m-2}, s_{m-1}, s_m, s_{m+1}, \ldots\right)\right] \\
= \; & L\left[\sigma\left(s_m - \tau_m, \ldots\right), \sigma\left(s_{m-1} - \tau_{m-1}, \ldots\right) \mid \sigma\left(\ldots, s_{m-1}, s_m, \ldots\right)\right] \\
= \; & L\left[\sigma\left(s_m - \tau_m, \ldots\right) \mid \sigma\left(\ldots, s_m, \ldots\right)\right] \cdot \\
& L\left[\sigma\left(s_{m-1} - \tau_{m-1}, \ldots\right) \mid \sigma\left(\ldots, s_{m-1}, \ldots\right)\right]
\end{aligned}$$

where, in the last step, we use the fact that $\langle \tau_n \rangle$ is a sequence of independent random variables.

Suppose that m_0 is honest/the queue is stable. Then we are able to conclude from Theorem 1.3.5 that $u_0 = s_0 - \tau_0$ is measurable w.r.t. $\mathcal{B}^{(m)}$. Now obviously s_0 is measurable w.r.t. $\mathcal{B}^{(m)} = \sigma(\langle s_n, n \in \mathbb{Z}\rangle)$ so that τ_0 is measurable w.r.t. this σ-field. The independence of $\langle s_n \rangle$ and $\langle \tau_n \rangle$ implies, however, that τ_0 is also independent of $\mathcal{B}^{(m)}$. Consequently τ_0 must be a constant, which contradicts the assumption of the corollary.

A contradiction arises in a completely analogous way if $\langle s_n \rangle$ is a sequence of non-constant mutually independent random variables. (This situation is considered in Loynes [47, *p.* 507].)

Hence m_0 cannot be honest/the queue is unstable if $\rho = 1$. We know from Theorem 1.3.1 that this is also the case when $\rho > 1$ but that the queue is stable when $\rho < 1$ (Theorem 1.3.2). This completes the proof.

□

Remark 1.3.7

The $GI/GI/1$ queues considered by Lindley [45] are incorporated into Corollary 1.3.5a.

Example 1.3.1

Suppose that, for some queue with $\rho = 1$, $\langle u_n \rangle$ is a Markov process and $m_n = g(u_n)$, where g is a measurable function.

First we observe that, because of the Markov property

$$\begin{aligned} & L\left[\sigma\left(u_n, u_{n+1}, \ldots\right), \sigma\left(u_{n-1}, u_{n-2}, \ldots\right) \mid \sigma\left(u_n\right)\right] \\ = \; & L\left[\sigma\left(u_n, u_{n+1}, \ldots\right) \mid \sigma\left(u_n\right)\right] \cdot L\left[\sigma\left(u_{n-1}, u_{n-2}, \ldots\right) \mid \sigma\left(u_n\right)\right] \end{aligned}$$

so $\mathcal{B}^{(n)} = \sigma\left(u_n\right)$ will do.

If the queue is to be stable, we know from Theorem 1.3.5(b) that u_n must be measurable w.r.t. $(\sigma(u_n), \sigma(u_{n+1}))$ and $m_{n+1} = m_n + u_n \Rightarrow g(u_{n+1}) = g(u_n) + u_n \;\; \forall\, n$, which is a strong condition.

Remark 1.3.8 (Stability of the Critical Queue)

From Theorem 1.3.5, Remark 1.3.7, Example 1.3.1 and especially Corollary 1.3.5a, it is clear that, unless $u_n = 0$ a.e. it is highly unlikely that the critical queue will be stable.

The set of equations (1.3.2a) is a further indication of this fact: it is possible to express m_0 in terms of both "past" values $\langle u_j, j \in \mathbb{Z}_- \rangle$ and "present/future" values $\langle u_j, j \in \mathbb{Z}_\oplus \rangle$.

CHAPTER 2

THE MULTISERVER QUEUE

2.1 INTRODUCTION

We now turn to an investigation of the $G/G/r$ (multiserver) queue.

If $\varphi_n = (\tau_n, s_n)$, then $\varphi = \langle \varphi_n; n \in \mathbb{Z} \rangle$ is stationary ergodic. In accordance with Brandt [15], we call $[\varphi, L(\varphi)]$ the CAP, the *complete arrival process*, for the queue under consideration. This notation for a CAP is slightly different from Brandt's canonical representation (Φ, P). In keeping with a comment towards the beginning of Chapter 0, φ (or ψ) may represent a random sequence or its realisation (whereas Brandt uses Φ & φ, Ψ & ψ). In addition to the assumptions of Chapter 1, we require that $\mathrm{E}(s_0)$ exists and is finite. Further on in the present chapter, more restrictions will be placed on φ.

Since we are considering a queueing system with r servers, it will become necessary to use r-dimensional (row) vectors. In particular, we write $\mathbf{1} = \mathbf{1}_r$ for the vector having $\{\mathbf{1}\}_i = 1 \;\forall\, i \in \mathbb{N}_r$, and $\boldsymbol{e}$ for the member of the standard basis of $\mathbb{R}^r$ determined by $\{\boldsymbol{e}\}_i = \delta_{1i}$.

When we write $\boldsymbol{x}^+$, we refer to a vector with ith component given by x_i^+. More generally, a vector $\boldsymbol{x} \in \mathbb{R}^r$ will be said to possess some property iff all of its components possess it.[1] For example, $\boldsymbol{x} \geq \boldsymbol{y} \Leftrightarrow x_i \geq y_i \;\forall\, i \in \mathbb{N}_r$. Then

[1] We make an exception to this general principle, namely, $\boldsymbol{x}$ will be said to be infinite iff $\exists i \in \mathbb{N}_r$ s.t. $x_i \in \{-\infty, \infty\}$.

the vector function $\boldsymbol{f} : \mathbb{R}^r \to \mathbb{R}^r$ is nondecreasing iff $\boldsymbol{x} > \boldsymbol{y} \Rightarrow \boldsymbol{f}(\boldsymbol{x}) \geq \boldsymbol{f}(\boldsymbol{y})$. With such assumptions, we are able to conclude that (in the first instance) parts (a)–(c) of Theorem 1.2.1 remain valid in r-dimensional ($r \in \overline{\mathbb{N}}_1$) Euclidean space (with vectors replacing scalars).

In Section 2.2 the appropriate Lindley equation is developed. The function which relates the waiting-times has a number of desirable and interesting properties which are then investigated in §2.3 that includes some results which, though fairly elementary, prove their worth in the more substantial theory further on in the chapter.

With the groundwork laid, it is possible to consider a theorem (due to Loynes [47]) in which conclusions are drawn for different values of ρ. A $G/G/r$ queueing system with a traffic intensity less than unity emerges as that of greatest interest for consideration in the remainder of the chapter and, for that matter, the thesis.

While less emphasis is placed on (distinguishing) (stationary) weak and strong solutions than is the case in Brandt [15], some details are provided in §2.5 where the question of uniqueness of solutions is raised. In this and subsequent sections, the existence and possible equality of extremal solutions: Loynes's minimal one ($\boldsymbol{m}_n(\varphi)$) and Brandt's maximal solution ($\boldsymbol{M}_n(\varphi)$) are investigated. The notation used here is clearly more in line with Loynes [47] than Brandt [15] & [16] who uses the symbols $\underline{w}_n(\Phi)$ and $\overline{w}_n(\Phi)$ respectively. When there is no ambiguity we shall also often suppress reference to φ (e.g., $\boldsymbol{m}_n$ is written for $\boldsymbol{m}_n(\varphi)$).

Some additional differences emerge. τ_n and q replace Brandt's A_n and d respectively. Finally, a.e. finite (waiting-time) sequences are tacitly assumed to have their zero-measure infinite realisations disregarded or "transformed". This will, for instance, render the distinction between sequences $\langle \underline{w}'_n(\Phi)\rangle$ and $\langle \underline{w}_n(\Phi)\rangle$ — a matter explicitly dealt with in Brandt [15] — unnecessary.

With technical preliminaries having been discussed, it is now possible to commence with the main investigation.

2.2 THE LINDLEY EQUATION FOR MULTISERVER QUEUES

Write

$$\boldsymbol{w}_n = (w_{n,1}, \ldots, w_{n,r}) \tag{2.2.1}$$

for the vector of "waiting times" of the nth customer. The ith component, $w_{n,i}$, represents the length of time the nth customer must wait until i channels become free of customers that arrived before him. It is immediately clear that

$$w_{n,i} \leq w_{n,i+1} \; \forall\, i \in \mathbb{N}_{r-1} \tag{2.2.2}$$

More especially, $w_{n,1} = \min_{i \in \mathbb{N}_r} w_{n,i}$ is the actual waiting time the nth customer experiences. Let

$$\mathcal{W}_r = \left\{ \boldsymbol{w} = (w_1, \ldots, w_r) \in \mathbb{R}^r_{\oplus} : w_i \leq w_{i+1} \; \forall\, i \in \mathbb{N}_{r-1} \right\} \tag{2.2.3}$$

Our first task is to find the appropriate function $\boldsymbol{f}$ of (1.2.6) which yields the $G/G/r$ Lindley-Loynes equation that generalises (1.1.1). To do so, we note the following. When the $(n+1)$st customer arrives, $w_{n,1}$ has increased by the service time, s_n, of the nth customer. In addition, a time of τ_n has passed since the arrival of the nth customer, so all "waiting times" experienced by the nth customer have decreased by τ_n, if they are not already zero. In conjunction with (2.2.2), it is clear that

$$\begin{aligned} w_{n+1,1} &= \min\left[(w_{n,1}+s_n-\tau_n)^+, (w_{n,2}-\tau_n)^+\right] \\ w_{n+1,r} &= \max\left[(w_{n,1}+s_n-\tau_n)^+, (w_{n,r}-\tau_n)^+\right] \end{aligned} \tag{2.2.4}$$

More generally, if we suppose that $w_{n,k} \le w_{n,1} + s_n \le w_{n,k+1}$, then

$$w_{n+1,i} = \begin{cases} (w_{n,i+1}-\tau_n)^+ & i = 1,\ldots,k-1 \\ (w_{n,1}+s_n-\tau_n)^+ & i = k \\ (w_{n,i}-\tau_n)^+ & i = k+1,\ldots,r \end{cases} \tag{2.2.5}$$

with obvious adjustments made for the cases $k = 1$ or $k = r$.

Denote by $\boldsymbol{R}(\boldsymbol{x})$ that vector which is obtained by (re)arranging the components of $\boldsymbol{x}$ in ascending order. In other words, if $\boldsymbol{x} = (x_1, \ldots, x_r)$, then $\boldsymbol{R}(\boldsymbol{x}) = \left(x_{\pi(1)}, \ldots, x_{\pi(r)}\right)$, where $\pi : \mathbb{N}_r \mapsto \mathbb{N}_r$ is a permutation of $\mathbb{N}_r$ such that $x_{\pi(i)} \le x_{\pi(i+1)}\ \forall\, i \in \mathbb{N}_{r-1}$. Then we may write that

$$\begin{aligned} \boldsymbol{w}_{n+1} &= \left[\boldsymbol{R}(\boldsymbol{w}_n + s_n\boldsymbol{e}) - \tau_n\mathbf{1}\right]^+ \\ &= \boldsymbol{f}(\boldsymbol{w}_n, s_n, \tau_n) \end{aligned} \tag{2.2.6}$$

where, typically, $\boldsymbol{w}_1 = \mathbf{0}$, although it can, of course, even be a random vector.

2.3 PROPERTIES OF THE POSITIVE PART $(\cdot)^+$, $\boldsymbol{R}$ AND $\boldsymbol{f}$

Lemma 2.3.1

The vector function $\boldsymbol{f}$ of (2.2.6) is nonnegative and, w.r.t. its first (vector) argument, it is nondecreasing and continuous from below.

<u>PROOF</u>

The nonnegativity is immediate. The left continuity is also not difficult to show. The monotonicity is less immediate so we reproduce the (rather tedious) details here.

Suppose that $\boldsymbol{w}^*_n > \boldsymbol{w}_n$ so $w^*_{n,i} > w_{n,i}\ \forall\, i \in \mathbb{N}_r$. Assume in addition that $w^*_{n,k^*} \leq w^*_{n,1}+s_n \leq w^*_{n,k^*+1}$. We wish to show that $\boldsymbol{w}^*_{n+1} = \boldsymbol{f}(\boldsymbol{w}^*_n) \geq \boldsymbol{f}(\boldsymbol{w}_n) = \boldsymbol{w}_{n+1}$. The proof differs according to the relative values of k ((2.2.5)) and k^*, because

$$w_{n+1,i} = \begin{cases} (w_{n,i+1} - \tau_n)^+ & i = 1,\ldots,k-1 \\ (w_{n,1} + s_n - \tau_n)^+ & i = k \\ (w_{n,i} - \tau_n)^+ & i = k+1,\ldots,r \end{cases}$$

$$w^*_{n+1,i} = \begin{cases} \left(w^*_{n,i+1} - \tau_n\right)^+ & i = 1,\ldots,k^*-1 \\ \left(w^*_{n,1} + s_n - \tau_n\right)^+ & i = k^* \\ \left(w^*_{n,i} - \tau_n\right)^+ & i = k^*+1,\ldots,r \end{cases}$$

Initially we suppose $k^* > k$.

If $i < k$, then $w^*_{n+1,i} = \left(w^*_{n,i+1} - \tau_n\right)^+ \geq (w_{n,i+1} - \tau_n)^+ = w_{n+1,i}$.

A similar result holds when $i > k^*$.

$w^*_{n+1,k} = \left(w^*_{n,k+1} - \tau_n\right)^+ \geq (w_{n,1} + s_n - \tau_n)^+ = w_{n+1,k}$ because

$w^*_{n,k+1} > w_{n,k+1} \geq w_{n,1} + s_n$;

$w^*_{n+1,k^*} = \left(w^*_{n,1} + s_n - \tau_n\right)^+ \geq (w_{n,k^*} - \tau_n)^+ = w_{n+1,k^*}$ because

$w^*_{n,1} + s_n \geq w^*_{n,k^*} > w_{n,k^*}$.

Finally, if $i = k+1, \ldots, k^* - 1$, then

$w^*_{n,i+1} \geq w^*_{n,i} > w_{n,i} \Rightarrow w^*_{n+1,i} = \left(w^*_{n,i+1} - \tau_n\right)^+ \geq (w_{n,i} - \tau_n)^+ = w_{n+1,i}$.

Next consider the case $k^* < k$.

If $i < k^*$ or $i > k$, then we proceed much as in the preceding argument. The inequalities $w^*_{n,i} \geq w^*_{n,1} + s_n > w_{n,1} + s_n \geq w_{n,j}$ $\forall\, i = k^* + 1, \ldots, k$ and $j = k^*, \ldots, k$ provide for the monotonicity of the waiting-time components in the remaining interval, $k^*, \ldots, k$.

The last possibility, $k^* = k$, follows directly from the inequality $w^*_{n,i} > w_{n,i}$ $\forall\, i \in \mathbb{N}_r$. The fact that some of the intervals in the previous two paragraphs may be empty does not present a problem. Hence the proof of the lemma is complete.

$\square$

Remark 2.3.1 (Commutativity and the Positive Part Function) In Chapter 1, indeed throughout the thesis, expressions such as $\left[\max_{j\in\mathbb{N}_m} z_j\right]^+$ are encountered.

Let $\sigma : \mathbb{N}_m \mapsto \mathbb{N}_m$ be a permutation of $\mathbb{N}_m$ such that $z_{\sigma(j)} \leq z_{\sigma(j+1)}\ \forall\, j \in \mathbb{N}_{m-1}$. In particular,

$$z_{\sigma(m)} = \max_{j\in\mathbb{N}_m} z_j \tag{2.3.1}$$

$$z_{\sigma(m)} \leq 0 \quad \Rightarrow \quad z_j \leq 0\ \forall\, j \in \mathbb{N}_m \quad \Rightarrow \quad z_j^+ = 0\ \forall\, j \in \mathbb{N}_m$$

$$z_{\sigma(m)} \geq 0 \quad \text{together with (2.3.1)} \quad \Rightarrow \quad z_{\sigma(m)}^+ = z_{\sigma(m)} \geq z_j^+\ \forall\, j \in \mathbb{N}_m$$

In either case

$$\left[\max_{j\in\mathbb{N}_m} z_j\right]^+ = z_{\sigma(m)}^+ = \max_{j\in\mathbb{N}_m}\left[z_j^+\right]$$

(since a similar proof with a permutation τ for which $z_{\tau(j)}^+ \leq z_{\tau(j+1)}^+\ \forall\, j \in \mathbb{N}_{m-1}$ yields the same result).

The intuitive conclusion is that *the maximum* (max) *and positive part* $(^+)$ *functions commute* and, analogously, *the supremum* (sup) *and positive part* $(^+)$ *functions commute.* It is equally obvious that (2.2.6) may be rewritten as $\boldsymbol{w}_{n+1} = [\boldsymbol{R}(\boldsymbol{w}_n + s_n\boldsymbol{e} - \tau_n\mathbf{1})]^+$ or $\boldsymbol{w}_{n+1} = \boldsymbol{R}\left[(\boldsymbol{w}_n + s_n\boldsymbol{e} - \tau_n\mathbf{1})^+\right]$: *the operations* $\boldsymbol{R}$ *and* $(\cdot)^+$ *commute.*

The positive part function fails to commute with a large number of other functions. A case in point is contained in Lemma 2.3.2, which follows shortly.

Remark 2.3.2 (Vector Norms)

As usual $\|\boldsymbol{x}\| = \sqrt{\boldsymbol{x}\cdot\boldsymbol{x}} = \sqrt{\sum_{i=1}^{r} x_i^2}$ denotes the *Euclidean norm* (of the Euclidean space $\mathbb{R}^r$). (Conventionally, of course, we write $|x|$ if $r = 1$.) $\|\boldsymbol{x}\|_{\max} = \max_{i\in\mathbb{N}_r} |x_i|$ shall be called the *maximum norm*. Since $|x_i| \geq 0$ and $x_i^2 \geq 0 \;\forall\, x_i \in \mathbb{R}$, it is clear that

$$\boldsymbol{x} = \mathbf{0} \;(\Leftrightarrow x_i = 0 \;\forall\, i) \Leftrightarrow \|\boldsymbol{x}\| = 0 \Leftrightarrow \|\boldsymbol{x}\|_{\max} = 0$$

Lemma 2.3.2 (Generalised Triangle Inequality)

$$x_i \in \mathbb{R} \;\forall\, i \;\Rightarrow\; \left(\sum_i x_i\right)^+ \leq \sum_i x_i^+$$

<u>Proof</u>

$$x_i^+ \;=\; \max(x_i, 0) \;\geq\; \left\{\begin{array}{lllll} x_i & \Rightarrow & \sum_i x_i^+ & \geq & \sum_i x_i \\ 0 & \Rightarrow & \sum_i x_i^+ & \geq & 0 \end{array}\right\} \Longrightarrow$$

$$\sum_i x_i^+ \;\geq\; \max\left(\sum_i x_i, 0\right) \;=\; \left(\sum_i x_i\right)^+$$

□

Lemma 2.3.3 (Inequalities)

(a) $x, y \in \mathbb{R} \Rightarrow |x^+ - y^+| \leq |x - y|$

(b) $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^r \Rightarrow \|\boldsymbol{x}^+ - \boldsymbol{y}^+\| \leq \|\boldsymbol{x} - \boldsymbol{y}\|$

PROOF

(a)

$$
\begin{aligned}
x^+ - y^+ &= [(x-y)+y]^+ - y^+ \\
&\leq [(x-y)^+ + y^+] - y^+ \quad (=) \text{ by Lemma 2.3.2} \\
&\leq (x-y)^+ \\
&\leq |x-y|
\end{aligned}
$$

and similarly

$$
\begin{aligned}
y^+ - x^+ &\leq |y-x| \\
\text{Therefore } |x^+ - y^+| &\leq |x-y|
\end{aligned}
$$

(b)

$$
\begin{aligned}
\|\boldsymbol{x}^+ - \boldsymbol{y}^+\| &= \sqrt{\sum_{i=1}^{r} \left|x_i^+ - y_i^+\right|^2} \\
&\leq \sqrt{\sum_{i=1}^{r} |x_i - y_i|^2} \quad (=) \text{ by (a)} \\
&\leq \|\boldsymbol{x} - \boldsymbol{y}\|
\end{aligned}
$$

□

Corollary 2.3.3*a*

If $\boldsymbol{x} \geq \mathbf{0}$ and $\exists\, j \in \mathbb{N}_r$ for which $y_j < 0$, then

$$\|\boldsymbol{x}^+ - \boldsymbol{y}^+\| < \|\boldsymbol{x} - \boldsymbol{y}\|$$

<u>PROOF</u>

$$\begin{array}{lllll} \boldsymbol{x} & \geq & \mathbf{0} & \Rightarrow & x_j^+ = x_j \geq 0 \qquad\qquad \text{and} \\ y_j & < & 0 & \Rightarrow & y_j^+ = 0 > y_j \end{array}$$

$$\Longrightarrow \quad \left|x_j^+ - y_j^+\right| \quad = \quad x_j - 0 \quad < \quad x_j - y_j \quad (=)$$

$$< \quad |x_j - y_j|$$

$$\text{but} \qquad \left|x_i^+ - y_i^+\right| \quad \leq \quad |x_i - y_i| \quad \forall\, i \in \mathbb{N}_r \setminus \{j\}$$

$$\Longrightarrow \qquad \|\boldsymbol{x}^+ - \boldsymbol{y}^+\| \quad < \quad \|\boldsymbol{x} - \boldsymbol{y}\|$$

$\square$

Lemma 2.3.4 (Contractive Property w.r.t. the Maximum Norm) Let $\boldsymbol{w}^{(1)}$, $\boldsymbol{w}^{(2)} \in \mathcal{W}_r$ and $s, \tau \in \mathbb{R}_{\oplus}$. Then

$$\left\| \boldsymbol{f}\left(\boldsymbol{w}^{(1)}, s, \tau\right) - \boldsymbol{f}\left(\boldsymbol{w}^{(2)}, s, \tau\right) \right\|_{\max} \leq \left\| \boldsymbol{w}^{(1)} - \boldsymbol{w}^{(2)} \right\|_{\max} \tag{2.3.2}$$

<u>PROOF</u>

Let

$$v_i^{(l)} = \left\{ \boldsymbol{v}^{(l)} \right\}_i = \left\{ \boldsymbol{f}\left(\boldsymbol{w}^{(l)}, s, \tau\right) \right\}_i = \left\{ \begin{array}{ll} \left(w_{i+1}^{(l)} - \tau\right)^+ & i = 1, \ldots, k_l - 1 \\ \left(w_1^{(l)} + s - \tau\right)^+ & i = k_l \\ \left(w_i^{(l)} - \tau\right)^+ & i = k_l + 1, \ldots, r \end{array} \right\} l = 1, 2$$

In the sequel, extensive use shall be made of Lemma 2.3.3 in conjunction with the above.

If $k_1 = k = k_2$, then

$$\begin{aligned} \left| v_i^{(1)} - v_i^{(2)} \right| &\leq \left\{ \begin{array}{ll} \left| w_{i+1}^{(1)} - w_{i+1}^{(2)} \right| & i = 1, \ldots, k-1 \\ \left| w_1^{(1)} - w_1^{(2)} \right| & i = k \\ \left| w_i^{(1)} - w_i^{(2)} \right| & i = k+1, \ldots, r \end{array} \right. \\ \Longrightarrow \quad \left| v_i^{(1)} - v_i^{(2)} \right| &\leq \max_{j \in \mathbb{N}_r} \left| w_j^{(1)} - w_j^{(2)} \right| \quad \forall\, i \in \mathbb{N}_r \end{aligned}$$

Otherwise we may assume w.l.g. that $k_1 < k_2$. The proof is tedious, but not difficult.

$$
\begin{array}{lll}
\textit{Case}\quad 1: & i = 1,\ldots,k_1-1 & \\
& \left|v_i^{(1)} - v_i^{(2)}\right| \;\le\; \left|w_{i+1}^{(1)} - w_{i+1}^{(2)}\right| &
\end{array}
$$

$$
\begin{array}{lll}
\textit{Case}\quad 2: & i = k_1 & \\
& \left|v_i^{(1)} - v_i^{(2)}\right| \;\le\; \left|w_1^{(1)} + s - w_{k_1+1}^{(2)}\right| & (=) \\
& \qquad\qquad\quad \le \begin{cases} \left|w_{k_1+1}^{(1)} - w_{k_1+1}^{(2)}\right| & w_1^{(1)} + s \ge w_{k_1+1}^{(2)} \\ \left|w_1^{(1)} - w_1^{(2)}\right| & w_1^{(1)} + s \le w_{k_1+1}^{(2)} \end{cases} &
\end{array}
$$

$$
\begin{array}{ll}
\textit{Case}\quad 3: & i = k_1+1,\ldots,k_2-1 \\
& \left|v_i^{(1)} - v_i^{(2)}\right| \;\le\; \left|w_i^{(1)} - w_{i+1}^{(2)}\right| \\
& \qquad\qquad\quad \le \begin{cases} \left|w_{i+1}^{(1)} - w_{i+1}^{(2)}\right| & w_i^{(1)} \ge w_{i+1}^{(2)} \\ \left|w_1^{(1)} - w_1^{(2)}\right| & w_i^{(1)} \le w_{i+1}^{(2)} \end{cases}
\end{array}
$$

The cases $i = k_2$ and $i = k_2+1,\ldots,r$ are proved similarly to Cases 2 and 1 respectively.

Again, therefore,

$$\left|v_i^{(1)} - v_i^{(2)}\right| \le \max_{j\in\mathbb{N}_r} \left|w_j^{(1)} - w_j^{(2)}\right| \qquad \forall\, i \in \mathbb{N}_r$$

which finally means that

$$\left\|\boldsymbol{v}^{(1)} - \boldsymbol{v}^{(2)}\right\|_{\max} \le \left\|\boldsymbol{w}^{(1)} - \boldsymbol{w}^{(2)}\right\|_{\max}$$

$\square$

Lemma 2.3.5

Let

$\boldsymbol{x} = (x_1, \ldots, x_r)$, $\boldsymbol{y} = (y_1, \ldots, y_r) \in \mathbb{R}^r$;

π and σ be (possibly equivalent) permutations of $\mathbb{N}_r$ so that

$\boldsymbol{R}(\boldsymbol{x}) = \left(x_{\pi(1)}, \ldots, x_{\pi(r)}\right)$ and $\boldsymbol{R}(\boldsymbol{y}) = \left(y_{\sigma(1)}, \ldots, y_{\sigma(r)}\right)$.

Then

$$\boldsymbol{x} \cdot \boldsymbol{y} \leq \boldsymbol{R}(\boldsymbol{x}) \cdot \boldsymbol{R}(\boldsymbol{y})$$

with equality iff $x_i = x_{\pi(j)} \Leftrightarrow y_i = y_{\sigma(j)} \ \forall\, i \in \mathbb{N}_r$
i.e. iff $\boldsymbol{x}$ and $\boldsymbol{y}$ are "ordered" in the same sense.

PROOF

We first show by indirect proof that the maximum scalar product occurs when $\boldsymbol{x}$ and $\boldsymbol{y}$ are ordered in the same sense. Suppose $\boldsymbol{x}$ and $\boldsymbol{y}$ are not similarly ordered. By the commutativity of addition and of multiplication, we may assume w.l.g. that $\boldsymbol{x} = \boldsymbol{R}(\boldsymbol{x})$ but that the components of $\boldsymbol{y}$ are not in ascending order. As a result $\exists\, m, n$ for which $x_m \leq x_n$ but $y_m > y_n$. Write $\tilde{\boldsymbol{y}}$ for the vector obtained by interchanging y_m and y_n. It follows that

$$\begin{array}{rl} & \boldsymbol{x} \cdot (\tilde{\boldsymbol{y}} - \boldsymbol{y}) \\ = & (x_1, \ldots, x_m, \ldots, x_n, \ldots, x_r) \cdot (0, \ldots, y_n - y_m, \ldots, y_m - y_n, \ldots, 0) \\ = & (x_n - x_m)(y_m - y_n) \\ \geq & 0 \\ \Longrightarrow & \boldsymbol{x} \cdot \tilde{\boldsymbol{y}} \text{ is at least as great as } \boldsymbol{x} \cdot \boldsymbol{y} \end{array}$$

Continuing in this way, we find that, after a finite number of such exchanges, $\boldsymbol{R}(\boldsymbol{x}) \cdot \boldsymbol{R}(\boldsymbol{y})$ yields the maximum scalar product. More details are provided in Hardy *et al.* [33, Chapter 10].

$\square$

Lemma 2.3.6 (Contractive Property w.r.t. the Euclidean Norm)
Let $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^r$ and $s, \tau \in \mathbb{R}$. Then

(a) (i) $\|\boldsymbol{R}(\boldsymbol{x}) - \boldsymbol{R}(\boldsymbol{y})\| \le \|\boldsymbol{x} - \boldsymbol{y}\|$

(ii) $\|\boldsymbol{f}(\boldsymbol{x}, s, \tau) - \boldsymbol{f}(\boldsymbol{y}, s, \tau)\| \le \|\boldsymbol{x} - \boldsymbol{y}\|$

(b) In particular, suppose that $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{W}_r$ and $s \in \mathbb{R}_\oplus$. If, in addition,

$$\|\boldsymbol{R}(\overline{\boldsymbol{x}}) - \boldsymbol{R}(\overline{\boldsymbol{y}})\| = \|\boldsymbol{x} - \boldsymbol{y}\|$$

where $\overline{\boldsymbol{x}} = \boldsymbol{x} + s\boldsymbol{e}$ and $\overline{\boldsymbol{y}} = \boldsymbol{y} + s\boldsymbol{e}$,

then $\exists$ permutation κ of $\mathbb{N}_r$ for which

$$\begin{cases} \boldsymbol{R}(\overline{\boldsymbol{x}}) &= \left(\overline{x}_{\kappa(1)}, \dots, \overline{x}_{\kappa(r)}\right) \\ \boldsymbol{R}(\overline{\boldsymbol{y}}) &= \left(\overline{y}_{\kappa(1)}, \dots, \overline{y}_{\kappa(r)}\right) \end{cases}$$

such that $y_1 + s < y_r \Rightarrow \kappa(r) = r$.

PROOF

(a) Let π and σ be permutations of $\mathbb{N}_r$ for which

$$\begin{cases} \boldsymbol{R}(\boldsymbol{x}) &= \left(x_{\pi(1)}, \dots, x_{\pi(r)}\right) \\ \boldsymbol{R}(\boldsymbol{y}) &= \left(y_{\sigma(1)}, \dots, y_{\sigma(r)}\right) \end{cases}$$

(i) From Lemma 2.3.5 $\sum_{j=1}^{r} x_j y_j \le \sum_{j=1}^{r} x_{\pi(j)} y_{\sigma(j)}$ which implies that

$$\begin{aligned}
\|\boldsymbol{R}(\boldsymbol{x}) - \boldsymbol{R}(\boldsymbol{y})\|^2 &= \sum_{j=1}^{r} \left[x_{\pi(j)}^2 + y_{\sigma(j)}^2\right] - 2\sum_{j=1}^{r} x_{\pi(j)} y_{\sigma(j)} \\
&\le \sum_{j=1}^{r} \left[x_j^2 + y_j^2\right] - 2\sum_{j=1}^{r} x_j y_j & (=) \\
&\le \|\boldsymbol{x} - \boldsymbol{y}\|^2
\end{aligned}$$

(ii)

$$\begin{aligned}
& \|\boldsymbol{f}(\boldsymbol{x}, s, \tau) - \boldsymbol{f}(\boldsymbol{y}, s, \tau)\| \\
= \ & \left\|[\boldsymbol{R}(\boldsymbol{x} + s\boldsymbol{e} - \tau\mathbf{1})]^+ - [\boldsymbol{R}(\boldsymbol{y} + s\boldsymbol{e} - \tau\mathbf{1})]^+\right\| \\
\le \ & \|\boldsymbol{R}(\boldsymbol{x} + s\boldsymbol{e} - \tau\mathbf{1}) - \boldsymbol{R}(\boldsymbol{y} + s\boldsymbol{e} - \tau\mathbf{1})\| && \text{by Lemma 2.3.3} \\
\le \ & \|(\boldsymbol{x} + s\boldsymbol{e} - \tau\mathbf{1}) - (\boldsymbol{y} + s\boldsymbol{e} - \tau\mathbf{1})\| && (=) \ \text{ from (i)} \\
\le \ & \|\boldsymbol{x} - \boldsymbol{y}\|
\end{aligned}$$

(b) If $s = 0$, then $\overline{\boldsymbol{x}} = \boldsymbol{x}$ and $\overline{\boldsymbol{y}} = \boldsymbol{y}$, but $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{W}_r \Rightarrow \overline{\boldsymbol{x}} = \boldsymbol{x} = \boldsymbol{R}(\boldsymbol{x}) = \boldsymbol{R}(\overline{\boldsymbol{x}})$; similarly $\overline{\boldsymbol{y}} = \boldsymbol{R}(\overline{\boldsymbol{y}})$. Consequently $\kappa = \iota$, the identity permutation of $\mathbb{N}_r$.

Otherwise $s > 0 \Rightarrow x_1 < x_1 + s$ and $y_1 < y_1 + s$. It is, therefore, meaningful to let $k = \max\{i : x_i < x_1 + s\}$ and $l = \max\{i : y_i < y_1 + s\}$.

Then

$$\{\boldsymbol{R}(\overline{\boldsymbol{x}})\}_i = \begin{cases} x_{i+1} & i = 1, \ldots, k-1 \\ x_1 + s & i = k \\ x_i & i = k+1, \ldots, r \end{cases}$$

$$= \overline{x}_{\overline{\sigma}(i)}$$

$$\{\boldsymbol{R}(\overline{\boldsymbol{y}})\}_i = \begin{cases} y_{i+1} & i = 1, \ldots, l-1 \\ y_1 + s & i = l \\ y_i & i = l+1, \ldots, r \end{cases}$$

$$= \overline{y}_{\overline{\pi}(i)}$$

For $k = l$, obviously

$$\kappa(i) = \begin{cases} i+1 & i = 1, \ldots, k-1 \\ 1 & i = k \\ i & i = k+1, \ldots, r \end{cases}$$

If, on the other hand, $k \neq l$, then we may assume w.l.g. that $k < l$.

Then

$$
\begin{aligned}
& \sum_{j=2}^{k}(x_j - y_j)^2 + \sum_{j=l+1}^{r}(x_j - y_j)^2 + (x_1 - y_1)^2 + \sum_{j=k+1}^{l}(x_j - y_j)^2 \\
= \quad & \|\boldsymbol{x} - \boldsymbol{y}\|^2 \\
= \quad & \|\boldsymbol{R}(\overline{\boldsymbol{x}}) - \boldsymbol{R}(\overline{\boldsymbol{y}})\|^2 \qquad \text{by assumption} \\
= \quad & \sum_{j=1}^{k-1}(x_{j+1} - y_{j+1})^2 + (x_1 + s - y_{k+1})^2 + \sum_{j=k+1}^{l-1}(x_j - y_{j+1})^2 \\
& +(x_l - y_1 - s)^2 + \sum_{j=l+1}^{r}(x_j - y_j)^2 \\
= \quad & \sum_{j=2}^{k}(x_j - y_j)^2 + \sum_{j=l+1}^{r}(x_j - y_j)^2 + \sum_{j=k+1}^{l}(x_j^2 + y_j^2) - 2\sum_{j=k+1}^{l-1} x_j y_{j+1} \\
& +(x_1 + s)^2 + (y_1 + s)^2 - 2(x_1 + s)y_{k+1} - 2x_l(y_1 + s)
\end{aligned}
$$

$$
\begin{aligned}
\Longrightarrow \quad & [(x_1 + s) - (y_1 + s)]^2 - 2\sum_{j=k+1}^{l} x_j y_j = \\
& (x_1 + s)^2 + (y_1 + s)^2 - 2\left[(x_1 + s)y_{k+1} + \sum_{j=k+2}^{l} x_{j-1}y_j + x_l(y_1 + s)\right] \\
\Rightarrow \quad & x_{k+1}y_{k+1} + \sum_{j=k+2}^{l} x_j y_j + (x_1 + s)(y_1 + s) = \\
& (x_1 + s)y_{k+1} + \sum_{j=k+2}^{l} x_{j-1}y_j + x_l(y_1 + s)
\end{aligned}
\tag{2.3.3}
$$

The right-hand side is equal to $\sum_{j=k}^{l} \overline{x}_{\overline{\sigma}(j)}\overline{y}_{\overline{\pi}(j)}$. From Lemma 2.3.5 we know this to be the maximum possible product of the components of the vectors $(x_1 + s, x_{k+1}, \ldots, x_{l-1}, x_l)$ and $(y_{k+1}, y_{k+2}, \ldots, y_l, y_1 + s)$. We also deduce from the same lemma that, because of the equality (2.3.3), $(x_{k+1}, x_{k+2}, \ldots, x_l, x_1 + s)$ and $(y_{k+1}, y_{k+2}, \ldots, y_l, y_1 + s)$ are similarly ordered which implies that

$$x_{k+1} \leq x_{k+2} \leq \ldots \leq x_l \leq x_1 + s \text{ but } x_1 + s \leq x_{k+1}$$

$$\begin{aligned}
\Rightarrow \quad & x_{k+1} = x_{k+2} = \ldots = x_l = x_1 + s \\
\Rightarrow \quad & \{\boldsymbol{R}(\overline{\boldsymbol{x}})\}_i = \begin{cases} x_{i+1} & i = 1, \ldots, k-1 \\ x_{i+1} & i = k, \ldots, l-1 \\ x_1 + s & i = l \\ x_i & i = l+1, \ldots, r \end{cases} \\
& \qquad = \overline{x}_{\overline{\pi}(i)} \\
\Rightarrow \quad & \kappa(i) = \overline{\pi}(i) \\
& \qquad = \begin{cases} i+1 & i = 1, \ldots, l-1 \\ 1 & i = l \\ i & i = l+1, \ldots, r \end{cases}
\end{aligned}$$

Finally, suppose $y_1 + s < y_r$. Then, since $\boldsymbol{y} \in \mathcal{W}_r$,

$$y_r = \max_{j \in \mathbb{N}_r} \overline{y}_j \Rightarrow \kappa(r) = r$$

$\square$

Remark 2.3.3

If $y_1 + s = y_r$, then $\kappa(r) = r$ or $\kappa(r) = 1$.

The following lemma, the last of the section, is concerned with a function similar to $f_*^{[\cdot]}$ in the proof of Lemma 1.2.2(b).

Lemma 2.3.7

Let

$$\begin{aligned} \boldsymbol{f}_1(\boldsymbol{w}, s_1, \tau_1) &= \boldsymbol{f}(\boldsymbol{w}, s_1, \tau_1) \\ \boldsymbol{f}_{k+1}\left(\boldsymbol{w}, (s_i, \tau_i;\ i \in \mathbb{N}_{k+1})\right) &= \boldsymbol{f}\left[\boldsymbol{f}_k\left(\boldsymbol{w}, (s_i, \tau_i;\ i \in \mathbb{N}_k)\right), s_{k+1}, \tau_{k+1}\right] \end{aligned} \tag{2.3.4}$$

so that $\boldsymbol{f}_k$ ($k \in \mathbb{N}$) is the k-fold iteration of $\boldsymbol{f}$. Then for any $k \in \mathbb{N}$, $\boldsymbol{f}_k$ inherits the following properties of $\boldsymbol{f}$:

(a) monotonicity in its first argument

(b) contractivity (in its first argument) w.r.t. the maximum norm

PROOF

(a) For $k = 1$, Lemma 2.3.1 may be applied. Suppose, therefore, that $\boldsymbol{f}_k\left(\boldsymbol{w}, (s_i, \tau_i;\ i \in \mathbb{N}_k)\right)$ is nondecreasing in $\boldsymbol{w}$ for some $k \in \mathbb{N}$. Then

$$\begin{aligned} & \boldsymbol{w} > \boldsymbol{w}^* \\ \Rightarrow\ & \boldsymbol{f}_k\left(\boldsymbol{w}, (s_i, \tau_i;\ i \in \mathbb{N}_k)\right) \\ & \geq\ \boldsymbol{f}_k\left(\boldsymbol{w}^*, (s_i, \tau_i;\ i \in \mathbb{N}_k)\right) \\ \Rightarrow\ & \boldsymbol{f}\left[\boldsymbol{f}_k\left(\boldsymbol{w}, (s_i, \tau_i;\ i \in \mathbb{N}_k)\right), s_{k+1}, \tau_{k+1}\right] \\ & \geq\ \boldsymbol{f}\left[\boldsymbol{f}_k\left(\boldsymbol{w}^*, (s_i, \tau_i;\ i \in \mathbb{N}_k)\right), s_{k+1}, \tau_{k+1}\right] \quad \text{monotonicity of } \boldsymbol{f} \\ \Rightarrow\ & \boldsymbol{f}_{k+1}\left(\boldsymbol{w}, (s_i, \tau_i;\ i \in \mathbb{N}_{k+1})\right) \\ & \geq\ \boldsymbol{f}_{k+1}\left(\boldsymbol{w}^*, (s_i, \tau_i;\ i \in \mathbb{N}_{k+1})\right) \qquad \text{by (2.3.4)} \end{aligned}$$

and induction completes the proof.

(b) Lemma 2.3.4 produces the desired result for $\boldsymbol{f}_1$. Assume, therefore, that $\boldsymbol{f}_k(\boldsymbol{w},(s_i,\tau_i;\ i\in\mathbb{N}_k))$ has the contractive property. Then

$$
\begin{aligned}
&\left\|\boldsymbol{f}_{k+1}\left(\boldsymbol{w}^{(1)},(s_i,\tau_i;\ i\in\mathbb{N}_{k+1})\right)-\boldsymbol{f}_{k+1}\left(\boldsymbol{w}^{(2)},(s_i,\tau_i;\ i\in\mathbb{N}_{k+1})\right)\right\|_{\max}\\
=\ &\left\|\begin{array}{l}\boldsymbol{f}\left[\boldsymbol{f}_k\left(\boldsymbol{w}^{(1)},(s_i,\tau_i;\ i\in\mathbb{N}_k)\right),s_{k+1},\tau_{k+1}\right]\\ -\boldsymbol{f}\left[\boldsymbol{f}_k\left(\boldsymbol{w}^{(2)},(s_i,\tau_i;\ i\in\mathbb{N}_k)\right),s_{k+1},\tau_{k+1}\right]\end{array}\right\|_{\max}\\
\leq\ &\left\|\boldsymbol{f}_k\left(\boldsymbol{w}^{(1)},(s_i,\tau_i;\ i\in\mathbb{N}_k)\right)-\boldsymbol{f}_k\left(\boldsymbol{w}^{(2)},(s_i,\tau_i;\ i\in\mathbb{N}_k)\right)\right\|_{\max}\\
\leq\ &\left\|\boldsymbol{w}^{(1)}-\boldsymbol{w}^{(2)}\right\|_{\max}
\end{aligned}
$$

□

Remark 2.3.4 (k-Fold Iteration $\boldsymbol{f}_k$)

(a) The properties of $\boldsymbol{f}$ inherited by $\boldsymbol{f}_k$ are not limited to those given in the lemma. For the purposes of this chapter, those given in Lemma 2.3.7 will, nonetheless, suffice.

(b) Using the notation of Loynes [47], one may write

$$
\begin{aligned}
\boldsymbol{w}_{n+1}(\boldsymbol{w}) &= \boldsymbol{f}_n(\boldsymbol{w},s_1,\ldots,s_n,\tau_1,\ldots,\tau_n)\\
&= \boldsymbol{f}_n(\boldsymbol{w},(s_i,\tau_i;\ i\in\mathbb{N}_n))
\end{aligned}
\tag{2.3.5}
$$

2.4 LOYNES'S MINIMAL SOLUTION, AND STABILITY

We again consider the $G/G/r$ queue according to the hypotheses of §2.1 (which include the fact that $\mathrm{E}(s_0)<\infty$) and the notation of the previous sections. It is convenient to write $u_n = s_n - r\tau_n$, which notation is consistent with the case $r = 1$ of the first chapter. The definition which follows establishes a connection with the stability introduced in that chapter.

Definition 2.4.1
Write $F_n = L(w_{n,1}, \ldots, w_{n,r})$. Then the $G/G/r$ queue shall be said to be

(a) *convergent stable* iff $F_n \xrightarrow[n\to\infty]{d} F$ where F is a proper distribution function

(b) *stable* iff the sequence $\langle w_{n,i}; n \in \mathbb{N}\rangle$ is stable $\forall\, i \in \mathbb{N}_r$

(c) *unstable* iff it is not stable

Lemma 2.4.1 (Equivalent Conditions for Stability)
The $G/G/r$ queue is stable iff either one of the following conditions holds:

(a) $\langle w_{n,r}; n \in \mathbb{N}\rangle$ is stable

(b) $\left\langle \sum\limits_{j=1}^{r} w_{n,j}; n \in \mathbb{N} \right\rangle$ is stable

PROOF

(a) *Necessity.* From Definition 2.4.1 stability of the $G/G/r$ queue $\Rightarrow \langle w_{n,i}; n \in \mathbb{N}\rangle$ stable $\forall\, i \in \mathbb{N}_r \Rightarrow \langle w_{n,r}; n \in \mathbb{N}\rangle$ stable.

Sufficiency. Suppose $\langle w_{n,r}; n \in \mathbb{N}\rangle$ is stable. From Lemma 1.2.1 and Remark 1.2.1, we know that there exists a proper distribution function

G such that $G(x) \le P[w_{n,r} \le x] \ \forall\, x \in \mathbb{R} \ \forall\, n \in \mathbb{N}$ but

$$\begin{aligned} w_{n,r} = \max_{i \in \mathbb{N}_r} w_{n,i} \ &\Rightarrow \ G(x) \le P[w_{n,r} \le x] \le P[w_{n,i} \le x] \ \forall\, i \in \mathbb{N}_r \\ &\Rightarrow \ \langle w_{n,i}; n \in \mathbb{N} \rangle \text{ stable } \forall\, i \in \mathbb{N}_r \\ &\Rightarrow \ \text{the queue is stable, as required} \end{aligned}$$

(b) *Necessity.* If the $G/G/r$ queue is stable, then each sequence $\langle w_{n,i}; n \in \mathbb{N} \rangle$ is stable. An extension, using induction, of Corollary 1.2.1(b) yields the fact that $\left\langle \sum_{j=1}^{r} w_{n,j}; n \in \mathbb{N} \right\rangle$ is also stable.

Sufficiency. As in (a), we make use of Lemma 1.2.1 and Remark 1.2.1. The proper distribution function H which bounds $P\left[\sum_{j=1}^{r} w_{n,j} \le x\right]$ from below is also a lower bound of $P[w_{n,i} \le x]$ for any $i \in \mathbb{N}_r$ since $\boldsymbol{w}_n \ge \boldsymbol{0} \Rightarrow H(x) \le P\left[\sum_{j=1}^{r} w_{n,j} \le x\right] \le P[w_{n,i} \le x] \ \ \forall\, i \in \mathbb{N}_r$.

□

Lemma 2.4.2 (The Essential Lemma)

Let

$$b_n = r w_{n,r} - \sum_{j=1}^{r} w_{n,j} = (r-1) w_{n,r} - \sum_{j=1}^{r-1} w_{n,j} = \sum_{j=1}^{r-1} (w_{n,r} - w_{n,j}) \quad \forall\, n \tag{2.4.1}$$

Then

(a) $w_{n,i} - w_{n,1} \le b_n \quad \forall\, i \in \mathbb{N}_r, \ \ \forall\, n$ (2.4.2)

(b) there exists a stationary a.e. finite sequence $\langle x_n \rangle$ such that

$$\frac{x_n}{n} \xrightarrow[\infty]{n} 0 \quad \text{and} \quad b_n \le (r-1) s_{n-1} + x_n + r w_{1,r} \tag{2.4.3}$$

(c) Thus

(i) $\lim_{y\to\infty} \liminf_{n\to\infty} P[b_n \leq y] = 1$

(ii) $\frac{1}{n}\left(rw_{n,i} - \sum_{j=1}^{r} w_{n,j}\right) \overset{n}{\underset{\infty}{\to}} 0 \quad \forall\, i \in \mathbb{N}_r, \quad \forall\, n$

Hence $w_{n,1}$ and $\boldsymbol{w}_n$ are stable or unstable together.

<u>PROOF</u>

Initially we note that, if $s_n = 0$ a.e., then $\boldsymbol{w}_1 = \mathbf{0} \Rightarrow \boldsymbol{w}_n = \mathbf{0}$ and even if the initial condition is (finite but) not $\mathbf{0}$, the system is stable. Hence there is no need to consider this situation further.

(a) The result is easily shown as follows:

$$\begin{aligned} w_{n,i} - w_{n,1} &\leq w_{n,r} - w_{n,1} \quad \forall\, i \in \mathbb{N}_r \\ &\leq (w_{n,r} - w_{n,1}) + \sum_{j=2}^{r-1}(w_{n,r} - w_{n,j}) \quad (=) \\ &\leq b_n \end{aligned}$$

(b)

Case 1: $w_{n,1} + s_n < w_{n,r} \Rightarrow b_{n+1} \le b_n - s_n$

From (2.2.4), $w_{n+1,r} = (w_{n,r} - \tau_n)^+$. More generally, write

$$w_{n+1,i} = \begin{cases} (w_{n,i+1} - \tau_n)^+ & i = 1, \ldots, k-1 \\ (w_{n,1} + s_n - \tau_n)^+ & i = k \\ (w_{n,i} - \tau_n)^+ & i = k+1, \ldots, r \end{cases}$$

as in (2.2.5) (with $k < r$).

Using Lemma 2.3.3(a) and noting that

$$\begin{aligned} \sum_{j=1}^{r-1} w_{n+1,j} &= \sum_{j=2}^{k} (w_{n,j} - \tau_n)^+ + (w_{n,1} + s_n - \tau_n)^+ \\ &\quad + \sum_{j=k+1}^{r-1} (w_{n,j} - \tau_n)^+ \text{, we obtain} \end{aligned}$$

$$\begin{aligned} b_{n+1} &= \sum_{j=2}^{r-1} \left[w_{n+1,r} - (w_{n,j} - \tau_n)^+ \right] + w_{n+1,r} \\ &\quad - (w_{n,1} + s_n - \tau_n)^+ \\ &= \sum_{j=2}^{r-1} \left| (w_{n,r} - \tau_n)^+ - (w_{n,j} - \tau_n)^+ \right| \\ &\quad + \left| (w_{n,r} - \tau_n)^+ - (w_{n,1} + s_n - \tau_n)^+ \right| \\ &\le \sum_{j=2}^{r-1} (w_{n,r} - w_{n,j}) + (w_{n,r} - w_{n,1}) - s_n \quad (=) \\ &\le (r-1) w_{n,r} - \sum_{j=1}^{r-1} w_{n,j} - s_n \quad (=) \\ &\le b_n - s_n \end{aligned}$$

Case 2: $w_{n,1} + s_n \geq w_{n,r} \Rightarrow b_{n+1} \leq (r-1)s_n$

Now (2.2.4) gives us $w_{n+1,r} = (w_{n,1} + s_n - \tau_n)^+$ and

$$w_{n+1,i} = \begin{cases} (w_{n,i+1} - \tau_n)^+ & i = 1, \ldots, r-1 \\ (w_{n,1} + s_n - \tau_n)^+ & i = r \end{cases}$$ which leads to

$$\begin{aligned}
b_{n+1} &= (r-1)w_{n+1,r} - \sum_{j=1}^{r-1} w_{n+1,j} \\
&= (r-1)\,(w_{n,1} + s_n - \tau_n)^+ - \sum_{j=2}^{r} (w_{n,j} - \tau_n)^+ \\
&= \sum_{j=2}^{r} \left[(w_{n,1} + s_n - \tau_n)^+ - (w_{n,j} - \tau_n)^+ \right] \\
&\leq \sum_{j=2}^{r} (w_{n,1} + s_n - w_{n,j}) && (=) \\
&\leq (r-1)w_{n,1} + (r-1)s_n - \sum_{j=2}^{r} w_{n,j} && (=) \\
&\leq (r-1)w_{n,1} + (r-1)s_n + (b_n + w_{n,1} - r w_{n,r}) && (=) \\
&\leq b_n + (r-1)s_n + r(w_{n,1} - w_{n,r}) \\
&\leq \left(r w_{n,r} - \sum_{j=1}^{r} w_{n,1} \right) + r(w_{n,1} - w_{n,r}) + (r-1)s_n && (=) \\
&\leq (r-1)s_n
\end{aligned}$$

where, in the penultimate step, we use the fact $w_{n,1} = \min_{i \in \mathbb{N}_r} w_{n,i}$. The first inequality is based on the results of Lemma 2.3.3(a).

Hence

$$b_{n+1} \leq \max[b_n - s_n, (r-1)s_n] \tag{2.4.4}$$

Next let

$$c_n = b_n - (r-1)s_{n-1} \quad \text{and} \quad v_n = (r-1)s_{n-1} - rs_n$$

Since $\langle s_n \rangle$ is a stationary sequence, so too is $\langle v_n \rangle$, and $\mathrm{E}(v_0) = -\mathrm{E}(s_0) \leq 0$. If $\mathrm{E}(s_0) = 0$, then $s_n = 0$ a.e. (because s_n is nonnegative) which, as explained already, is not of interest. Also

$$c_n + v_n = b_n - rs_n$$

and

$$\begin{array}{rcll} c_{n+1} & = & b_{n+1} - (r-1)s_n & \\ & \leq & \max[b_n - s_n, (r-1)s_n] - (r-1)s_n & (=) \\ & \leq & \max[b_n - rs_n, 0] & (=) \\ & \leq & (c_n + v_n)^+ & \end{array}$$

As indicated above, we need only consider the case in which $\mathrm{E}(v_0) < 0$. (2.4.4) does not allow us to determine the sign of c_1, but obviously $c_1^+ \geq 0$. We may, therefore, use c_1^+ as an initial condition $w_1 = c_1^+$. Thereafter we may consider a queue, the waiting-time sequence of which is given by $w_{n+1} = (w_n + v_n)^+ \geq 0$. With $\mathrm{E}(v_0) < 0$ we are thus in a position to apply the results of Theorem 1.3.2 and to conclude that there exists a unique a.e. finite stationary solution sequence $\langle \widetilde{m}_n \rangle$ for which $\frac{\widetilde{m}_n}{n} \xrightarrow[\infty]{n} 0$. (2.4.5)

We claim that the following relationship exists between $\langle \widetilde{m}_n \rangle$ and $\langle c_n \rangle$:

$$c_n \leq \widetilde{m}_n + c_1^+ \tag{2.4.6}$$

Obviously $c_1 \leq c_1^+ \leq \widetilde{m}_1 + c_1^+$, so suppose $c_n \leq \widetilde{m}_n + c_1^+$ holds for some n. Then

$$\begin{aligned} c_{n+1} &\leq (c_n + v_n)^+ \\ &\leq \left((\widetilde{m}_n + c_1^+) + v_n\right)^+ \quad (=) \quad \text{by hypothesis} \\ &\leq \left((\widetilde{m}_n + v_n) + c_1^+\right)^+ \\ &\leq (\widetilde{m}_n + v_n)^+ + c_1^+ \qquad (=) \quad \text{by Lemma 2.3.2} \\ &\leq \widetilde{m}_{n+1} + c_1^+ \qquad\qquad\quad \text{as required} \end{aligned}$$

From (2.4.6) and the definition of c_1, we obtain

$$\begin{aligned} b_n - (r-1)s_{n-1} &\leq \widetilde{m}_n + [b_1 - (r-1)s_0]^+ \qquad (=) \\ &\leq \widetilde{m}_n + \left(rw_{1,r} - \left[\sum_{j=1}^{r} w_{1,j} + (r-1)s_0\right]\right)^+ \\ &\leq \widetilde{m}_n + rw_{1,r} \\ \text{or} \quad rw_{n,r} - \sum_{j=1}^{r} w_{n,j} &\leq (r-1)s_{n-1} + \widetilde{m}_n + rw_{1,r} \end{aligned}$$

In view of this and (2.4.5), $x_n = \widetilde{m}_n$ will do.

(c) From (b)

(i)

$$\begin{aligned} \lim_{y\to\infty} \liminf_{n\to\infty} P[b_n \leq y] &\geq \lim_{y\to\infty} \liminf_{n\to\infty} P[(r-1)s_{n-1} + x_n + rw_{1,r} \leq y] \\ &= 1 \end{aligned}$$

since the random variables on the right-hand side of the inequality are stationary and, more important, they are honest.

Hence $\langle b_n \rangle$ is a stable sequence.

(ii)

$$
\begin{aligned}
r w_{n,i} - \sum_{j=1}^{r} w_{n,j} &\leq b_n \quad (w_{n,i} \leq w_{n,r}) \\
\Rightarrow \quad \frac{1}{n}\left(r w_{n,i} - \sum_{j=1}^{r} w_{n,j} \right) &\leq (r-1)\frac{s_{n-1}}{n} + \frac{x_n}{n} + \frac{r w_{1,r}}{n} \xrightarrow[\infty]{n} 0
\end{aligned}
$$

because $\frac{s_{n-1}}{n}$ is the difference of two sums, each of which converges to $\mathrm{E}(s_0) < \infty$ and $\frac{x_n}{n} \xrightarrow[\infty]{n} 0$ from (b).

Combining (a) and (c)(i) we obtain

$$
\begin{aligned}
\lim_{y \to \infty} \liminf_{n \to \infty} P[w_{n,i} - w_{n,1} \leq y] &\geq \lim_{y \to \infty} \liminf_{n \to \infty} P[b_n \leq y] \\
&\geq 1 \\
\Longrightarrow \quad \lim_{y \to \infty} \liminf_{n \to \infty} P[w_{n,i} - w_{n,1} \leq y] &= 1 \\
\Rightarrow \quad w_{n,i} - w_{n,1} \text{ is stable } \forall\, i \in \mathbb{N}_r &
\end{aligned}
$$

In particular $w_{n,r} - w_{n,1} = \max_{i \in \mathbb{N}_r} w_{n,i} - w_{n,1}$ is stable. Consequently the stability of $w_{n,r}$ implies the stability of $w_{n,1}$ and conversely. Now Lemma 2.4.1 implies $\boldsymbol{w}_n$ is stable $\Leftrightarrow w_{n,r}$ is stable. Finally, therefore, $\boldsymbol{w}_n$ is stable $\Leftrightarrow w_{n,1}$ is stable.

$\square$

Theorem 2.4.1

Consider a $G/G/r$ queue with CAP $[\varphi, L(\varphi)]$. Write $u_n = s_n - r\tau_n$ and $\rho = \frac{\mathrm{E}(s_0)}{r\mathrm{E}(\tau_0)}$.

(a) The stability of the multiserver queue implies the stability of the single-server queue with $\varphi'_n = (r\tau_n, s_n)$

(b) If $\rho > 1$, then the multiserver queue cannot be stable

(c) For all values of ρ

$$\frac{rw_{n,i}}{n} \xrightarrow[\infty]{n} [\mathrm{E}(u_0)]^+ \text{ a.e. } \quad \forall\, i \in \mathbb{N}_r$$

(d) If $\rho = 1$, then there exists a minimal stationary sequence $\langle \boldsymbol{m}_n = \boldsymbol{m}_n(\varphi) \rangle$ satisfying

$$\boldsymbol{m}_{n+1} = \boldsymbol{R}(\boldsymbol{m}_n + s_n \boldsymbol{e}) - \tau_n \boldsymbol{1} \tag{2.4.7}$$

(e) If $\rho < 1$ (or, equivalently, $\mathrm{E}(u_0) < 0$), then the multiserver queue is stable and there exists a minimal a.e. finite stationary sequence $\langle \boldsymbol{m}_n = \boldsymbol{m}_n(\varphi) \rangle$ (so $\boldsymbol{m}_n \in \mathcal{W}_r$ a.s.) such that

$$\boldsymbol{m}_{n+1} = \boldsymbol{f}(\boldsymbol{m}_n, s_n, \tau_n)$$

PROOF

From Lemma 2.3.1, we know that (the vector function) $\boldsymbol{f}$ satisfies the properties stipulated in Lemma 1.2.2. Consequently, as stated in the introduction to the present chapter, most of Theorem 1.2.1 remains valid in r-dimensional Euclidean space.

We first prove some results of general interest in the theorem. Initially we need to show that

$$\sum_{j=1}^{r} w_{n+1,j} \geq \left[\sum_{j=1}^{r} w_{n,j} + u_n\right]^+ \tag{2.4.8}$$

for any n. To do so, let $\boldsymbol{w}_{n+1}$ be as in (2.2.5). Then

$$\begin{aligned}
\left[\sum_{j=1}^{r} w_{n,j} + u_n\right]^+ &= \left[\sum_{j=1}^{r} w_{n,j} + s_n - r\tau_n\right]^+ \\
&= \left[(w_{n,1} + s_n - \tau_n) + \sum_{j=2}^{r}(w_{n,j} - \tau_n)\right]^+ \\
&\leq (w_{n,1} + s_n - \tau_n)^+ + \sum_{j=2}^{r}(w_{n,j} - \tau_n)^+ \qquad (=) \quad \text{Lemma 2.3.2} \\
&\leq \sum_{j=1}^{k-1}(w_{n,j+1} - \tau_n)^+ + (w_{n,1} + s_n - \tau_n)^+ \\
&\quad + \sum_{j=k+1}^{r}(w_{n,j} - \tau_n)^+ \qquad (=) \\
&\leq \sum_{j=1}^{r} w_{n+1,j}
\end{aligned}$$

Next we prove that

$$\sum_{j=1}^{r} w_{n+1,j} \geq \sum_{m=1}^{n} u_m \quad \forall\, n \in \mathbb{N} \tag{2.4.9}$$

From (2.4.8), if $n = 1$

$$\sum_{j=1}^{r} w_{2,j} \geq \left[\sum_{j=1}^{r} w_{1,j} + u_1\right]^+ \geq u_1^+ \geq \sum_{m=1}^{1} u_m$$

Suppose (2.4.9) to be true for some $n \in \mathbb{N}$. Then

$$\begin{aligned}
\sum_{j=1}^{r} w_{n+2,j} &\geq \left[\sum_{j=1}^{r} w_{n+1,j} + u_{n+1}\right]^{+} && \text{by (2.4.8)} \\
&\geq \left[\sum_{m=1}^{n} u_m + u_{n+1}\right]^{+} && (=) \ \text{by induction} \\
&\geq \left(\sum_{m=1}^{n+1} u_m\right)^{+} \\
&\geq \sum_{m=1}^{n+1} u_m
\end{aligned}$$

Hence, from (2.4.9)

$$\begin{aligned}
\liminf_{n\to\infty} \tfrac{1}{n+1} \sum_{j=1}^{r} w_{n+1,j} &\geq \liminf_{n\to\infty} \tfrac{1}{n+1} \sum_{m=1}^{n} u_m \quad (=) \\
&\geq \lim_{n\to\infty} \tfrac{1}{n} \sum_{m=1}^{n} u_m \qquad (=) \\
&\geq \mathrm{E}(u_0) \qquad \text{Theorem 0.2.2}
\end{aligned}$$

Furthermore, from Lemma 2.4.2(c)(ii)

$$\begin{aligned}
0 &= \lim_{n\to\infty} \tfrac{1}{n+1} \left(r w_{n+1,i} - \sum_{j=1}^{r} w_{n+1,j} \right) \\
&= \liminf_{n\to\infty} \tfrac{1}{n+1} \left(r w_{n+1,i} - \sum_{j=1}^{r} w_{n+1,j} \right) \\
&= \liminf_{n\to\infty} \tfrac{r w_{n+1,i}}{n+1} - \liminf_{n\to\infty} \tfrac{1}{n+1} \sum_{j=1}^{r} w_{n+1,j} \\
&\leq \liminf_{n\to\infty} \tfrac{r w_{n+1,i}}{n+1} - \mathrm{E}(u_0)
\end{aligned}$$

$$\Longrightarrow \liminf_{n\to\infty} \tfrac{r w_{n,i}}{n} \geq \mathrm{E}(u_0) \ \text{a.s.} \tag{2.4.10}$$

(where we may proceed directly to (2.4.10) if the limit on the left is infinite).

(a) Suppose that the multiserver queue is stable. From Lemma 1.2.1, there exists a proper distribution G such that

$$\begin{aligned} G(x) &\leq P\left[\sum_{j=1}^{r} w_{n+1,j} \leq x\right] \\ &\leq P\left[\left(\sum_{j=1}^{r} w_{n,j} + u_n\right)^{+} \leq x\right] \quad \text{from (2.4.8)} \end{aligned}$$

which yields the stability of the single-server queue.

(b) When $\rho > 1$, then Theorem 1.3.1 implies instability of the single-server queue with $\varphi_n' = (r\tau_n, s_n)$ which, in conjunction with (a) above, means that the multiserver queue must also be unstable.

(c) When $\mathrm{E}(u_0) > 0$, (2.4.10) allows us to conclude that $w_{n,i} \xrightarrow[\infty]{n} \infty \;\; \forall\, i \in \mathbb{N}_r$.

Consequently, for sufficiently large $n \geq N$, the positive part in (2.2.5) becomes redundant and we may write

$$w_{n+1,i} = \begin{cases} w_{n,i+1} - \tau_n & i = 1, \ldots, k-1 \\ w_{n,1} + s_n - \tau_n & i = k \\ w_{n,i} - \tau_n & i = k+1, \ldots, r \end{cases}$$

$$\begin{aligned} \Rightarrow \quad \sum_{j=1}^{r} w_{n+1,j} &= \sum_{j=1}^{r} w_{n,j} + u_n \\ \Rightarrow \quad \sum_{j=1}^{r} w_{n+1,j} - \sum_{j=1}^{r} w_{N,j} &= \sum_{m=N}^{n} u_m && \text{from Lemma 0.2.2} \\ \Rightarrow \quad \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{r} w_{n,j} &= \mathrm{E}(u_0) && \text{by Theorem 0.2.2} \\ \Rightarrow \quad \lim_{n\to\infty} \frac{r w_{n,i}}{n} &= \mathrm{E}(u_0) \text{ a.s. } \forall\, i \in \mathbb{N}_r && \text{by Lemma 2.4.2(c)} \end{aligned}$$

Recall how Theorem 1.3.1(d) ($\frac{w_n(\alpha)}{n} \underset{\infty}{\overset{n}{\rightarrow}} \mathrm{E}(u_0)$ a.s. for the supercritical queue) was used to obtain the result $\frac{w_n(\alpha)}{n} \underset{\infty}{\overset{n}{\rightarrow}} 0$ a.s. for the critical and subcritical queues. Using this method, we reach the analogous conclusion $\frac{1}{n}\sum_{j=1}^{r} w_{n,i} \underset{\infty}{\overset{n}{\rightarrow}} 0$ leading to $\frac{rw_{n,i}}{n} \underset{\infty}{\overset{n}{\rightarrow}} 0 \;\; \forall\, i \in \mathbb{N}_r$. Hence the result.

(d) From Lemma 2.4.2, we know that $w_{n,1}$ and $\boldsymbol{w}_n$ are stable or unstable together. For certain initial conditions (especially if the servers are all initially free) $L(w_{n,1}) \underset{\infty}{\overset{n}{\rightarrow}} L(m_{0,1})$ and $L(\boldsymbol{w}_n) \underset{\infty}{\overset{n}{\rightarrow}} L(\boldsymbol{m}_0)$.

Suppose, in particular, that both $\langle w_{n,1}\rangle$ and $\langle \boldsymbol{w}_n\rangle$ are stable. Then there exist proper distribution functions G_1 and G such that

$$\left\{\begin{array}{lcl} G_1(x) & \leq & P[w_{n,1} \leq x] \\ G(\boldsymbol{x}) & \leq & P[\boldsymbol{w}_n \leq \boldsymbol{x}] \end{array}\right. \quad \forall\, n, \;\; \forall\, x$$

$$\Rightarrow \left\{\begin{array}{lclcl} G_1(x) & \leq & \lim\limits_n P[w_{n,1} \leq x] & = & P[m_{0,1} \leq x] \\ G(\boldsymbol{x}) & \leq & \lim\limits_n P[\boldsymbol{w}_n \leq \boldsymbol{x}] & = & P[\boldsymbol{m}_0 \leq \boldsymbol{x}] \end{array}\right.$$

so both $m_{0,1}$ and $\boldsymbol{m}_0$ are stable. The case of instability follows in a similar way, with the result that $m_{0,1}$ and $\boldsymbol{m}_0$ are honest or dishonest together. Using this fact, in conjunction with stationarity, we conclude that, for any n, $P[\boldsymbol{m}_n \text{ infinite}] = P[\boldsymbol{m}_0 \text{ infinite}]$.

Let $A = [m_{n,i} = \infty \;\forall\, i \in \mathbb{N}_r]$. Since $m_{n,1} = \min\limits_{i \in \mathbb{N}_r} m_{n,i}$, we may write

$A = [m_{n,1} = \infty]$. Now

$$\begin{aligned} P(A \Delta T^{-1} A) &= P[\{m_{n-1,1} = \infty\} \cap \{m_{n,1} < \infty\}] \\ &\quad + P[\{m_{n-1,1} < \infty\} \cap \{m_{n,1} = \infty\}] \\ &= 0 \end{aligned}$$

since

$$\begin{aligned} m_{n,1} &= \min[(m_{n-1,1} + s_{n-1} - \tau_{n-1})^+, (m_{n-1,2} - \tau_{n-1})^+] \\ & \begin{cases} = \infty & \text{if } m_{n-1,1} = \infty \ (\Rightarrow m_{n-1,2} = \infty) \\ < \infty & \text{if } m_{n-1,1} < \infty \end{cases} \end{aligned}$$

Thus A is an invariant event (Definition 0.2.4) and with $\langle (s_n, \tau_n) \rangle$ ergodic, this means that $P(A) = 0$ or 1. Consequently $\boldsymbol{m}_0$ is a proper random vector or each one of its components is infinite. In the latter case (2.4.7) is obviously true.

Next consider the case of $\boldsymbol{m}_0$ proper. With $\boldsymbol{m}_1$ honest (because $\boldsymbol{m}_0$ is) we may have $\boldsymbol{w}_1 = \boldsymbol{m}_1 \Rightarrow \boldsymbol{w}_n = \boldsymbol{m}_n$ a.e. Using (c), therefore, gives us $\frac{r m_{n,r}}{n} \xrightarrow[\infty]{n} 0$ a.s.

Much as in Theorem 1.3.3(a), let

$$\tilde{u}_n = \sum_{j=1}^{r} m_{n+1,j} - \sum_{j=1}^{r} m_{n,j} \geq u_n$$

which leads to $\tilde{u}_n = u_n$ a.s. $(\forall\, n)$ in an analogous way, implying

$$\begin{array}{rrcl} & \sum_{j=1}^{r} m_{n+1,j} & = & \sum_{j=1}^{r} m_{n,j} + u_n \\ \Rightarrow & (m_{n,1} + s_n - \tau_n)^+ + \sum_{j=2}^{r} (m_{n,j} - \tau_n)^+ & = & (m_{n,1} + s_n - \tau_n) \\ & & & + \sum_{j=2}^{r} (m_{n,j} - \tau_n) \\ \Rightarrow & (x_{n,1}^+ - x_{n,1}) + \sum_{j=2}^{r} (x_{n,j}^+ - x_{n,j}) & = & 0 \end{array}$$

where $$x_{n,j} = \begin{cases} m_{n,1} + s_n - \tau_n & j = 1 \\ m_{n,j} - \tau_n & j \in \mathbb{N}_r \setminus \{1\} \end{cases}$$

but $$x_{n,j}^+ \geq x_{n,j} \quad \forall\, j \in \mathbb{N}_r$$

$\Rightarrow$ $$x_{n,j}^+ = x_{n,j} \quad \forall\, j \in \mathbb{N}_r$$

$\Longrightarrow$ $$\boldsymbol{m}_{n+1} = \boldsymbol{R}(\boldsymbol{m}_n + s_n \boldsymbol{e}) - \tau_n \mathbf{1}$$

(e) As in (d), the remark at the start of the proof establishes part of the desired result. Now we will be concerned mostly with stability and the honesty of $\langle \boldsymbol{m}_n \rangle$.

$\boldsymbol{w}_1 \leq w_{1,r}\mathbf{1} \leq \boldsymbol{m}_1 + w_{1,r}\mathbf{1}$ since $\boldsymbol{m}_1 \geq \mathbf{0}$
If we now assume that $\boldsymbol{w}_n \leq \boldsymbol{m}_n + w_{1,r}\mathbf{1}$ for some n, we obtain

$$\begin{array}{rcll} \boldsymbol{w}_{n+1} & = & \boldsymbol{f}(\boldsymbol{w}_n, s_n, \tau_n) & \\ & \leq & \boldsymbol{f}(\boldsymbol{m}_n + w_{1,r}\mathbf{1}, s_n, \tau_n) \qquad (=) & \text{by monotonicity} \\ & \leq & [[\boldsymbol{R}(\boldsymbol{m}_n + s_n \boldsymbol{e}) - \tau_n \mathbf{1}] + w_{1,r}\mathbf{1}]^+ & \\ & \leq & \boldsymbol{m}_{n+1} + w_{1,r}\mathbf{1} & \text{by Lemma 2.3.2} \end{array}$$

Consequently, by induction

$$\boldsymbol{w}_n \;\leq\; \boldsymbol{m}_n + w_{1,r}\mathbf{1} \quad \forall\, n \in \mathbb{N}$$

If $\boldsymbol{m}_n$ is honest $\forall\, n$ (or simply, $\boldsymbol{m}_0$ is honest because of stationarity), then

$$G(\boldsymbol{x}) \;=\; P[\boldsymbol{m}_n + w_{1,r}\mathbf{1} \leq \boldsymbol{x}] \;\leq\; P[\boldsymbol{w}_n \leq \boldsymbol{x}] \quad \forall\, n, \;\; \forall\, \boldsymbol{x}$$

where G is a proper distribution function because $\boldsymbol{w}_1$ is also honest. By Lemma 1.2.1 and Remark 1.2.1, $\boldsymbol{w}_n$ is stable. Let us, therefore, show that $\boldsymbol{m}_0$ is honest. For the sake of simplicity, but w.l.g., we consider the situation in which $\boldsymbol{w}_1 = \mathbf{0}$.

Recall the random variable $b_n = rw_{n,r} - \sum_{j=1}^{r} w_{n,j} \leq (r-1)s_{n-1} + x_n$ (Lemma 2.4.2(b) with $w_{1,r} = 0$), and define another, related nonnegative random variable

$$\begin{aligned}
d_n &= \sum_{j=1}^{r} w_{n,j} - rw_{n,1} \\
&= (rw_{n,r} - b_n) - rw_{n,1} \\
&= r(w_{n,r} - w_{n,1}) - b_n \\
&\leq (r-1)b_n \qquad\qquad w_{n,r} - w_{n,1} \leq b_n \text{ by Lemma 2.4.2(a)}
\end{aligned}$$

Now $\exists\, m \in \mathbb{N}_n$; $\;w_{m,1} = 0$ but $w_{l,1} > 0 \;\forall\, l \in \mathbb{N}_n \setminus \mathbb{N}_m$ (i.e., there is a last one of $w_{i,1}$ $(i \in \mathbb{N}_n)$ which vanishes). (In particular, if $w_{n,1} = 0$ *and* $w_{n+1,1} = 0$, then $m = n$.)

Thus $d_m = \sum_{j=1}^{r} w_{m,j}$ and since $0 < w_{l,1} = \min_{i \in \mathbb{N}_r} w_{l,i}$, we have that $\forall\, l \in \mathbb{N}_n \setminus \mathbb{N}_m$

$$\begin{aligned}
\boldsymbol{w}_{l+1} &= \boldsymbol{R}(\boldsymbol{w}_l + s_l\boldsymbol{e}) - \tau_l \mathbf{1} && \forall\, l \in \{m, \ldots, n-1\} \\
\Rightarrow \quad \sum_{j=1}^{r} w_{l+1,j} &= \sum_{j=1}^{r} w_{l,j} + u_{n-1} && \forall\, l \in \{m, \ldots, n-1\} \\
\Rightarrow \quad \sum_{j=1}^{r} w_{n,j} &= \sum_{j=1}^{r} w_{m,j} + \sum_{l=m}^{n-1} u_l && \text{by Lemma 0.2.2} \\
&= d_m + \sum_{l=m}^{n-1} u_l
\end{aligned}$$

(which is obvious if $m = n$).

Thus

$$\begin{aligned}
\sum_{j=1}^{r} w^0_{n,j} &= \sum_{j=1}^{r} w_{n,j} \\
&= \sup_{m \in \mathbb{N}_n} \sum_{j=1}^{r} w_{n,j} \\
&= \sup_{m \in \mathbb{N}_n} \left[d_m + \sum_{l=m}^{n} u_l - u_n \right] \\
&= \sup_{m \in \mathbb{N}_n} \left[d_m + \sum_{l=m}^{n} u_l \right] - u_n \\
&\leq \sup_{m \in \mathbb{N}_n} \left[(r-1) b_m + \sum_{l=m}^{n} u_l \right] - u_n \\
&\leq \sup_{m \in \mathbb{N}_n} \left[(r-1)[(r-1)s_{m-1} + x_m] + \sum_{l=m}^{n} u_l \right] - u_n
\end{aligned}$$

using the bounds for d_m and b_m in the second-last and last steps respectively.

It follows that

$$
\begin{aligned}
& \sum_{j=1}^{r} w_{0,j}^{n} + u_0 \\
\leq \quad & \sup_{m \in \mathbb{N}_n} \left[(r-1)^2 s_{m-1-n} + (r-1) x_{m-n} + \sum_{l=m-n}^{0} u_l \right] && (=) \\
\leq \quad & \sup_{0 \leq n-m \leq n-1} \left[(r-1)^2 s_{-(n-m)-1} + (r-1) x_{-(n-m)} + \sum_{l=0}^{n-m} u_{-l} \right] && (=) \\
\leq \quad & \sup_{0 \leq m \leq n-1} \left[(r-1)^2 s_{-m-1} + (r-1) x_{-m} + \sum_{l=0}^{m} u_{-l} \right] && (=)
\end{aligned}
$$

$$
\begin{aligned}
\Longrightarrow \qquad \sum_{j=1}^{r} m_{0,j} \quad &= \quad \lim_{n \to \infty} \sum_{j=1}^{r} w_{0,j}^{n} \\
&\leq \quad \sup_{m \in \mathbb{Z}_{\oplus}} z_m - u_0 \\
&\leq \quad \sup_{m \in \mathbb{Z}_{\oplus}} \left[y_m + \sum_{l=0}^{m} u_{-l} \right] - u_0
\end{aligned}
$$

where $y_m = (r-1)^2 s_{-m-1} + (r-1) x_{-m}$.

Next observe that $\lim_{m \to \infty} \frac{y_m}{m} = 0$ for essentially the same reasoning used to obtain the result in Lemma 2.4.2(c)(ii). The law of large numbers gives $\lim_{m \to \infty} \frac{1}{m} \sum_{l=0}^{m} u_{-l} = \mathrm{E}(u_0)$ a.e. Thus $\lim_{m \to \infty} z_m = -\infty$. Consequently $\sup_{m \in \mathbb{Z}_{\oplus}} z_m$ is finite (because it is attained for finite m). This implies that $\boldsymbol{m}_0$ is honest and, as shown earlier, we are then able to deduce that $\boldsymbol{w}_n$ is stable. Thus the multiserver queue is stable (using Definition 2.4.1 and Lemma 2.4.1 in conjunction with Lemma 2.4.2).

□

Example 2.4.1 (Converse of Theorem 2.4.1(a))

We now illustrate that, although the stability of the single-server queue with $\varphi'_n = (r\tau_n, s_n)$ is a necessary condition for the stability of the associated multiserver queue, it is not sufficient.

Suppose $u_n = s_n - r\tau_n = 0 \ \forall n$ a.s. where $\langle \tau_n \rangle$ is a sequence of stationary positive mutually independent unbounded random variables with finite mean $\mathrm{E}\tau$. From Theorem 1.3.4 and Remark 1.3.7, we deduce that the single-server queue is stable. Obviously we also have $\mathrm{E}(u_0) = 0$ so that (2.4.7) holds. Thus

$$\begin{array}{rcll} 0 & \leq & m_{n+1,1} & \\ & \leq & \min[m_{n,1} + s_n - \tau_n, m_{n,2} - \tau_n] \quad (=) & \text{by } (2.2.4) \\ & \leq & m_{n,2} - \tau_n & \\ \Longrightarrow \tau_n & \leq & m_{n,2} & \end{array}$$

Also $m_{n,2}$ is independent of τ_n (in part a consequence of the fact that $m_{n,2} = \max[m_{n-1,1} + s_{n-1} - \tau_{n-1}, m_{n-1,2} - \tau_{n-1}]$). Combining these facts, we obtain

$$\begin{array}{rcll} 0 & = & P[m_{n,2} < c < \tau_n] & \text{since } \tau_n \leq m_{n,2} \\ & = & P[m_{n,2} < c]P[\tau_n > c] & \text{by independence} \\ \Rightarrow \ P[m_{n,2} < c] & = & 0 \ \forall n, \ \forall c > 0 & \langle \tau_n \rangle \text{ an unbounded sequence} \end{array}$$

$\Rightarrow$ $m_{n,2}$ is dishonest

$\Rightarrow$ $\boldsymbol{m}_n$ is dishonest

$\Rightarrow$ the multiserver queue is unstable

At this point, it is worthwhile to establish a connection between the notation of Loynes [47] and Brandt [15].

As indicated in Theorem 2.4.1, the existence of the minimal (vector) sequence $\langle \boldsymbol{m}_n, n \in \mathbb{Z} \rangle$ is determined by a process analogous to that in Lemma 1.2.2 and Theorem 1.2.1. So we would have

$$\begin{aligned} \boldsymbol{m}_n &= \lim_{m \to \infty} \boldsymbol{w}_n^m \quad \text{where} \\ \boldsymbol{w}_n^m &= \begin{cases} \mathbf{0} & n \leq 1 - m \\ \boldsymbol{f}\left(\boldsymbol{w}_{n-1}^m, s_{n-1}, \tau_{n-1}\right) & n > 1 - m \end{cases} \end{aligned} \tag{2.4.11}$$

As usual, let $\varphi = \langle \varphi_n, n \in \mathbb{Z} \rangle$ where $\varphi_n = (\tau_n, s_n) \in \mathbb{R}_\oplus^2$, and $\boldsymbol{w} \in \mathcal{W}_r$.

Denote by $\langle \boldsymbol{w}_n(\boldsymbol{w}, \varphi), n \in \mathbb{N} \rangle$ the sequence given by

$$\boldsymbol{w}_n(\boldsymbol{w}, \varphi) = \begin{cases} \boldsymbol{w} & n = 1 \\ \boldsymbol{f}\left(\boldsymbol{w}_{n-1}(\boldsymbol{w}, \varphi), s_{n-1}, \tau_{n-1}\right) & n \in \overline{\mathbb{N}}_1 \end{cases} \tag{2.4.12}$$

which may be equated with Loynes's $\langle \boldsymbol{w}_n(\boldsymbol{w}), n \in \mathbb{N} \rangle$. In particular, when $\boldsymbol{w} = \mathbf{0}$, we could write $\left\langle \boldsymbol{w}_n^{\mathbf{0}}, n \in \mathbb{N} \right\rangle$ (as indicated in Remark 1.2.2).

Returning to (2.4.12), we are able to deduce that $\boldsymbol{w}_k(\boldsymbol{w}, T^l\varphi)$ is the workload vector at the arrival moment of the $(k + l)$th customer, determined by starting the queue at the time of arrival of the $(l + 1)$st customer ($(k - 1)$ arrivals previously) with the initial condition $\boldsymbol{w}$. (2.4.13)

To see why this is the case, note that

$$\begin{aligned}
T^l\varphi_n &= (\tau_{n+l}, s_{n+l}) \\
\boldsymbol{w}_1(\boldsymbol{w}, T^l\varphi) &= \boldsymbol{w} \\
\boldsymbol{w}_2(\boldsymbol{w}, T^l\varphi) &= \boldsymbol{f}\left(\boldsymbol{w}, T^l s_1, T^l \tau_1\right) \\
&= \boldsymbol{f}\left(\boldsymbol{w}, s_{l+1}, \tau_{l+1}\right)
\end{aligned}$$

$\Longrightarrow$ initial condition $\boldsymbol{w}$ is the workload vector of the $(l+1)$st customer

$\Rightarrow$ $\boldsymbol{w}_k(\boldsymbol{w}, T^l\varphi)$ is the workload vector for $(k-1)$ customers thereafter, i.e., for the $(k+l)$th customer

Obviously we are also able to write

$$\boldsymbol{w}_{k+1}(\boldsymbol{w}, T^l\varphi) = \boldsymbol{f}\left(\boldsymbol{w}_k(\boldsymbol{w}, T^l\varphi), s_{k+l}, \tau_{k+l}\right) \tag{2.4.14}$$

From (2.4.11), we deduce that $\boldsymbol{w}_n^m$ is the workload vector, at the moment of arrival of the nth customer, determined by starting the queue with workload $\mathbf{0}$ at the arrival epoch of the $(1-m)$th customer. On applying (2.4.13), we obtain

$$\boldsymbol{w}_n^m = \boldsymbol{w}_{m+n}(\mathbf{0}, T^{-m}\varphi) \tag{2.4.15}$$

or, by change of variables ($m = j + 1 - n$ (which $\to \infty$ as $j \to \infty$))

$$\left.\begin{aligned}
\boldsymbol{w}_n^{j+1-n} &= \boldsymbol{w}_{j+1}(\mathbf{0}, T^{n-j-1}\varphi) \\
\text{and} \qquad \boldsymbol{m}_n &= \lim_{j\to\infty} \boldsymbol{w}_{j+1}(\mathbf{0}, T^{n-j-1}\varphi)
\end{aligned}\right\} \quad j \in \mathbb{Z}_{\oplus},\ n \in \mathbb{Z} \tag{2.4.16}$$

From now on we will tend to use the latter notation, in order to fit in with Brandt [15]. In this spirit, we prove the vector analogue of Lemma 1.2.2 in a more elegant fashion.

Lemma 2.4.3

$$\boldsymbol{w}_{j+1}(\mathbf{0}, T^{n-j-1}\varphi) \leq \boldsymbol{w}_{j+2}(\mathbf{0}, T^{n-j-2}\varphi) \quad \forall\, j \in \mathbb{Z}_{\oplus},\ \forall\, n \in \mathbb{Z} \tag{2.4.17}$$

<u>PROOF</u>

When $j = 0$, we have

$$\begin{aligned} \boldsymbol{w}_2(\mathbf{0}, T^{n-2}\varphi) &= \boldsymbol{f}\left(\boldsymbol{w}_1(\mathbf{0}, T^{n-2}\varphi), s_{n-1}, \tau_{n-1}\right) \\ &= \boldsymbol{f}(\mathbf{0}, s_{n-1}, \tau_{n-1}) \\ &\geq \mathbf{0} && (=) \\ &\geq \boldsymbol{w}_1(\mathbf{0}, T^{n-1}\varphi) && \forall\, n \in \mathbb{Z} \end{aligned}$$

Suppose (2.4.17) to be true for some (fixed) $j \in \mathbb{Z}_{\oplus}$. Then

$$\begin{aligned} \boldsymbol{w}_{j+3}(\mathbf{0}, T^{n-j-3}\varphi) &= \boldsymbol{f}\left(\boldsymbol{w}_{j+2}(\mathbf{0}, T^{n-j-3}\varphi), s_{n-1}, \tau_{n-1}\right) \\ &= \boldsymbol{f}\left(\boldsymbol{w}_{j+2}(\mathbf{0}, T^{(n-1)-j-2}\varphi), s_{n-1}, \tau_{n-1}\right) \\ &\geq \boldsymbol{f}\left(\boldsymbol{w}_{j+1}(\mathbf{0}, T^{(n-1)-j-1}\varphi), s_{n-1}, \tau_{n-1}\right) \quad (=) \\ &\geq \boldsymbol{w}_{j+2}(\mathbf{0}, T^{n-j-2}\varphi) \end{aligned}$$

Hence (2.4.17) follows by induction.

□

Remark 2.4.1

Several of the other results used in Theorem 2.4.1 could also be proved in a similar way, using this different notation.

In the remaining sections of this chapter we will make the assumptions (some of which have been stipulated already) that

$$\begin{array}{rrcll} & \rho & = & \frac{\mathrm{E}s_0}{r\mathrm{E}\tau_0} & < \ 1 \\ \text{where} & & & & \\ & 0 & < & \mathrm{E}(s_0) & < \ \infty \\ & 0 & < & \mathrm{E}(\tau_0) & < \ \infty \\ \Longrightarrow & -\infty & < & \mathrm{E}(u_0) & < \ \infty \end{array}$$

By a solution of (2.2.6) we shall mean a finite (and stationary) solution (an element of $\mathcal{W}_r$).

2.5 THE EXISTENCE OF NON-MINIMAL SOLUTIONS

In Theorem 1.3.2, the uniqueness of the a.e. finite stationary minimal solution $\langle m_n \rangle$ to the equation $w_{n+1} = (w_n + u_n)^+$ was established for a subcritical $G/G/1$ queue.

The example which follows shortly illustrates that this uniqueness cannot be guaranteed for a subcritical $G/G/r$ queue. Before presenting this example, it is convenient to formalise the concept of a solution to (2.2.6).

Definition 2.5.1 (Weak and Strong Solutions)
Let $\varphi = \langle \varphi_n, n \in \mathbb{Z} \rangle$ ($\varphi_n = (\tau_n, s_n)$), and consider CAP $[\varphi, L(\varphi)]$. Analogously, let $\tilde{\varphi} = \langle \tilde{\varphi}_n, n \in \mathbb{Z} \rangle$ ($\tilde{\varphi}_n = (\tilde{\tau}_n, \tilde{s}_n)$), and write $\phi = \langle \phi_n; n \in \mathbb{Z} \rangle = \langle \tilde{\varphi}_n, \widetilde{\boldsymbol{w}}_n \rangle$ for the corresponding stationary sequence with distribution $L(\phi)$. Then $[\phi, L(\phi)]$ shall be said to be a *weak solution of* (2.2.6) *for the CAP* $[\varphi, L(\varphi)]$ iff

$$\begin{aligned} \tilde{\varphi} &=_d \varphi \\ \widetilde{\boldsymbol{w}}_{n+1} &= \boldsymbol{f}(\widetilde{\boldsymbol{w}}_n, \tilde{s}_n, \tilde{\tau}_n) \quad \forall n \in \mathbb{Z} \text{ a.s.} \end{aligned} \tag{2.5.1}$$

In addition, the weak solution $[\phi, L(\phi)]$ shall be called a *strong solution of* (2.2.6) *for the CAP* $[\varphi, L(\varphi)]$ iff $\exists$ sequence $\langle h_n; n \in \mathbb{Z} \rangle$ of measurable functions

$$\left.\begin{aligned} & h_n : \left(\mathbb{R}^2_{\oplus}\right)^{\mathbb{Z}} \to \mathcal{W}_r \\ \text{such that} \quad & \widetilde{\boldsymbol{w}}_n = h_n(\tilde{\varphi}) \qquad \text{a.s.} \end{aligned}\right\} \quad \forall n \in \mathbb{Z} \tag{2.5.2}$$

Remark 2.5.1

(a) Since a strong solution is by definition also a weak solution, it is convenient to use the term *properly weak solution* for a (weak) solution which is not strong.

(b) A weak solution $[\phi, L(\phi)]$ is unique iff for all weak solutions $[\phi', L(\phi')]$ $L(\phi) = L(\phi')$. On the other hand, $[\phi, L(\phi)]$ and $[\phi', L(\phi')]$ are *different* weak solutions iff $L(\phi) \neq L(\phi')$. Thus solutions determined by (independent) copies of identically distributed random variables are, sensibly enough, not considered to be different.

(c) The existence of more than one weak solution implies the existence of infinitely many weak solutions. We examine this issue further in Remark 2.6.4.

(d) For a subcritical $G/G/1$ queue the solution $\langle(\tilde{\tau}_n, \tilde{s}_n), m_n\rangle$ (with $(\tilde{\tau}_n, \tilde{s}_n) =_d (\tau_n, s_n)$) where $m_n = \left[\sup_{k\in\mathbb{N}} \sum_{j=1}^{k} u_{n-j}\right]^+$ is the unique strong solution (which is also the unique weak solution).

It was indicated in the introduction to this section that the uniqueness mentioned in Remark (d) above does not extend to $G/G/r$ $(r > 1)$ subcritical queues. The example below demonstrates that fact.

Example 2.5.1 (Several Strong Solutions)

Consider the CAP $[\varphi, L(\varphi)]$ with $P[\varphi = \omega] = \frac{1}{2} = P[\varphi = T\omega]$ where T is the usual shift transformation. (This assures us of stationarity.)

Let a $G/G/2$ queue be determined by $[\varphi, L(\varphi)]$ such that (with a slight abuse of notation)

$$\begin{aligned} \tau_n(\omega) &= \tau > 0 & \forall\, n \in \mathbb{Z} \\ s_n(\omega) &= \begin{cases} \sigma_1 & n = 2k \\ \sigma_2 & n = 2k+1 \end{cases} & k \in \mathbb{Z} \end{aligned}$$

where

$$\tau \ < \ \sigma_1 \ < \ \sigma_2 \ = \ 2\tau \tag{2.5.3}$$

$$\Rightarrow \rho \ = \ \frac{\mathrm{E}s}{2\mathrm{E}\tau} \ = \ \tfrac{1}{4}(\sigma_1 + \sigma_2) \ < \ 1$$

so that $\tau_n(T\omega) = \tau$ and $s_n(T\omega) = s_{n+1}(\omega)$.

We now claim that the waiting-time sequence $\langle \widetilde{\boldsymbol{w}}_n^a(\omega), n \in \mathbb{Z} \rangle$ (more especially, a sequence of measurable functions) given by

$$\widetilde{\boldsymbol{w}}_n^a(\omega) = \begin{cases} (0, a) & n = 2k \\ (a - \tau, \sigma_1 - \tau) & n = 2k+1 \end{cases} \quad k \in \mathbb{Z},\ a \in [\tau, \sigma_1] \tag{2.5.4}$$

is a solution to the equation

$$\boldsymbol{w}_{n+1}(\omega) = [\boldsymbol{R}(\boldsymbol{w}_n(\omega) + s_n(\omega)\boldsymbol{e}) - \tau_n(\omega)\mathbf{1}]^+ \tag{2.5.5}$$

Suppose $\widetilde{\boldsymbol{w}}_{2k_1}^a(\omega) = (0, a)$ for some $k_1 \in \mathbb{Z}$. Then the right-hand side of (2.5.5) gives

$$\begin{aligned} [\boldsymbol{R}((0,a) + \sigma_1(1,0)) - \tau(1,1)]^+ &= [\boldsymbol{R}((\sigma_1 - \tau, a - \tau))]^+ \\ &= (a - \tau, \sigma_1 - \tau) \qquad \text{since } \tau \leq a \leq \sigma_1 \\ &= \widetilde{\boldsymbol{w}}_{2k_1+1}^a(\omega) \end{aligned}$$

Conversely, suppose $\widetilde{\boldsymbol{w}}_{2k_2+1}^a(\omega) = (a - \tau, \sigma_1 - \tau)$ for some $k_2 \in \mathbb{Z}$. It follows that

$$\begin{aligned} & \left[\boldsymbol{R}\left(\widetilde{\boldsymbol{w}}_{2k_2+1}^a(\omega) + s_{2k_2+1}(\omega)\boldsymbol{e}\right) - \tau_{2k_2+1}(\omega)\mathbf{1}\right]^+ \\ = \ & [\boldsymbol{R}((a + \sigma_2 - 2\tau, \sigma_1 - 2\tau))]^+ \\ = \ & (0, a) \qquad \text{by (2.5.3)} \\ = \ & \widetilde{\boldsymbol{w}}_{2k_2+2}^a(\omega) \qquad \text{as required} \end{aligned}$$

In a completely analogous fashion $\widetilde{\boldsymbol{w}}_n^a(T\omega) = \widetilde{\boldsymbol{w}}_{n+1}^a(\omega)$ satisfies (2.5.4), as should be expected.

Next we note that $\widetilde{\boldsymbol{w}}_{n-1} = (0,0)$ is transformed by (2.5.5) into

$$\widetilde{\boldsymbol{w}}_n(\omega) = \left\{ \begin{array}{ll} (0, \sigma_2 - \tau) = (0, \tau) & n = 2k \\ (0, \sigma_1 - \tau) & n = 2k+1 \end{array} \right\} = \widetilde{\boldsymbol{w}}_n^\tau(\omega)$$

If we write $\phi^a = \langle(\tau_n, s_n), \widetilde{\boldsymbol{w}}_n^a(\omega)\rangle$, then $[\phi^a, L(\phi^a)]$ is a strong solution for our complete arrival process. In particular we deduce from the previous discussion that $[\phi^\tau, L(\phi^\tau)] = \phi_{\min}$, the minimal solution, and $\forall\, a \in (\tau, \sigma_1]$ (which is a nonempty interval by (2.5.3)) $L(\phi^a) \neq L(\phi_{\min})$. In accordance with Definition 2.5.1, we conclude that there are several (in fact, a continuum of) strong solutions.

The next example illustrates the difference between strong and properly weak solutions.

Example 2.5.2 (Weak Solutions which are not Strong)
Return to the $G/G/2$ queue of Example 2.5.1, and consider a continuous random variable Y with support $[\tau, \sigma_1]$, and which is distributed independently of φ.

By comparison with (2.5.4), we obtain the following weak solution of (2.5.5):

$$\widetilde{\boldsymbol{w}}_n^Y(\omega) = \left\{ \begin{array}{ll} (0, Y) & n = 2k \\ (Y - \tau, \sigma_1 - \tau) & n = 2k+1 \end{array} \right. \quad k \in \mathbb{Z}$$

i.e., $\phi^Y = \left\langle (\tau_n, s_n), \widetilde{\boldsymbol{w}}_n^Y(\omega) \right\rangle$, together with its distribution, is a weak solution.

The randomness of Y, however, implies that $\widetilde{\boldsymbol{w}}_n^Y(\omega)$ is not a measurable function so that $[\phi^Y, L(\phi^Y)]$ is not a strong solution.

The non-uniqueness of solutions demonstrated in Example 2.5.1 raises the question of the existence of a maximal solution, as a counterpart to the minimal solution. This topic is covered in more detail in the next section where the following proposition will be used.

Proposition 2.5.1

Consider the notation used in Definition 2.5.1. Given weak solution $[\phi, L(\phi)]$, we have that $\forall\, n \in \mathbb{Z}$, almost surely

$$\text{(a)} \quad \|\boldsymbol{m}_n(\tilde{\varphi}) - \widetilde{\boldsymbol{w}}_n\|_{\max} = \|\boldsymbol{m}_1(\tilde{\varphi}) - \widetilde{\boldsymbol{w}}_1\|_{\max} \tag{2.5.6}$$

$$\text{(b)} \quad \|\boldsymbol{m}_n(\tilde{\varphi}) - \widetilde{\boldsymbol{w}}_n\| = \|\boldsymbol{m}_1(\tilde{\varphi}) - \widetilde{\boldsymbol{w}}_1\| \tag{2.5.7}$$

<u>PROOF</u>

Since $[\phi, L(\phi)]$ is a weak solution, it is a stationary solution (Definition 2.5.1) which implies that $\langle \widetilde{\boldsymbol{w}}_n; n \in \mathbb{Z} \rangle$ is a stationary sequence. From Theorem 1.2.1, $\langle \boldsymbol{m}_n(\tilde{\varphi}); n \in \mathbb{Z} \rangle$ is also stationary. Consequently $\langle \boldsymbol{m}_n(\tilde{\varphi}) - \widetilde{\boldsymbol{w}}_n \rangle$, $\langle \|\boldsymbol{m}_n(\tilde{\varphi}) - \widetilde{\boldsymbol{w}}_n\|; n \in \mathbb{Z} \rangle$ and $\langle \|\boldsymbol{m}_n(\tilde{\varphi}) - \widetilde{\boldsymbol{w}}_n\|_{\max}; n \in \mathbb{Z} \rangle$ are stationary sequences.

(a) From the above and from Remark 0.2.2, we deduce that

$$\|\boldsymbol{m}_n(\tilde{\varphi}) - \widetilde{\boldsymbol{w}}_n\|_{\max} \;=_d\; \|\boldsymbol{m}_1(\tilde{\varphi}) - \widetilde{\boldsymbol{w}}_1\|_{\max} \tag{2.5.8}$$

but the contractive property (Lemma 2.3.4) gives us

$$\|\boldsymbol{m}_n(\tilde{\varphi}) - \widetilde{\boldsymbol{w}}_n\|_{\max} \;\leq\; \|\boldsymbol{m}_1(\tilde{\varphi}) - \widetilde{\boldsymbol{w}}_1\|_{\max} \tag{2.5.9}$$

If we combine (2.5.8) and (2.5.9), we obtain

$$P\left[\|\boldsymbol{m}_n(\tilde{\varphi}) - \widetilde{\boldsymbol{w}}_n\|_{\max} = \|\boldsymbol{m}_1(\tilde{\varphi}) - \widetilde{\boldsymbol{w}}_1\|_{\max}\right] = 1$$

which is what we require.

(b) In view of our conclusion about the stationarity of $\langle\|\boldsymbol{m}_n(\tilde{\varphi}) - \widetilde{\boldsymbol{w}}_n\|; n \in \mathbb{Z}\rangle$, the proof will proceed by considering equations analogous to (2.5.8) and (2.5.9) for the ordinary, rather than the maximum, norm.

□

2.6 IN SEARCH OF A MAXIMAL SOLUTION

The construction of the maximal solution proceeds in stages, and is based on taking a pair of limits. We begin with a lemma which involves a sequence of terms of the form

$$\boldsymbol{w}_{k+1}\left(\boldsymbol{m}_{n-k}(\varphi) + q\mathbf{1}, T^{n-k-1}\varphi\right) \tag{2.6.1}$$

which, we recall ((2.4.13)) is the workload vector at the arrival moment of the nth customer, determined by starting the queue at the arrival epoch of the $(n-k)$th customer, with initial condition $\boldsymbol{m}_{n-k}(\varphi)+q\mathbf{1}$.

Lemma 2.6.1
$\left\langle \boldsymbol{w}_{k+1}\left(\boldsymbol{m}_{n-k}(\varphi)+q\mathbf{1}, T^{n-k-1}\varphi\right), k \in \mathbb{Z}_{\oplus}\right\rangle$ for $n \in \mathbb{Z}$ and $q \in \mathbb{R}_{\oplus}$ fixed is a nonincreasing sequence.

<u>PROOF</u>

Proceed by induction on k.

Initially we compare the first two terms of the sequence.

$$\begin{aligned}
\boldsymbol{w}_1\left(\boldsymbol{m}_n(\varphi)+q\mathbf{1}, T^{n-1}\varphi\right) &= \boldsymbol{m}_n(\varphi)+q\mathbf{1} \\
&= \boldsymbol{f}\left(\boldsymbol{m}_{n-1}(\varphi), s_{n-1}, \tau_{n-1}\right)+q\mathbf{1}
\end{aligned}$$

and

$$\begin{aligned}
\boldsymbol{w}_2\left(\boldsymbol{m}_{n-1}(\varphi)+q\mathbf{1}, T^{n-2}\varphi\right) &= \boldsymbol{f}\left[\boldsymbol{w}_1\left(\boldsymbol{m}_{n-1}(\varphi)+q\mathbf{1}, T^{n-2}\varphi\right), s_{n-1}, \tau_{n-1}\right] \\
&= \boldsymbol{f}\left(\boldsymbol{m}_{n-1}(\varphi)+q\mathbf{1}, s_{n-1}, \tau_{n-1}\right)
\end{aligned}$$

Now $\forall\, n \in \mathbb{Z}$ (because of the monotonicity of $\boldsymbol{f}$)

$$\begin{aligned}
\boldsymbol{f}\left(\boldsymbol{m}_{n-1}(\varphi)+q\mathbf{1}, s_{n-1}, \tau_{n-1}\right)-\boldsymbol{f}\left(\boldsymbol{m}_{n-1}(\varphi), s_{n-1}, \tau_{n-1}\right) &\leq q\mathbf{1} \\
\Rightarrow \quad \boldsymbol{w}_2\left(\boldsymbol{m}_{n-1}(\varphi)+q\mathbf{1}, T^{n-2}\varphi\right)-\boldsymbol{w}_1\left(\boldsymbol{m}_n(\varphi)+q\mathbf{1}, T^{n-1}\varphi\right) &\leq 0
\end{aligned}$$

as required.

Next suppose that

$$\boldsymbol{w}_{k+1}\left(\boldsymbol{m}_{n-k}(\varphi)+q\mathbf{1}, T^{n-k-1}\varphi\right) \leq \boldsymbol{w}_k\left(\boldsymbol{m}_{n-k+1}(\varphi)+q\mathbf{1}, T^{n-k}\varphi\right)$$

for some $k \in \mathbb{N}$ and $\forall\, n \in \mathbb{Z}$. Then

$$\begin{aligned}
& \boldsymbol{w}_{k+2}\left(\boldsymbol{m}_{n-k-1}(\varphi) + q\mathbf{1}, T^{n-k-2}\varphi\right) \\
= \;& \boldsymbol{f}\left[\boldsymbol{w}_{k+1}\left(\boldsymbol{m}_{n-k-1}(\varphi) + q\mathbf{1}, T^{n-k-2}\varphi\right), s_{n-1}, \tau_{n-1}\right] \\
= \;& \boldsymbol{f}\left[\boldsymbol{w}_{k+1}\left(\boldsymbol{m}_{(n-1)-k}(\varphi) + q\mathbf{1}, T^{(n-1)-k-1}\varphi\right), s_{n-1}, \tau_{n-1}\right] \\
\leq \;& \boldsymbol{f}\left[\boldsymbol{w}_{k}\left(\boldsymbol{m}_{(n-1)-k+1}(\varphi) + q\mathbf{1}, T^{(n-1)-k}\varphi\right), s_{n-1}, \tau_{n-1}\right] \qquad (=) \quad \boldsymbol{f} \text{ monotone} \\
\leq \;& \boldsymbol{f}\left[\boldsymbol{w}_{k}\left(\boldsymbol{m}_{n-k}(\varphi) + q\mathbf{1}, T^{n-k-1}\varphi\right), s_{n-1}, \tau_{n-1}\right] \qquad (=) \\
\leq \;& \boldsymbol{w}_{k+1}\left(\boldsymbol{m}_{n-k}(\varphi) + q\mathbf{1}, T^{n-k-1}\varphi\right)
\end{aligned}$$

which completes our proof.

$\square$

Remark 2.6.1

The result is not unreasonable. The earlier $q\mathbf{1}$ is added into the sequence $\langle \boldsymbol{m}_n \rangle$, the greater is its "absorption by the positive-part function".

Lemma 2.6.2

The limit

$$\boldsymbol{w}_n^q(\varphi) \;=\; \lim_{k\to\infty} \boldsymbol{w}_{k+1}\left(\boldsymbol{m}_{n-k}(\varphi) + q\mathbf{1}, T^{n-k-1}\varphi\right) \tag{2.6.2}$$

exists and is finite.

PROOF

Use the following sequence of inequalities, obtained from Lemma 2.6.1:

$$
\begin{aligned}
0 &\le \boldsymbol{w}_{k+1}\left(\boldsymbol{m}_{n-k}(\varphi) + q\mathbf{1}, T^{n-k-1}\varphi\right) \\
&\le \boldsymbol{w}_1\left(\boldsymbol{m}_n(\varphi) + q\mathbf{1}, T^{n-1}\varphi\right) \qquad (=) \\
&\le \boldsymbol{m}_n(\varphi) + q\mathbf{1} \\
&< \infty
\end{aligned}
$$

□

Proposition 2.6.1

Consider a $G/G/r$ queue with CAP $[\varphi, L(\varphi)]$. Then

(a) $\forall\, q \in \mathbb{R}_{\oplus}$

$$[\psi^q, L(\psi^q)] : \psi^q = \langle(\tau_n, s_n), \boldsymbol{w}_n^q(\varphi)\rangle \tag{2.6.3}$$

is an ergodic strong solution of (2.2.6).

(b) $\boldsymbol{w}_n^q(\varphi)$ is a nondecreasing function of q.

PROOF

(a) First we note that ψ^q is a solution to (2.2.6).

$$\begin{aligned}
\boldsymbol{w}_{n+1}^q(\varphi) &= \lim_{k\to\infty} \boldsymbol{w}_{k+1}\left(\boldsymbol{m}_{n+1-k}(\varphi) + q\mathbf{1}, T^{n-k}\varphi\right) \\
&= \lim_{k\to\infty} \boldsymbol{f}\left[\boldsymbol{w}_k(\boldsymbol{m}_{n+1-k}(\varphi) + q\mathbf{1}, T^{n-k}\varphi), s_n, \tau_n\right] \\
&= \boldsymbol{f}\left[\lim_{k\to\infty} \boldsymbol{w}_k(\boldsymbol{m}_{n+1-k}(\varphi) + q\mathbf{1}, T^{n-k}\varphi), s_n, \tau_n\right] \qquad \text{cont. of } \boldsymbol{f} \\
&= \boldsymbol{f}\left[\lim_{k\to\infty} \boldsymbol{w}_{k+1}(\boldsymbol{m}_{n-k}(\varphi) + q\mathbf{1}, T^{n-k-1}\varphi), s_n, \tau_n\right] \\
&= \boldsymbol{f}\left(\boldsymbol{w}_n^q(\varphi), s_n, \tau_n\right)
\end{aligned}$$

The stationarity is immediate:

$$\begin{aligned}
L[\boldsymbol{w}_n^q(\varphi)] &= L\left[\boldsymbol{w}_n^q(T\varphi)\right] \qquad \varphi \text{ stat.} \\
&= L\left[\lim_{k\to\infty} \boldsymbol{w}_{k+1}\left(\boldsymbol{m}_{n-k}(T\varphi) + q\mathbf{1}, T^{n-k-1}(T\varphi)\right)\right] \\
&= L\left[\lim_{k\to\infty} \boldsymbol{w}_{k+1}\left(\boldsymbol{m}_{n+1-k}(\varphi) + q\mathbf{1}, T^{n-k}\varphi\right)\right] \\
&= L\left[\boldsymbol{w}_{n+1}^q(\varphi)\right]
\end{aligned}$$

Similarly, the ergodicity is a consequence of the ergodicity of φ. Finally, the measurability of $\boldsymbol{w}_n^q$ derives from the measurability of $\boldsymbol{m}_n$.

(b) Let $q > q^*$. Then if we use Lemma 2.3.7(a), we obtain

$$\begin{aligned}
\boldsymbol{w}_n^q(\varphi) &= \lim_{k\to\infty} \boldsymbol{w}_{k+1}\left(\boldsymbol{m}_{n-k}(\varphi) + q\mathbf{1}, T^{n-k-1}\varphi\right) \\
&= \lim_{k\to\infty} \boldsymbol{f}_k\left(\boldsymbol{m}_{n-k}(\varphi) + q\mathbf{1}, s_{n-k}, \ldots, s_{n-1}, \tau_{n-k}, \ldots, \tau_{n-1}\right) \\
&\geq \lim_{k\to\infty} \boldsymbol{f}_k\left(\boldsymbol{m}_{n-k}(\varphi) + q^*\mathbf{1}, s_{n-k}, \ldots, s_{n-1}, \tau_{n-k}, \ldots, \tau_{n-1}\right) \quad (=) \\
&\geq \lim_{k\to\infty} \boldsymbol{w}_{k+1}\left(\boldsymbol{m}_{n-k}(\varphi) + q^*\mathbf{1}, T^{n-k-1}\varphi\right) \quad (=) \\
&\geq \boldsymbol{w}_n^{q^*}(\varphi)
\end{aligned}$$

□

The existence of the limit

$$\boldsymbol{M}_n(\varphi) = \lim_{q\to\infty} \boldsymbol{w}_n^q(\varphi) \tag{2.6.4}$$

is proved by a method analogous to that used in the proof of Lemma 2.6.2, and is an immediate consequence of Proposition 2.6.1.

Proposition 2.6.2

Consider a $G/G/r$ queue with CAP $[\varphi, L(\varphi)]$.

(a) If

$$B = \left\{k \in \overline{\mathbb{N}}_1 : \{\boldsymbol{R}(\boldsymbol{m}_{k-1}(\varphi) + s_{k-1}\boldsymbol{e} - \tau_{k-1}\mathbf{1})\}_1 < 0\right\} \tag{2.6.5}$$

and

$$K = K(\varphi) = \begin{cases} \min\{k : k \in B\} & B \neq \emptyset \\ \infty & B = \emptyset \end{cases} \tag{2.6.6}$$

then

$$P\left[K(\varphi) < \infty\right] = 1 \tag{2.6.7}$$

(b)

$$P\left[w^q_{K,1} = 0\right] = 1 \;\; \forall\, q \in \mathbb{N} \tag{2.6.8}$$

(c) $\boldsymbol{M}_1$ is finite w.p.1.

Hence

$$P\left[\boldsymbol{M}_n(\varphi) \in \mathcal{W}_r\right] = 1 \tag{2.6.9}$$

PROOF

To simplify the proof we may assume w.l.g. that q runs through $\mathbb{N}$ because

(1) $\boldsymbol{w}^q_n$ is an nondecreasing function of q

and

(2) we are concerned with a limit in which we let $q \to \infty$.

(a) If we prove (2.6.7), then we are saying that, with probability one, $\exists\, k > 1$ s.t.

$$\left\{\boldsymbol{R}(\boldsymbol{m}_{k-1}(\varphi) + s_{k-1}\boldsymbol{e} - \tau_{k-1}\mathbf{1})\right\}_1 < 0 \tag{2.6.10}$$

which, of course, means that $m_{k,1} = 0$. In the first place, we recall our assumptions that $\mathrm{E}(\tau_0)$ and $\mathrm{E}(s_0)$ are both finite, $\rho < 1$ and the fact that the minimal solution is stationary ergodic. This immediately establishes a connection with Theorem 1.3.2 where it was concluded that $\exists\, k > 1$ for which $P\left[m_k = 0\right] = 1$. Example 2.5.1 implies that it is wrong to assume in general that $\exists\, k > 1$ s.t. $P\left[\boldsymbol{m}_k = \mathbf{0}\right] = 1$

in the more general case of $r > 1$ servers. It is, however, valid that $P\left[m_{k,1} = 0\right] = 1$. Renewal events are considered in greater detail in Borovkov [13] & [14, Chapter 4]. A related issue is discussed in Wolff [59, *pp.* 495–496].

Before continuing with the main argument of the proof, we note that, if

$$A = \begin{array}{l} \{\varphi : \varphi_n \in \mathbb{R}^2_{\oplus}, \boldsymbol{m}_n(\varphi) \in \mathcal{W}_r, \\ \quad \|\boldsymbol{w}^q_n(\varphi) - \boldsymbol{m}_n(\varphi)\| = \|\boldsymbol{w}^q_1(\varphi) - \boldsymbol{m}_1(\varphi)\|, n \in \mathbb{Z}, q \in \mathbb{N}\} \end{array}$$

then

$$P[\varphi \in A] = 1 \tag{2.6.11}$$

because of the results given in Theorem 2.4.1(e) and Proposition 2.5.1.

(b) Assume, to the contrary, that (2.6.8) is not true. Then

$$\exists\, q \in \mathbb{N} \text{ s.t. } P\left[w^q_{K,1} > 0\right] > 0$$

which implies that

$$P\left[\boldsymbol{w}^q_K \geq \mathbf{0}\right] > 0 \tag{2.6.12}$$

Next we consider (2.6.12) and (2.6.10) in conjunction with Corollary 2.3.3a and Lemma 2.3.6 to obtain — with probability one —

$$\begin{aligned} & \|\boldsymbol{w}^q_K - \boldsymbol{m}_K\| \\ = \; & \left\| \left[\boldsymbol{R}\left(\boldsymbol{w}^q_{K-1} + s_{K-1}\boldsymbol{e} - \tau_{K-1}\mathbf{1}\right)\right]^+ - \left[\boldsymbol{R}\left(\boldsymbol{m}_{K-1} + s_{K-1}\boldsymbol{e} - \tau_{K-1}\mathbf{1}\right)\right]^+ \right\| \\ < \; & \left\| \boldsymbol{R}\left(\boldsymbol{w}^q_{K-1} + s_{K-1}\boldsymbol{e} - \tau_{K-1}\mathbf{1}\right) - \boldsymbol{R}\left(\boldsymbol{m}_{K-1} + s_{K-1}\boldsymbol{e} - \tau_{K-1}\mathbf{1}\right) \right\| \quad (\leq) \\ < \; & \left\| \boldsymbol{w}^q_{K-1} - \boldsymbol{m}_{K-1} \right\| \end{aligned}$$

Thus

$$P\left[\|\boldsymbol{w}_K^q - \boldsymbol{m}_K\| < \|\boldsymbol{w}_{K-1}^q - \boldsymbol{m}_{K-1}\|\right] > 0$$

which contradicts (2.6.11). Hence (2.6.8) is valid.

(c) Consider now the possibility that $\boldsymbol{M}_1 = \lim_{q\to\infty} \boldsymbol{w}_1^q$ is infinite ($\Leftrightarrow w_{1,r}^q = \infty$ (since $w_{1,r}^q = \max_{i\in\mathbb{N}_r} w_{1,i}^q$)). With $\boldsymbol{m}_1$ stationary and finite, the assumption that $\boldsymbol{M}_1$ is unbounded is equivalent to

$$\begin{aligned} \|\boldsymbol{w}_K^q - \boldsymbol{m}_K\| &= \|\boldsymbol{w}_1^q - \boldsymbol{m}_1\| \xrightarrow[\infty]{q} \infty \\ \Longrightarrow \qquad w_{K,r}^q &\xrightarrow[\infty]{q} \infty \end{aligned} \tag{2.6.13}$$

where Proposition 2.6.1 and (2.6.4) have also been used.

We can find

$$N > K \text{ s.t. } \sum_{i=K}^{N-1} \tau_i > m_{K,r} \tag{2.6.14}$$

and from (2.6.13)

$$\exists\, Q \in \mathbb{N} \text{ s.t. } w_{K,r}^Q > \sum_{i=K}^{N-1} (s_i + \tau_i) \tag{2.6.15}$$

We need to show that the conditions for Lemma 2.3.6(b) are met. Write

$\overline{\boldsymbol{w}}_n^Q = \boldsymbol{w}_n^Q + s\boldsymbol{e}$, $\overline{\boldsymbol{m}}_n = \boldsymbol{m}_n + s\boldsymbol{e}$. From (2.6.11), we have

$$\begin{aligned}
& \left\|\boldsymbol{w}_n^Q - \boldsymbol{m}_n\right\| \\
= \quad & \left\|\boldsymbol{w}_1^Q - \boldsymbol{m}_1\right\| \\
= \quad & \left\|\boldsymbol{w}_{n+1}^Q - \boldsymbol{m}_{n+1}\right\| \\
= \quad & \left\|\left[\boldsymbol{R}(\overline{\boldsymbol{w}}_n^Q) - \tau\mathbf{1}\right]^+ - [\boldsymbol{R}(\overline{\boldsymbol{m}}_n) - \tau\mathbf{1}]^+\right\| \\
\leq \quad & \left\|\left[\boldsymbol{R}(\overline{\boldsymbol{w}}_n^Q) - \tau\mathbf{1}\right] - [\boldsymbol{R}(\overline{\boldsymbol{m}}_n) - \tau\mathbf{1}]\right\| \qquad (=) \quad \text{Lemma 2.3.3(b)} \\
\leq \quad & \left\|\boldsymbol{R}(\overline{\boldsymbol{w}}_n^Q) - \boldsymbol{R}(\overline{\boldsymbol{m}}_n)\right\| \\
\leq \quad & \left\|\boldsymbol{w}_n^Q - \boldsymbol{m}_n\right\| \qquad \text{Lemma 2.3.6(a)} \\
\Longrightarrow \quad & \left\|\boldsymbol{R}(\overline{\boldsymbol{w}}_n^Q) - \boldsymbol{R}(\overline{\boldsymbol{m}}_n)\right\| = \left\|\boldsymbol{w}_n^Q - \boldsymbol{m}_n\right\|
\end{aligned}$$

Consequently $\forall\, n$ there exists a permutation $\kappa(n)$ of $\mathbb{N}_r$ such that

$$\boldsymbol{R}(\overline{\boldsymbol{m}}_n) = \left(\overline{\boldsymbol{m}}_{\kappa_n(1)}, \ldots, \overline{\boldsymbol{m}}_{\kappa_n(r)}\right)$$

and

$$\boldsymbol{R}(\overline{\boldsymbol{w}}_n^Q) = \left(\overline{\boldsymbol{w}}_{\kappa_n(1)}^Q, \ldots, \overline{\boldsymbol{w}}_{\kappa_n(r)}^Q\right)$$

but

$$\boldsymbol{m}_{n+1} = [\boldsymbol{R}(\overline{\boldsymbol{m}}_n) - \tau_n\mathbf{1}]^+$$

Thus $$\boldsymbol{m}_{n+1} = \left(\overline{\boldsymbol{m}}_{\kappa_n(1)} - \tau_n, \ldots, \overline{\boldsymbol{m}}_{\kappa_n(r)} - \tau_n\right)^+$$

and similarly $$\boldsymbol{w}_{n+1}^Q = \left(\overline{\boldsymbol{w}}_{\kappa_n(1)}^Q - \tau_n, \ldots, \overline{\boldsymbol{w}}_{\kappa_n(r)}^Q - \tau_n\right)^+ \tag{2.6.16}$$

In particular, suppose $n \in \{K, \ldots, N-1\}$. We claim that, for these values of n, $w_{n,1}^{Q} + s_n < w_{n,r}^{Q}$ in which case Lemma 2.3.6(b) implies that $\kappa_n(r) = r$. The proof of this is carried out in two stages.

First we demonstrate that $\forall\, n \in \overline{\mathbb{N}}_{K-1}$

$$w_{n,1}^{Q} + s_n \leq \sum_{i=K}^{n} s_i \tag{2.6.17}$$

From (2.6.8) we have

$$w_{K,1}^{Q} + s_K = s_K = \sum_{i=K}^{K} s_i$$

Suppose (2.6.17) is true for some $n \in \overline{\mathbb{N}}_{K-1}$. Then

$$\begin{aligned} w_{n+1,1}^{Q} + s_{n+1} &= \min\left[\left(w_{n,1}^{Q} + s_n - \tau_n\right)^{+}, \left(w_{n,2}^{Q} - \tau_n\right)^{+}\right] + s_{n+1} \\ &\leq \left(w_{n,1}^{Q} + s_n - \tau_n\right)^{+} + s_{n+1} \\ &\leq \left(\sum_{i=K}^{n} s_i - \tau_n\right)^{+} + s_{n+1} \\ &\leq \sum_{i=K}^{n+1} s_i \end{aligned}$$

Hence (2.6.17) holds $\forall\, n \in \overline{\mathbb{N}}_{K-1}$, as claimed.

Next we show that $\forall\, n \in \{K, \ldots, N-1\}$

$$w_{n,r}^{Q} > \sum_{i=K}^{N-1} s_i + \sum_{i=n}^{N-1} \tau_i \tag{2.6.18}$$

From (2.6.15) we have

$$w^Q_{K,r} > \sum_{i=K}^{N-1} s_i + \sum_{i=K}^{N-1} \tau_i$$

as required.

If $K = N-1$, then the sub-proof is complete. Otherwise, suppose (2.6.18) is valid for some $n \in \{K, \ldots, N-2\}$. Then

$$\begin{aligned} w^Q_{n,r} &> \sum_{i=K}^{N-1} s_i && (\geq) \\ &> \sum_{i=K}^{n} s_i && (\geq) \\ &> w^Q_{n,1} + s_n && \text{by (2.6.17)} \end{aligned}$$

so (2.2.4) $\Rightarrow$

$$\begin{aligned} w^Q_{n+1,r} &= \left(w^Q_{n,r} - \tau_n\right)^+ \\ &> \left(\sum_{i=K}^{N-1} s_i + \sum_{i=n}^{N-1} \tau_i - \tau_n\right)^+ \quad (=) \quad \text{by the inductive hypothesis} \\ &> \sum_{i=K}^{N-1} s_i + \sum_{i=n+1}^{N-1} \tau_i \end{aligned}$$

Thus $\forall\, n \in \{K, \ldots, N-1\}$

$$\begin{aligned} w^Q_{n,r} &> \sum_{i=K}^{N-1} s_i + \sum_{i=n}^{N-1} \tau_i && (\geq) \\ &> \sum_{i=K}^{n} s_i && (\geq) \\ &> w^Q_{n,1} + s_n && \text{from (2.6.17)} \end{aligned}$$

From (2.6.18) we have $\forall\, n \in \{K, \ldots, N-1\}$

$$w^Q_{n,r} - \tau_n > \sum_{i=K}^{N-1} s_i + \sum_{i=n+1}^{N-1} \tau_i \geq 0 \tag{2.6.19}$$

Next for $n \in \{K, \ldots, N-1\}$

$$\begin{aligned}
\sum_{i=1}^{r} \left(w^Q_{n,i} - m_{n,i}\right)^2 &= \left\|\boldsymbol{w}^Q_n - \boldsymbol{m}_n\right\|^2 \\
&= \left\|\boldsymbol{w}^Q_{n+1} - \boldsymbol{m}_{n+1}\right\|^2 \quad \text{from (2.6.11)} \\
&= \begin{array}{l} \displaystyle\sum_{i \in \mathbb{N}_r \setminus \{l\}} \left[\left(w^Q_{n,i} - \tau_n\right)^+ - \left(m_{n,i} - \tau_n\right)^+\right]^2 \\ + \left[\left(w^Q_{n,l} + s_n - \tau_n\right)^+ - \left(m_{n,l} + s_n - \tau_n\right)^+\right]^2 \end{array}
\end{aligned}$$

where $l = \kappa_n^{-1}(1) < r$, and (2.6.16) is applied in the last step. Thus

$$\sum_{i=1}^{r} \left[\left(x_{n,i} - y_{n,i}\right)^2 - \left(x^+_{n,i} - y^+_{n,i}\right)^2\right] = 0$$

where

$$\begin{aligned}
x_{n,i} &= \begin{cases} w^Q_{n,i} - \tau_n & i \in \mathbb{N}_r \setminus \{l\} \\ w^Q_{n,i} + s_n - \tau_n & i = l \end{cases} \\
y_{n,i} &= \begin{cases} m_{n,i} - \tau_n & i \in \mathbb{N}_r \setminus \{l\} \\ m_{n,i} + s_n - \tau_n & i = l \end{cases}
\end{aligned}$$

but from Lemma 2.3.3(a), $\left(x_{n,i} - y_{n,i}\right)^2 - \left(x^+_{n,i} - y^+_{n,i}\right)^2 \geq 0$ $\forall i,\ \forall n \in \{K, \ldots, N-1\}$ so we must have that $\forall i,\ \forall n \in \{K, \ldots, N-1\}$

$$\left(x_{n,i} - y_{n,i}\right)^2 = \left(x^+_{n,i} - y^+_{n,i}\right)^2 \tag{2.6.20}$$

$$\begin{aligned}
\text{From (2.6.19)} \quad x_{n,r} > 0 \quad &\Rightarrow \quad x_{n,r} = x^+_{n,r} \\
&\Rightarrow \quad y_{n,r} = y^+_{n,r} \qquad \text{by (2.6.20)} \\
\text{or} \qquad m_{n,r} - \tau_n \quad &= \quad (m_{n,r} - \tau_n)^+ \\
&= \quad m_{n+1,r} \qquad \forall\, n \in \{K, \ldots, N-1\}
\end{aligned}$$

which, from Lemma 0.2.2, means that

$$\begin{aligned}
m_{N,r} \quad &= \quad m_{K,r} - \sum_{n=K}^{N-1} \tau_n \\
&< \quad 0 \qquad\qquad \text{by (2.6.14)}
\end{aligned}$$

which contradicts $m_{N,r} \geq 0$.

Thus $\boldsymbol{M}_1$ is finite (w.p.1).

□

The next theorem contains the most important results of this section.

Theorem 2.6.1 (The Maximal Solution)
Consider a $G/G/r$ queue with CAP $[\varphi, L(\varphi)]$ $(\varphi_n = (\tau_n, s_n))$.

(a) In addition to the minimal solution

$$[\psi_{\min}, L(\psi_{\min})] : \psi_{\min} = \langle(\tau_n, s_n), \boldsymbol{m}_n(\varphi)\rangle \tag{2.6.21}$$

of §2.4, there exists the following ergodic strong solution of (2.2.6):

$$[\psi_{\max}, L(\psi_{\max})] : \psi_{\max} = \langle(\tau_n, s_n), \boldsymbol{M}_n(\varphi)\rangle \tag{2.6.22}$$

(b) $[\psi_{\max}, L(\psi_{\max})]$ is the maximal solution of (2.2.6). In other words, for any stationary strong solution $\langle(\tau_n, s_n), \boldsymbol{x}_n\rangle$ of (2.2.6), the following relationship holds:

$$\boldsymbol{M}_n(\varphi) \quad \geq \quad \boldsymbol{x}_n \quad \forall\, n \in \mathbb{Z} \text{ a.s.} \tag{2.6.23}$$

<u>PROOF</u>

(a) The proof is similar to that of Proposition 2.6.1(a).

$$\begin{aligned}
\boldsymbol{M}_{n+1}(\varphi) &= \lim_{q\to\infty} \boldsymbol{w}^q_{n+1}(\varphi) \\
&= \lim_{q\to\infty} \boldsymbol{f}\left(\boldsymbol{w}^q_n(\varphi), s_n, \tau_n\right) \\
&= \boldsymbol{f}\left(\lim_{q\to\infty} \boldsymbol{w}^q_n(\varphi), s_n, \tau_n\right) \quad \text{continuity of } \boldsymbol{f} \\
&= \boldsymbol{f}\left(\boldsymbol{M}_n(\varphi), s_n, \tau_n\right) \qquad \text{a.s.}
\end{aligned}$$

Next we have

$$\begin{aligned}
L[\boldsymbol{M}_n(\varphi)] &= L[\boldsymbol{M}_n(T\varphi)] \qquad \text{stationarity of } \varphi \\
&= L\left[\lim_{q\to\infty} \boldsymbol{w}^q_n(T\varphi)\right] \\
&= L\left[\lim_{q\to\infty} \boldsymbol{w}^q_{n+1}(\varphi)\right] \\
&= L[\boldsymbol{M}_{n+1}(\varphi)]
\end{aligned}$$

In an analogous way, ergodicity of the solution is a result of the ergodicity of φ. This completes (a) (because the aspect of measurability may (in part) be determined from the measurability of $\boldsymbol{w}^q_n$).

(b) From the fact that each of $\langle \boldsymbol{x}_n \rangle$ and $\langle \boldsymbol{m}_n(\varphi) \rangle$ is a stationary a.e. finite sequence, we have

$$P\left[\bigcup_{q=1}^{\infty} \bigcap_{n=-\infty}^{\infty} \{\boldsymbol{x}_n \leq \boldsymbol{m}_n(\varphi) + q\mathbf{1}\}\right] = 1 \tag{2.6.24}$$

which expresses the fact that

$$\forall\, \epsilon > 0 \;\; \exists\, q(\epsilon) \text{ s.t. } P\{\omega : \boldsymbol{x}_n(\omega) \leq \boldsymbol{m}_n(\omega, \varphi) + q(\epsilon)\} > 1 - \epsilon$$

Again, because of the stationarity, we note that

$$\begin{aligned}
& \{\boldsymbol{x}_n \leq \boldsymbol{m}_n(\varphi) + q(\epsilon)\mathbf{1}\} \\
\Rightarrow\;& \{\boldsymbol{x}_{n-k} \leq \boldsymbol{m}_{n-k}(\varphi) + q(\epsilon)\mathbf{1} \;\; \forall\, k \in \mathbb{Z}_{\oplus},\ \forall\, n \in \mathbb{Z}\} \\
\Rightarrow\;& \{\boldsymbol{w}_{k+1}(\boldsymbol{x}_{n-k}, T^{n-k-1}\varphi) \leq \boldsymbol{w}_{k+1}(\boldsymbol{m}_{n-k}(\varphi) + q(\epsilon)\mathbf{1}, T^{n-k-1}\varphi) \\
& \hspace{18em} \forall\, k \in \mathbb{Z}_{\oplus},\ \forall\, n \in \mathbb{Z}\} \\
\Rightarrow\;& \{\boldsymbol{x}_n \leq \boldsymbol{w}_{k+1}(\boldsymbol{m}_{n-k}(\varphi) + q(\epsilon)\mathbf{1}, T^{n-k-1}\varphi) \;\; \forall\, k \in \mathbb{Z}_{\oplus},\ \forall\, n \in \mathbb{Z}\} \\
\Rightarrow\;& \{\boldsymbol{x}_n \leq \boldsymbol{w}_n^{q(\epsilon)}(\varphi) \leq \lim_{q \to \infty} \boldsymbol{w}_n^q(\varphi) \;\; \forall\, n \in \mathbb{Z},\ \forall\, q(\epsilon)\} \\
& \text{i.e., the bound is uniform in } q \\
\Rightarrow\;& \{\boldsymbol{x}_n \leq \boldsymbol{M}_n(\varphi)\}
\end{aligned} \tag{2.6.25}$$

Combination of (2.6.24) and (2.6.25) yields

$$
\begin{aligned}
1 &= P\left[\bigcup_{q=1}^{\infty}\bigcap_{n=-\infty}^{\infty}\{\boldsymbol{x}_n \le \boldsymbol{m}_n(\varphi)+q\mathbf{1}\}\right] \\
&\le P\left[\bigcup_{q=1}^{\infty}\bigcap_{n=-\infty}^{\infty}\bigcap_{k=0}^{\infty}\left\{\boldsymbol{x}_n \le \boldsymbol{w}_{k+1}\left(\boldsymbol{m}_{n-k}(\varphi)+q\mathbf{1}, T^{n-k-1}\varphi\right)\right\}\right] \quad (=) \\
&\le P\left[\bigcup_{q=1}^{\infty}\bigcap_{n=-\infty}^{\infty}\{\boldsymbol{x}_n \le \boldsymbol{w}_n^q(\varphi)\}\right] \quad (=) \\
&\le P\left[\boldsymbol{x}_n \le \boldsymbol{M}_n(\varphi);\ n \in \mathbb{Z}\right]
\end{aligned}
$$

which gives (2.6.23).

□

Remark 2.6.2

From Brandt [15], Loynes [47] and Theorem 2.6.1, we conclude that $[\psi_{\min}, L(\psi_{\min})]$ and $[\psi_{\max}, L(\psi_{\max})]$ give us two ergodic strong solutions. In addition

$$\boldsymbol{m}_n(\varphi) \;\le\; \boldsymbol{x}_n \;\le\; \boldsymbol{M}_n(\varphi) \quad \forall\, n \in \mathbb{Z} \text{ a.s.} \tag{2.6.26}$$

where the notation of the theorem is used. A weak solution $\langle(\tilde{\tau}_n, \tilde{s}_n), \tilde{\boldsymbol{x}}_n\rangle$ is characterised by

$$\boldsymbol{m}_n(\tilde{\varphi}) \;\le\; \tilde{\boldsymbol{x}}_n \;\le\; \boldsymbol{M}_n(\tilde{\varphi}) \quad \forall\, n \in \mathbb{Z} \text{ a.s.} \tag{2.6.27}$$

but $\boldsymbol{m}_n(\tilde{\varphi}) =_d \boldsymbol{m}_n(\varphi)$ and $\boldsymbol{M}_n(\tilde{\varphi}) =_d \boldsymbol{M}_n(\varphi)$ which implies that

$$\boldsymbol{m}_n(\varphi) \;\le_{st}\; \tilde{\boldsymbol{x}}_n \;\le_{st}\; \boldsymbol{M}_n(\varphi) \quad \forall\, n \in \mathbb{Z} \tag{2.6.28}$$

From (2.6.28), we deduce that in the class of weak solutions (of which the collection of strong solutions is a subset) $\boldsymbol{m}_n(\varphi)$ and $\boldsymbol{M}_n(\varphi)$ obviously are minimal and maximal respectively. The coincidence of these solutions is considered next.

Corollary 2.6.1*a* (Unique Solution)

For a $G/G/r$ queue with CAP $[\varphi, L(\varphi)]$, $[\psi_{\min}, L(\psi_{\min})]$ is the unique weak solution iff

$$P[\psi_{\min} = \psi_{\max}] \quad = \quad P[\boldsymbol{m}_n(\varphi) = \boldsymbol{M}_n(\varphi)\ \forall\, n \in \mathbb{Z}] \quad = \quad 1 \qquad (2.6.29)$$

PROOF

Necessity. Suppose $[\psi_{\min}, L(\psi_{\min})]$ is the unique weak solution.
Then, since $\boldsymbol{M}_n(\varphi)$ is a weak solution $\boldsymbol{M}_n(\varphi) =_{st} \boldsymbol{m}_n(\varphi)\ \forall\, n \in \mathbb{Z}$, or $P[\boldsymbol{m}_n(\varphi) = \boldsymbol{M}_n(\varphi)\ \forall n \in \mathbb{Z}] = 1$, and the minimal and maximal solutions coincide.

Sufficiency. Suppose $P[\boldsymbol{m}_n(\varphi) = \boldsymbol{M}_n(\varphi)\ \forall\, n \in \mathbb{Z}] = 1$. (2.6.28) yields $\boldsymbol{m}_n(\varphi) =_{st} \tilde{\boldsymbol{x}}_n =_{st} \boldsymbol{M}_n(\varphi)\ \forall\, n \in \mathbb{Z}$ for every weak solution sequence $\langle \tilde{\boldsymbol{x}}_n \rangle$. Hence the result.

□

Remark 2.6.3 (Uniqueness)

Uniqueness in the class of weak solutions immediately implies uniqueness in the (sub)class of strong solutions (because every strong solution is also weak).

Remark 2.6.4 (Remark 2.5.1(c) revisited)

With more theory, terminology and notation now at our disposal, we are able to provide an example which verifies Remark 2.5.1(c) (on the existence of infinitely many weak solutions when the minimal and maximal solutions do not coincide). Consider

$$\psi_p \;=_d\; \begin{cases} \psi_{\min} & \text{with probability } p \\ \psi_{\max} & \text{with probability } 1-p \end{cases} \qquad \forall\, p \in (0,1)$$

where the "Bernoulli choice" may be incorporated into a random initial condition. (Naturally, therefore, this does not imply infinitely many *strong* solutions.)

2.7 LIMITING BEHAVIOUR AND UNIQUENESS

Proposition 2.7.1

As before, consider a $G/G/r$ queue with CAP $[\varphi, L(\varphi)]$ and the random initial condition $\boldsymbol{w}$ which satisfies

$$\boldsymbol{w} \;\leq\; \boldsymbol{m}_1(\varphi) \quad \text{a.s.} \tag{2.7.1}$$

Then

$$\|\boldsymbol{w}_n(\boldsymbol{w},\varphi) - \boldsymbol{m}_n(\varphi)\|_{\max} \overset{n}{\underset{\infty}{\to}} 0 \;\; \text{a.s.} \tag{2.7.2}$$

so that

$$\boldsymbol{w}_n(\boldsymbol{w},\varphi) \xrightarrow[n\to\infty]{d} \boldsymbol{m}_1(\varphi) \tag{2.7.3}$$

PROOF

$$
\begin{array}{rrcl}
 & \boldsymbol{w} & \leq & \boldsymbol{m}_1(\varphi) \\
\Rightarrow & \boldsymbol{w}_n(\boldsymbol{w},\varphi) & \leq & \boldsymbol{m}_n(\varphi) =_d \boldsymbol{m}_1(\varphi) \\
\Rightarrow & \boldsymbol{w}' = \lim_{n\to\infty} \boldsymbol{w}_n(\boldsymbol{w},\varphi) & \leq_{st} & \boldsymbol{m}_1(\varphi)
\end{array}
$$

but $\boldsymbol{w}'$ is a stationary solution

$$
\begin{array}{rrcll}
\Rightarrow & \boldsymbol{m}_1(\varphi) & \leq_{st} & \boldsymbol{w}' & \text{minimality} \\
\Rightarrow & \boldsymbol{w}' & =_{st} & \boldsymbol{m}_1(\varphi) &
\end{array}
$$

which leads to (2.7.2) and (2.7.3).

□

The next example demonstrates that without (2.7.1) we cannot guarantee (2.7.2) (or (2.7.3)).

Example 2.7.1

We again consider the $G/G/2$ queue of Example 2.5.1 but with the additional assumption that $\sigma_1 = \frac{3}{2}\tau$. Initially we also restrict ourselves to $n \in \mathbb{N}$ in this example in which the minimal waiting-time sequence is then given by

$$
\boldsymbol{m}_n(\omega) \quad = \quad \widetilde{\boldsymbol{w}}_n^{\tau}(\omega) \quad = \quad \begin{cases} (0,\tau) & n = 2k \\ (0,\frac{1}{2}\tau) & n = 2k+1 \end{cases}
$$

$$
\Rightarrow \ \boldsymbol{m}_n(\varphi) \ < \ \boldsymbol{w} \ = \ (0,\sigma_1) \ = \ (0,\tfrac{3}{2}\tau) \ \ \forall\, n \in \mathbb{Z}
$$

where $\boldsymbol{w}$ is the fixed initial state we will be considering.

Then

$$\boldsymbol{w}_n(\boldsymbol{w}, T\omega) = \begin{cases} (\frac{1}{2}\tau, \frac{1}{2}\tau) & n = 2k \\ (0, \frac{3}{2}\tau) & n = 2k+1 \end{cases} \qquad \text{(since } \boldsymbol{w}_n(\boldsymbol{w}, T\omega) = \widetilde{\boldsymbol{w}}_n^{\sigma_1}(T\omega))$$

$$\boldsymbol{w}_n(\boldsymbol{w}, \omega) = \begin{cases} (0, \frac{3}{2}\tau) & n = 1 \\ (\frac{1}{2}\tau, \tau) & n = 2 \\ (0, \tau) & n = 3 \\ \boldsymbol{m}_n(\omega) & n \in \overline{\mathbb{N}}_3 \end{cases}$$

For sufficiently large n certainly[2]

$$\begin{aligned} \boldsymbol{w}_n(\boldsymbol{w}, \omega) - \boldsymbol{m}_n(\omega) &= (0,0) \\ \Rightarrow \qquad \|\boldsymbol{w}_n(\boldsymbol{w}, \omega) - \boldsymbol{m}_n(\omega)\|_{\max} &= 0 \end{aligned}$$

and

$$\begin{aligned} \boldsymbol{w}_n(\boldsymbol{w}, T\omega) - \boldsymbol{m}_n(T\omega) &= \begin{cases} (\frac{1}{2}\tau, 0) & n = 2k \\ (0, \frac{1}{2}\tau) & n = 2k+1 \end{cases} \\ \Rightarrow \quad \|\boldsymbol{w}_n(\boldsymbol{w}, T\omega) - \boldsymbol{m}_n(T\omega)\|_{\max} &= \tfrac{1}{2}\tau \end{aligned}$$

but $P[\varphi = \omega] = \frac{1}{2} = P[\varphi = T\omega]$. Hence

$$P\left[\|\boldsymbol{w}_n(\boldsymbol{w}, \varphi) - \boldsymbol{m}_n(\varphi)\|_{\max} = 0\right] = \tfrac{1}{2} = P\left[\|\boldsymbol{w}_n(\boldsymbol{w}, \varphi) - \boldsymbol{m}_n(\varphi)\|_{\max} = \tfrac{1}{2}\tau\right]$$

so obviously (2.7.2) cannot hold.

From Example 2.5.1 we know that there is more than one weak solution of (2.2.6) for the $G/G/2$ queue considered in that example. This suggests a

[2]With an initial condition "imposed at $-\infty$" the result would actually hold $\forall\, n \in \mathbb{Z}$.

potential connection between a failure to satisfy (2.7.2) and the existence of more than one weak solution. The next theorem confirms this suspicion.

Theorem 2.7.1

Consider a $G/G/r$ queue with CAP $[\varphi, L(\varphi)]$. Then $[\psi_{\min}, L(\psi_{\min})]$ is the unique weak solution for $[\varphi, L(\varphi)]$ iff any one of the following equivalent conditions holds for an arbitrary initial state $\boldsymbol{w}$.

$$\|\boldsymbol{w}_n(\boldsymbol{w},\varphi) - \boldsymbol{m}_n(\varphi)\| \underset{\infty}{\overset{n}{\to}} 0 \text{ a.s.} \tag{2.7.4}$$

$$\langle \boldsymbol{w}_{n+k}(\boldsymbol{w},\varphi), n \in \mathbb{N}\rangle \xrightarrow[k\to\infty]{d} \langle \boldsymbol{m}_n(\varphi), n \in \mathbb{N}\rangle \tag{2.7.5}$$

$$\boldsymbol{w}_n(\boldsymbol{w},\varphi) \overset{d}{\to} \boldsymbol{m}_1(\varphi) \tag{2.7.6}$$

<u>PROOF</u>

(i) <u>*Uniqueness of* $[\psi_{\min}, L(\psi_{\min})] \Rightarrow$ (2.7.4)</u>

Let $q > 0$,

$$\begin{aligned}
\boldsymbol{d}^q(k,n) &= \boldsymbol{d}^q(k,n,\varphi) \\
&= \boldsymbol{f}_k(\boldsymbol{m}_n(\varphi) + q\mathbf{1}, s_n, \ldots, s_{n+k-1}, \tau_n, \ldots, \tau_{n+k-1}) \\
&\quad -\boldsymbol{m}_{n+k}(\varphi) \\
&= \boldsymbol{w}_{k+1}(\boldsymbol{m}_n(\varphi) + q\mathbf{1}, T^{n-1}\varphi) \\
&\quad -\boldsymbol{w}_{k+1}(\boldsymbol{m}_n(\varphi), T^{n-1}\varphi) \\
&\leq q\mathbf{1} \\
\Longrightarrow \quad \|\boldsymbol{d}^q(k,n)\|_{\max} &\leq q
\end{aligned} \tag{2.7.7}$$

and

$$D_n^q = D_n^q(\varphi) = \{k \in \mathbb{N}: \|\boldsymbol{d}^q(k,n,\varphi)\|_{\max} < q\} \tag{2.7.8}$$

Define

$$\delta_n^q = \delta_n^q(\varphi) = \begin{cases} \min\{k: k \in D_n^q\} & D_n^q \neq \emptyset \\ \infty & D_n^q = \emptyset \end{cases} \tag{2.7.9}$$

$$\Delta_n^q = \Delta_n^q(\varphi) = \begin{cases} q - \|\boldsymbol{d}^q(\delta_n^q, n)\|_{\max} & \delta_n^q < \infty \\ 0 & \delta_n^q = \infty \end{cases} \tag{2.7.10}$$

Next consider the sequence for fixed q

$$\Lambda^q = \langle(\tau_n, s_n, \delta_n^q, \Delta_n^q)\rangle \tag{2.7.11}$$

For a reasoning similar to that used in Proposition 2.6.1(a) and Theorem 2.6.1(a), Λ^q is stationary ergodic (because φ is).

If we now assume that

$$\exists\, q_1 > 0 \text{ s.t. } P[\delta_1^{q_1} = \infty] = 1 \tag{2.7.12}$$

then the stationarity implies $P[\delta_n^{q_1} = \infty\ \forall\, n \in \mathbb{Z}] = 1$.

In accordance with (2.7.7), (2.7.8), (2.7.9) and stationarity, we have

$$
\begin{aligned}
& P[\|\boldsymbol{d}^{q_1}(k,n,\varphi)\|_{\max} = q_1 \;\; \forall\, n \in \mathbb{Z},\ \forall\, k \in \mathbb{N}] && = 1 \\
\Rightarrow\quad & P[\|\boldsymbol{d}^{q_1}(k,n,T^{-k}\varphi)\|_{\max} = q_1 \;\; \forall\, n \in \mathbb{Z},\ \forall\, k \in \mathbb{N}] && = 1 \\
\Rightarrow\quad & P[\|\boldsymbol{w}_{k+1}(\boldsymbol{m}_{n-k}(\varphi) + q_1\mathbf{1}, T^{n-k-1}\varphi) - \boldsymbol{m}_n(\varphi)\|_{\max} \\
& \qquad = q_1 \;\; \forall\, n \in \mathbb{Z},\ \forall\, k \in \mathbb{N}] && = 1 \\
\Rightarrow\quad & P[\|\lim_{k\to\infty} \boldsymbol{w}_{k+1}(\boldsymbol{m}_{n-k}(\varphi) + q_1\mathbf{1}, T^{n-k-1}\varphi) - \boldsymbol{m}_n(\varphi)\|_{\max} \\
& \qquad = q_1 \;\; \forall\, n \in \mathbb{Z}] && = 1 \\
\Rightarrow\quad & P[\|\boldsymbol{w}_n^{q_1}(\varphi) - \boldsymbol{m}_n(\varphi)\|_{\max} = q_1 > 0 \;\; \forall\, n \in \mathbb{Z}] && = 1
\end{aligned}
\tag{2.7.13}
$$

but, of course, $\boldsymbol{m}_n(\varphi) \leq \boldsymbol{w}_n^{q_1}(\varphi)$ so (2.7.13) would then imply that $L(\psi_{\min}) \neq L(\psi^q)$ which contradicts the assumption on the uniqueness of $[\psi_{\min}, L(\psi_{\min})]$.

Thus (2.7.12) is invalid so $\forall\, q > 0$

$$P[\delta_1^q(\varphi) < \infty] > 0 \tag{2.7.14}$$

hence

$$
\begin{aligned}
& P[\exists\, k \in \mathbb{N} \text{ s.t. } \|\boldsymbol{d}^q(k,1)\|_{\max} < q] && > 0 \\
\Rightarrow\quad & P[\exists\, k \in \mathbb{N} \text{ s.t. } q - \|\boldsymbol{d}^q(k,1)\|_{\max} > 0] && > 0 \\
\Rightarrow\quad & P[\Delta_1^q > 0] && > 0
\end{aligned}
\tag{2.7.15}
$$

From (2.7.14) and (2.7.15) we thus have that

$$\exists\, \epsilon(q) > 0 \text{ s.t. } P\left[\{\delta_1^q(\varphi) < \infty\} \cap \{\Delta_1^q(\varphi) \geq \epsilon(q)\}\right] > 0 \tag{2.7.16}$$

Next define

$$\begin{aligned} B^q(\varphi) &= \{n \in \mathbb{N}: \{\delta_n^q(\varphi) < \infty\} \cap \{\Delta_n^q(\varphi) \geq \epsilon(q)\}\} \\ \mathcal{B}^q &= \left\{\varphi \in (\mathbb{R}_\oplus^2)^{\mathbb{Z}}: |B^q(\varphi)| = \infty\right\} \\ \text{and} \quad \mathcal{B} &= \bigcap_{k=1}^{\infty} \mathcal{B}^{k^{-1}} \end{aligned} \tag{2.7.17}$$

The stationarity and ergodicity of Λ^q, in conjunction with (2.7.16), (because the probability of an invariant event is 0 or 1 under these conditions) gives

$$\begin{aligned} P[\varphi \in \mathcal{B}^q] &= 1 \quad \forall\, q \\ \Rightarrow \quad P[\varphi \in \mathcal{B}^{k^{-1}}] &= 1 \quad \forall\, k \in \mathbb{N} \\ \Rightarrow \quad P[\varphi \in \mathcal{B}] &= 1 \end{aligned} \tag{2.7.18}$$

Consider $\varphi_0 \in \mathcal{B}$ and an arbitrary initial condition $\overline{\boldsymbol{w}} \in \mathcal{W}_r$ satisfying

$$\boldsymbol{m}_1(\varphi_0) \leq \overline{\boldsymbol{w}} \tag{2.7.19}$$

We prove that

$$\begin{aligned} z_n &= z_n(\overline{\boldsymbol{w}}, \varphi_0) = \|\boldsymbol{v}_n\|_{\max} \xrightarrow[\infty]{n} 0 \\ \text{where} \quad \boldsymbol{v}_n &= \boldsymbol{v}_n(\overline{\boldsymbol{w}}, \varphi_0) = \boldsymbol{w}_n(\overline{\boldsymbol{w}}, \varphi_0) - \boldsymbol{m}_n(\varphi_0) \end{aligned} \tag{2.7.20}$$

by a *reductio ad absurdum* argument.

Now $\langle z_n, n \in \mathbb{N}\rangle$ is a nonnegative nonincreasing sequence (Lemma 2.3.4) so that if (2.7.20) is assumed to be false, we must have $z_n > 0 \; \forall n \in \mathbb{N}$. Consequently

$$\exists\, k_0 \in \mathbb{N} \text{ s.t. } q_0 = k_0^{-1} < z_1 \leq z_n \quad \forall\, n \in \mathbb{N} \tag{2.7.21}$$

Since $\varphi_0 \in \mathcal{B}$, $\varphi_0 \in \mathcal{B}^{k^{-1}} \; \forall\, k \in \mathbb{N} \Rightarrow \varphi_0 \in \mathcal{B}^{q_0}$. Thus there exists a subsequence $\langle i_l\rangle$ of the natural numbers such that $\delta_{i_l}^{q_0}(\varphi_0) < \infty \;\forall\, i_l \; (l \in \mathbb{N})$ and $\Delta_{i_l}^{q_0}(\varphi_0) \geq \epsilon_{q_0}$. From this denumerable subsequence a sub-subsequence $\langle j'(i_l)\rangle$ (that we will abbreviate as $\langle j_l\rangle$) satisfying

$$j_l + \delta_{j_l}^{q_0}(\varphi_0) < j_{l+1} \quad \text{and} \quad \Delta_{j_l}^{q_0}(\varphi_0) \geq \epsilon_{q_0} \tag{2.7.22}$$

may be extracted.

Fix $l \in \mathbb{N}$ and, to simplify notation, write

$$\boldsymbol{w}_{\delta+1}(\cdot, T^{j_l-1}\varphi_0) \;=\; \boldsymbol{f}_\delta(\cdot, s_{j_l}, \ldots, s_{j_l+\delta-1}, \tau_{j_l}, \ldots, \tau_{j_l+\delta-1})$$

where $\delta = \delta_{j_l}^{q_0}$.

If we define

$$\begin{aligned}\boldsymbol{w}_{j_l}^*(\overline{\boldsymbol{w}}, \varphi_0) &= \boldsymbol{m}_{j_l}(\varphi_0) + q_0 \frac{\boldsymbol{v}_{j_l}}{\|\boldsymbol{v}_{j_l}\|_{\max}} \\ &\in [\boldsymbol{m}_{j_l}(\varphi_0), \boldsymbol{m}_{j_l}(\varphi_0) + q_0\mathbf{1}]\end{aligned} \tag{2.7.23}$$

then

$$
\begin{aligned}
\boldsymbol{w}_{j_l}(\overline{\boldsymbol{w}},\varphi_0)-\boldsymbol{w}^*_{j_l}(\overline{\boldsymbol{w}},\varphi_0) &= [\boldsymbol{v}_{j_l}+\boldsymbol{m}_{j_l}(\varphi_0)]-\left[\boldsymbol{m}_{j_l}(\varphi_0)+q_0\frac{\boldsymbol{v}_{j_l}}{\|\boldsymbol{v}_{j_l}\|_{\max}}\right] \\
&= \boldsymbol{v}_{j_l}-q_0\frac{\boldsymbol{v}_{j_l}}{\|\boldsymbol{v}_{j_l}\|_{\max}} \\
&= \frac{\boldsymbol{v}_{j_l}}{\|\boldsymbol{v}_{j_l}\|_{\max}}[z_{j_l}-q_0] \\
&\geq \mathbf{0}
\end{aligned}
\tag{2.7.24}
$$

because of (2.7.21). Using Lemma 2.3.7 several times

$$
\begin{aligned}
& z_{j_{(l+1)}} \\
= \;& \|\boldsymbol{v}_{j_{(l+1)}}\|_{\max} \\
= \;& \left\|\boldsymbol{w}_{j_{(l+1)}}(\overline{\boldsymbol{w}},\varphi_0)-\boldsymbol{m}_{j_{(l+1)}}(\varphi_0)]\right\|_{\max} \\
\leq \;& \|\boldsymbol{w}_{j_l+\delta}(\overline{\boldsymbol{w}},\varphi_0)-\boldsymbol{m}_{j_l+\delta}(\varphi_0)]\|_{\max} && j_{(l+1)}>\delta+j_l \text{ by } (2.7.22) \\
\leq \;& \left\|\boldsymbol{w}_{j_l+\delta}(\overline{\boldsymbol{w}},\varphi_0)-\boldsymbol{w}^*_{j_l+\delta}(\overline{\boldsymbol{w}},\varphi_0)]\right\|_{\max} \\
& +\left\|\boldsymbol{w}^*_{j_l+\delta}(\overline{\boldsymbol{w}},\varphi_0)-\boldsymbol{m}_{j_l+\delta}(\varphi_0)]\right\|_{\max} \\
\leq \;& \left\|\boldsymbol{w}_{j_l}(\overline{\boldsymbol{w}},\varphi_0)-\boldsymbol{w}^*_{j_l}(\overline{\boldsymbol{w}},\varphi_0)\right\|_{\max} && \text{using}(2.7.23) \\
& +\|\boldsymbol{w}_{\delta+1}\left(\boldsymbol{m}_{j_l}(\varphi_0)+q_0\mathbf{1},T^{j_l-1}\varphi_0\right)-\boldsymbol{w}_{\delta+1}\left(\boldsymbol{m}_{j_l}(\varphi_0),T^{j_l-1}\varphi_0\right)\|_{\max} \\
\leq \;& (z_{j_l}-q_0)+\left(q_0-\Delta^{q_0}_{j_l}\right) && \text{by } (2.7.24) \text{ and } (2.7.10) \text{ resp.} \\
\leq \;& z_{j_l}-\epsilon_{q_0} && \Delta^{q_0}_{j_l}\geq\epsilon_{q_0} \text{ by } (2.7.22)
\end{aligned}
$$

Thus $z_{j_l}\leq z_{j_1}-\epsilon_{q_0}(l-1)\ \forall\, l\in\mathbb{N}$ by suitable adaptation of Lemma

0.2.2. For sufficiently large l, this leads to a contradiction (because of Lemma 2.3.7(b)) which implies that the assumption $\lim_{n\to\infty} z_n \neq 0$ is false. Hence (2.7.20) is valid.

In Proposition 2.7.1 (equation (2.7.2)) we deduced that

$$\boldsymbol{w} \;\leq\; \boldsymbol{m}_1(\varphi) \quad \text{a.s.} \quad \Rightarrow \quad \|\boldsymbol{w}_n(\boldsymbol{w},\varphi) - \boldsymbol{m}_n(\varphi)\|_{\max} \overset{n}{\underset{\infty}{\to}} 0 \quad \text{a.s.}$$

whereas we have just shown ((2.7.20)) that

$$\boldsymbol{m}_1(\varphi) \;\leq\; \boldsymbol{w} \quad \text{a.s.} \quad \Rightarrow \quad \|\boldsymbol{w}_n(\boldsymbol{w},\varphi) - \boldsymbol{m}_n(\varphi)\|_{\max} \overset{n}{\underset{\infty}{\to}} 0 \quad \text{a.s.}$$

It remains to investigate more general instances of the initial state $\boldsymbol{w} \in \mathcal{W}_r$.

To this end, define two $\mathcal{W}_r$-vectors $\boldsymbol{w}^{(-)}$ and $\boldsymbol{w}^{(+)}$ where

$$\left.\begin{array}{lll} w_i^{(-)} & = \min[m_{1,i}(\varphi), w_i] & \leq \left\{\begin{array}{l} w_i \\ m_{1,i}(\varphi) \end{array}\right. \\ \\ w_i^{(+)} & = \max[m_{1,i}(\varphi), w_i] & \geq \left\{\begin{array}{l} w_i \\ m_{1,i}(\varphi) \end{array}\right. \end{array}\right. \quad \forall\, i \in \mathbb{N}_r \tag{2.7.25}$$

which implies that

$$\boldsymbol{w}_n(\boldsymbol{w}^{(-)},\varphi) \;\leq\; \boldsymbol{w}_n(\boldsymbol{w},\varphi) \;\leq\; \boldsymbol{w}_n(\boldsymbol{w}^{(+)},\varphi) \tag{2.7.26}$$

(2.7.25) in conjunction with (2.7.2) gives us

$$\|\boldsymbol{w}_n(\boldsymbol{w}^{(-)},\varphi) - \boldsymbol{m}_n(\varphi)\|_{\max} \overset{n}{\underset{\infty}{\to}} 0 \quad \text{a.s.} \tag{2.7.27}$$

and (2.7.25) combined with (2.7.20) yields

$$\|\boldsymbol{w}_n(\boldsymbol{w}^{(+)},\varphi) - \boldsymbol{m}_n(\varphi)\|_{\max} \overset{n}{\underset{\infty}{\to}} 0 \quad \text{a.s.} \tag{2.7.28}$$

Consequently

$$\begin{aligned}
0 &\le \|\boldsymbol{w}_n(\boldsymbol{w},\varphi) - \boldsymbol{m}_n(\varphi)\|_{\max} \\
&\le \|\boldsymbol{w}_n(\boldsymbol{w},\varphi) - \boldsymbol{w}_n(\boldsymbol{w}^{(+)},\varphi)\|_{\max} + \|\boldsymbol{w}_n(\boldsymbol{w}^{(+)},\varphi) - \boldsymbol{m}_n(\varphi)\|_{\max} \\
&\le \|\boldsymbol{w}_n(\boldsymbol{w}^{(+)},\varphi) - \boldsymbol{w}_n(\boldsymbol{w}^{(-)},\varphi)\|_{\max} + \|\boldsymbol{w}_n(\boldsymbol{w}^{(+)},\varphi) - \boldsymbol{m}_n(\varphi)\|_{\max} \\
&\le 2\|\boldsymbol{w}_n(\boldsymbol{w}^{(+)},\varphi) - \boldsymbol{m}_n(\varphi)\|_{\max} + \|\boldsymbol{w}_n(\boldsymbol{w}^{(-)},\varphi) - \boldsymbol{m}_n(\varphi)\|_{\max} \overset{n}{\underset{\infty}{\to}} 0 \text{ a.s.} \\
&\Rightarrow \|\boldsymbol{w}_n(\boldsymbol{w},\varphi) - \boldsymbol{m}_n(\varphi)\| \overset{n}{\underset{\infty}{\to}} 0 \text{ a.s. by Remark 2.3.2}
\end{aligned}$$

(ii) (2.7.4) $\Rightarrow$ (2.7.5)

Let $X = \{\boldsymbol{x} = \langle \boldsymbol{x}_n, n \in \mathbb{N}\rangle\,; \boldsymbol{x}_n \in \mathbb{R}^m\}$. If $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)} \in X$, then define

$$d(\boldsymbol{x}^{(1)},\boldsymbol{x}^{(2)}) = \sum_{n=1}^{\infty} d^*(\boldsymbol{x}_n^{(1)},\boldsymbol{x}_n^{(2)})2^{-n} = \sum_{n=1}^{\infty} \frac{\|\boldsymbol{x}_n^{(1)} - \boldsymbol{x}_n^{(2)}\|}{1 + \|\boldsymbol{x}_n^{(1)} - \boldsymbol{x}_n^{(2)}\|} 2^{-n}$$

(Compare Billingsley [11, *pp.* 218–219] and Franken *et al.* [31, *p.* 68].)

d^* is a metric on $\mathbb{R}^m$ (compare Apostol [1, *p.* 67, 3.29]) so that $2^{-n}d^*$ ($\forall\, n \in \mathbb{N}$) is another metric on that space.[3] More significantly, (X, d) is a complete separable metric space.

Write

$$\begin{aligned}
\boldsymbol{\xi}_k &= \langle(\tau_{n+k}, s_{n+k}, \boldsymbol{m}_{n+k}(\varphi)), n \in \mathbb{N}\rangle \\
\boldsymbol{\zeta}_k &= \langle(\tau_{n+k}, s_{n+k}, \boldsymbol{w}_{n+k}(\boldsymbol{w},\varphi)), n \in \mathbb{N}\rangle
\end{aligned}$$

It is easy to prove that

$$\|\boldsymbol{x}_{n+k}^{(1)} - \boldsymbol{x}_{n+k}^{(2)}\| \le \|\boldsymbol{x}_n^{(1)} - \boldsymbol{x}_n^{(2)}\| \Rightarrow d^*(\boldsymbol{x}_{n+k}^{(1)},\boldsymbol{x}_{n+k}^{(2)}) \le d^*(\boldsymbol{x}_n^{(1)},\boldsymbol{x}_n^{(2)}) \tag{2.7.29}$$

[3] 2^{-n} ensures convergence of the series.

(by considering the difference $d^*(\boldsymbol{x}_n^{(1)}, \boldsymbol{x}_n^{(2)}) - d^*(\boldsymbol{x}_{n+k}^{(1)}, \boldsymbol{x}_{n+k}^{(2)})$).

Now suppose $m = r + 2$. Then

$$\begin{aligned}
& P[d(\boldsymbol{\xi}_k, \boldsymbol{\zeta}_k) \geq \epsilon] \\
= \; & P\left[\sum_{n=1}^{\infty} d^*(\boldsymbol{w}_{n+k}(\boldsymbol{w}, \varphi), \boldsymbol{m}_{n+k}(\varphi))2^{-n} \geq \epsilon\right] \\
\leq \; & P\left[\sum_{n=1}^{\infty} d^*(\boldsymbol{w}_{k+1}(\boldsymbol{w}, \varphi), \boldsymbol{m}_{k+1}(\varphi))2^{-n} \geq \epsilon\right] \qquad (=) \quad \text{by Lemma 2.3.6 and (2.7.29)} \\
\leq \; & P\left[d^*(\boldsymbol{w}_{k+1}(\boldsymbol{w}, \varphi), \boldsymbol{m}_{k+1}(\varphi))\sum_{n=1}^{\infty} 2^{-n} \geq \epsilon\right] \qquad (=) \\
\leq \; & P[\|\boldsymbol{w}_{k+1}(\boldsymbol{w}, \varphi) - \boldsymbol{m}_{k+1}(\varphi)\| \geq \epsilon(1 + \|\boldsymbol{w}_{k+1}(\boldsymbol{w}, \varphi) - \boldsymbol{m}_{k+1}(\varphi)\|)\,] \qquad (=) \\
\leq \; & P\left[\|\boldsymbol{w}_{k+1}(\boldsymbol{w}, \varphi) - \boldsymbol{m}_{k+1}(\varphi)\| \geq \frac{\epsilon}{1-\epsilon}\right] \overset{k}{\underset{\infty}{\to}} 0
\end{aligned} \tag{2.7.30}$$

Stationarity gives $\boldsymbol{\xi}_k =_d \boldsymbol{\xi}_0$ so obviously $\boldsymbol{\xi}_k \xrightarrow[k\to\infty]{d} \boldsymbol{\xi}_0$ and (2.7.30) implies that $d(\boldsymbol{\xi}_k, \boldsymbol{\zeta}_k) \xrightarrow[k\to\infty]{p} 0$.

Consequently (Billingsley [11, *p.* 25, Theorem 4.1]) $\boldsymbol{\zeta}_k \xrightarrow[k\to\infty]{d} \boldsymbol{\xi}_0$.

In other words

$$\langle(\tau_{n+k}, s_{n+k}, \boldsymbol{w}_{n+k}(\boldsymbol{w}, \varphi)), n \in \mathbb{N}\rangle \xrightarrow[k\to\infty]{d} \langle(\tau_n, s_n, \boldsymbol{m}_n(\varphi)), n \in \mathbb{N}\rangle$$

such that

$$\langle\boldsymbol{w}_{n+k}(\boldsymbol{w}, \varphi), n \in \mathbb{N}\rangle \xrightarrow[k\to\infty]{d} \langle\boldsymbol{m}_n(\varphi), n \in \mathbb{N}\rangle$$

(iii) <u>(2.7.5) $\Rightarrow$ (2.7.6)</u>

From (2.7.5) we may deduce that

$$\boldsymbol{w}_{k+1}(\boldsymbol{w},\varphi) \xrightarrow[k\to\infty]{d} \boldsymbol{m}_1(\varphi)$$

$$\Rightarrow \quad \boldsymbol{w}_n(\boldsymbol{w},\varphi) \xrightarrow[n\to\infty]{d} \boldsymbol{m}_1(\varphi)$$

(by "letting $k = n - 1$").

(iv) <u>(2.7.6) $\Rightarrow$ *Uniqueness of* $[\psi_{\min}, L(\psi_{\min})]$</u>

(2.7.6) gives

$$\boldsymbol{w}_n(\boldsymbol{M}_1(\varphi),\varphi) \xrightarrow[n\to\infty]{d} \boldsymbol{m}_1(\varphi)$$

but $P[\boldsymbol{w}_n(\boldsymbol{M}_1(\varphi),\varphi) = \boldsymbol{M}_n(\varphi)] = 1$ which implies that

$$\begin{aligned}
L[\boldsymbol{M}_1(\varphi)] &= \lim_{n\to\infty} L[\boldsymbol{M}_1(\varphi)] \\
&= \lim_{n\to\infty} L[\boldsymbol{M}_n(\varphi)] && \text{by stationarity} \\
&= \lim_{n\to\infty} L[\boldsymbol{w}_n(\boldsymbol{M}_1(\varphi),\varphi)] \\
&= L[\boldsymbol{m}_1(\varphi)]
\end{aligned}$$

hence the uniqueness.

$\square$

2.8 CONCLUSION

The two chapters of this part of the thesis are important both in their own right, and collectively as a strong basis for the forthcoming work on networks.

Before we commence with that work, we make a brief mention of a result from the paper by Brandt & Lisek [17] on the continuity of $G/GI/r$ queues: Given $G/GI/r$ queues with CAPs $\varphi = \langle(\tau_n, s_n)\rangle$ and $\varphi^k = \left\langle\left(\tau_n^k, s_n^k\right)\right\rangle$, if φ^k convergences in distribution to φ, and $\mathrm{E}(\tau_0^k)$ and $\mathrm{E}(s_0^k)$ converge to $\mathrm{E}(\tau_0)$ and $\mathrm{E}(s_0)$ respectively, then the unique stationary waiting-time distribution $L\left(\boldsymbol{m}_n(\varphi^k)\right) \xrightarrow[\infty]{k} L(\boldsymbol{m}_n(\varphi))$. It is worth noting that in this paper, as well as in Brandt [16], consideration is also given to cyclic queueing disciplines.

PART II
ON A CLASS OF QUEUEING AND COMMUNICATION NETWORKS

A theoretical foundation for a broad class of queues was established in Part I of the thesis. The tendency in this second part will be to supplement the theory with practical applications and a simulation study. In keeping with the title (of this part and, more generally, the thesis), several examples, as well as an explanation of the terminology, from communications and computing are included.

To a certain extent, the notation and assumptions of the previous part of the thesis will apply but inevitably there is a need to make some changes and additions, which will be specified in the individual chapters. Most of the references cited in the introduction to Part I include at least one chapter on the theory and application of networks. Kleinrock [43] is particularly relevant. In addition there is the detailed review paper, Kelbert & Sukhov [38].

While some background knowledge of networks is assumed, it is appropriate to explain a few terms, many of which have their origin in queueing theory but are applied in this second part on networks.

Franken *et al.* [31, *pp.* 60–61] differs somewhat from Gross & Harris [32, *p.* 360] and Kleinrock [42, *p.* 208] & [43, *p.* 113] in defining *work conservation*. We adopt the definition of the latter group of authors. A queueing system (or the discipline which governs its operation) is said to be *conservative* or *work conserving* if no work (service requirement) is created or destroyed within the system. Destruction of work would occur if a customer were to leave the system before completing his service (*balking* and *reneging* are extreme examples). Creation of work might correspond to a situation in which *a server stands idle in the face of a nonempty queue.*

We next examine the issue of *priority queueing systems.* Not only will we borrow from the language of such systems, but they will also provide us with another example of "non-conservation". While a *nonpreemptive* service rule prevents interruption of the service of a customer by another of higher priority, a *preemptive* rule stipulates that the customer with a high priority (relatively-speaking) is allowed to enter service immediately if a customer with a lower priority is in service when the higher-priority customer enters the system. The preempted customer is forced to return to the queue. Two major types of preemptive priority schemes may be distinguished:

(a) *Preemptive resume priority systems* are characterised by the fact that, when the preempted customer returns to the service facility, he resumes service from the point of preemption.

(b) *Preemptive repeat priority systems* require that the preempted customer's service starts afresh.

Clearly (b) is a case in which additional work is created (i.e., the conservation principle is not satisfied).

A third idea worthy of mention is that of the *virtual waiting time* w_t which, in its more general form (usually for $FCFS$ queues) is defined as the time a (fictitious) customer would have to wait were he to arrive at time t. In a single-server queue, w_t may be defined equivalently as the length of the interval from t until the server becomes free if no further customers are allowed to join the queue, hence it is also the *work backlog* or remaining work.[1] In multiserver queues, however, remaining work generally differs from the virtual wait because only one server needs to be idle for a customer to start his service.

The general objectives and approach with regard to proofs, examples and remarks are essentially the same as in Part I (although in Part II, (a) remarks are less frequent and (b) not all proofs are given in full detail). While a fair proportion of the work involves a review of articles and extracts from books, an attempt has been made at (consolidation and) amalgamation of these articles with one another and with original or new work. Diagrams, tables and graphs reinforce the ideas and explanations, as do the results of a simulation program written by the author as an original contribution to the thesis.

A brief insight into the interdependence between telecommunications and networks is provided in Chapter 3 which is based on many different sources, including Baccelli & Makowski [5]. Chapter 4 focuses on a specific type of network and the disciplines which may govern its operation. This chapter, which includes the simulation results and analysis, forms a strong foundation for Chapters 5, 6 & 8, in the first two of which (especially) generalisations and extensions of the Lindley equation feature prominently.

[1] It is a well-known fact that the actual and virtual (steady-state) waiting-time distributions coincide iff the input is Poisson. (Compare this with the discussion in Appendix B on the excess-lifetime distribution for a Poisson process.)

Baccelli *et al.* [4] and Karpelevich *et al.* [35] form the basis of Chapters 5 and 6 respectively. The network discipline used in these two chapters also appears in a more general form in the article by Berezner & Malyshev [8] which is reviewed in Chapter 7. Chapter 8 is concerned with some recent work (Berezner *et al.* [9]).

CHAPTER 3

NETWORKS AND TELECOMMUNICATIONS

3.1 INTRODUCTION

A link in a communications network establishes a connection between two different points. A direct connection between every pair of points in a system with N sites (hence $\binom{N}{2}$ links) is not feasible, particularly when N is large. What is required is a partially-connected design which allows all N sites to communicate in an efficient manner with fewer than $\binom{N}{2}$ links. In graph-theoretical terminology (Appendix A) a connected graph — but not a complete graph — is desired.

The configuration of transmission facilities into a network capable of serving a number of distributed sites falls under the heading of *networking*. *Switching*, on the other hand, is a reference to the appropriate allocation of transmission devices in the partially connected topology. The *switch* in "switching networks" may be thought of as a flexible allocation device for opening, closing and directing an electrical circuit. Together, networking and switching contribute to an economical and efficient communications system. One of the criteria requiring efficient management is the transmission *bandwidth* (i.e., range of frequencies).

The next two sections are devoted to the topics of switching and network topology. §§ 4 & 5 serve as a transition from the computing/engineering arena to a more mathematical approach.

3.2 SWITCHING NETWORKS

Circuit-switched (or *line-switched networks*) are most commonly associated with the transmission of voice information, such as in telephone systems. They are also, however, sometimes applied in the transmission of data. We shall model the circuit-switched network using a finite connected multi-graph (Appendix A). The terminology used below derives from telephone and telecommunication networks. A node $i \in \mathbb{N}_K$ is a *switching centre* and an edge $e = ij$ is a *trunk group* or *link*. The multiplicity m_e of that edge is indicative of the capacity of the trunk group, i.e., the maximum number of calls a single trunk group can carry simultaneously, one call per *trunk* (*circuit*). The arrival streams are typically assumed to be Poisson processes. (This is not an unreasonable assumption.)

In the circuit-switched network, a communicating path of connected links is established from the source to the destination before transmission begins. Suppose it is necessary to establish a path between nodes i_0 and i_n. The *routing scheme* in operation will determine how that specific path may be established. A special signalling message (known as the *call*) is generated by source node i_0 and proceeds from there to switching centre i_1 on an idle circuit. Then that circuit is no longer available to another call. At node i_1, the routing scheme dictates which of r allowable trunk groups the call is offered. If that "first-choice" trunk group is busy (so that all of its trunks are engaged with other calls), then our call is offered an alternative trunk group. The process continues in this way until either a free trunk group is found or the call is rejected. If the call is successful in reaching node i_2, the entire procedure repeats itself. Ultimately, if a complete i_0-i_n path is established (electrically), the call is (naturally enough) said to be *completed*; otherwise, it is *blocked*. In the former case, a return signal informs the source that the path is completed, whereafter the source is able to transmit. (The path remains connected and dedicated to a call until it is released by the communicating parties.) During the *holding (conversation) time*, therefore, the end-users essentially have a continuous circuit connecting them. When this period has elapsed, the trunks occupied by that call again become idle.

In data communications, information exchange occurs in the form of blocks of data referred to as *messages*. In a *message-switched* network, a message that arrives at a node from a link or source is directed to an adjacent node along a link which, again, is selected according to a routing scheme. The process is repeated, with the message (physically) advancing in a node-by-node fashion. (These switching nodes are often computers with processing and storage capabilities.) The situation is, however, different from that in circuit-switching. In message-switching, there is no need to "keep control over" previously-visited links. Furthermore, if transmission capacity is not available on some link, then, rather than being rejected, the message is placed in a queue associated with that link, behind messages which preceded it (storage). The procedure is thus also known as the *store-and-forward* technique. It is interesting to note that message-switching is often applied in *email* (electronic mail) technology.

Circuit-switching is effective in a situation where an established circuit will be exposed to a steady flow of information (such as during a normal telephone conversation). On the other hand, message-switching is better suited to inter-computer communication which is typically *bursty*: varied in load and prone to random gaps inherent in the message-generation process. (For instance, many devices, such as keyboard terminals, transmit traffic in bursts. The data channel is idle while the user inputs more data or pauses to think about a problem.) Furthermore, although the communication facilities are used comparatively seldom, when they are utilized, the machinery must respond with relatively high speed. The end-to-end circuit characterising circuit-switching is not an appropriate transmission method because of the low channel utilisation.

Message-switching is preferred in the bursty communication scenario because messages occupy only one link at a time, and only for the period of transmission between its two switches. During the rest of the time, the link is available to other messages. The dynamic allocation of bandwidth (for a particular message and link) is an advantage which message-switching has over circuit-switching.

Full advantage is taken of the dynamic and fair allocation of bandwidth in *packet-switching* which is a particular, and more recent, form of message-switching. In a packet-switched network, constraints on message length means that those blocks of data exceeding the limits are divided into packets of fixed length. As before, each packet is forwarded along with address information from node to node until it reaches its destination, where the message is reconstructed from all the constituent packets.

In a packet-switched network, many packets from the same message may be transmitted simultaneously over successive links of a path from source to destination. This mostly reduces end-to-end transmission delays. The problem of transmission error is also reduced because packets are more manageable than the messages they form. So, if multipath routing is implemented, then resource utilisation and network reliability are increased. Although there are drawbacks, such as the need to reorganise the packets into their correct sequence at the destination node, packet-switching has become an economically viable method for the transmission of data.

Typically, the performance of a circuit-switched network is based on the overall *blocking probability*. The effectiveness of a message-switched network, on the other hand, is commonly judged by the average delay for all messages entering the network. Clearly the routing rule used will affect the performance: hybrid switching and routing schemes could prove the most efficient. (*Hybrid switching* combines circuit switching and message (or packet) switching. For a particular instance and more details, see Kelbert *et al.* [37] and Saito & Kawashima [55]. In the latter, other switching techniques, including *burst switching*, are also considered.) Further relevant references (many of which were used in writing this section and other sections of the present chapter) are those numbered [10], [21], [22], [25], [26], [34], [39], [48], [50], [52] and [57].

Mathematical tools from a wide range of fields including nonlinear, integer and dynamic programming, graph theory, numerical methods and, of course, applied probability and decision theory, are used in the design and evaluation of switching networks. The role of mathematics and applied probability in communication technology is examined in later in this chapter and in subsequent chapters.

3.3 NETWORK TOPOLOGY

In the introduction, mention was made of distributed sites. These sites may fall within a local, metropolitan or wide area. Correspondingly a network may be designated as a LAN, MAN or WAN. The first is the familiar local area network found in schools, universities and office buildings.

Network topology is a description of the shape or physical connectivity of the network. As ever, efficiency and economy are major factors in determining which configuration to use. The most commonly encountered LAN topologies are the *star, bus* and *ring* configurations, the names of which are indicative of their respective forms. The *hierarchical* or *tree* topology is also encountered, typically in mainframe environments.

(a) **Star Topology**

This configuration is also mentioned in Remark A.4.1. The primary feature is the access which each station has to a single central node or hub. In a *switching star network*, the central hub provides buffering and switching functions; the *broadcast star network* merely makes all incoming signals available to all outgoing lines.

(b) **Bus (Horizontal) Topology**

The stations are connected to a backbone cable, the *bus*, which acts as the transmission medium. Stations receive one another's transmissions in the case of a *logical bus* or *broadcast* topology. In the *physical bus* network, each station may sense and receive network transmissions but cannot interfere with existing network data. A drawback in using the horizontal topology is that there is usually only one transmission channel available. If that channel fails, then the network is lost.

(c) **Ring Topology**

Stations are arranged in a closed loop. Interference with common data by a station is allowed. Usually each station receives messages from only one station and transmits to only one station (those occupying the positions in the ring adjacent to it). In other words, data flows in one direction only. The inevitable drawback is that the serial connection of the stations will mean that a failed unit is a failed system (unless, of course, backup is provided).

The metropolitan-, and especially the wide-, area networks (MANs, WANs) often have a so-called *general mesh topology* (*web*). Switching centres are connected across potentially long distances; they, in turn, are linked to regional network systems, and eventually to the LANs considered earlier in this section. Another characteristic of the mesh topology is that several different paths of equal length (Appendix A) may connect any pair of stations.

With this background established, it is now appropriate to investigate the operation of a queueing network in greater detail.

3.4 SYNCHRONIZATION CONSTRAINTS

Consider a typical queueing network with $J \in \mathbb{N}$ nodes (labelled as $1, \ldots, J$). In this section, the following notation (similar to that in Part I) will be used for each node $k \in \mathbb{N}_J$.

t_n^k time of nth arrival into node k
σ_n^k time at which the nth service in the kth node begins
s_n^k service time for nth arrival in node k
δ_n^k time of nth departure from the kth node

The property that $t_n^k \leq t_{n+1}^k \; \forall \, n \in \mathbb{N}$ cannot really be regarded as a constraint. Rather it is a logical consequence of the "definition" of t_n^k. When a particular queue discipline such as $FCFS$ is in operation at a node, then the sequence of arrival times $\left\langle t_n^k, n \in \mathbb{N} \right\rangle$ imposes a *total order* on the input to node k. For it is well-known that, in the case of a $FIFO$ queue, the index n will refer to the nth customer throughout and so:

$$t_n^k \leq \sigma_n^k \leq \delta_n^k = \sigma_n^k + s_n^k \tag{3.4.1}$$

Moreover, $\sigma_n^k - t_n^k$ and $\delta_n^k - t_n^k$ are the waiting and sojourn times respectively of the nth customer in k.

Admittedly, constraints such as those in (3.4.1) are inherent to the definition of the operation of the network, nonetheless, we decide to formulate the following definition (Baccelli & Makowski [5]).

Definition 3.4.1 (Synchronization Constraints)
By *synchronization* between (possibly distinct) nodes $j, k \in \mathbb{N}_J$ shall be meant a global constraint satisfied by two sequences of times associated with

these nodes.

The sequel is devoted to specific instances of synchronization. Examples of particular interest are given.

(a) **Arrival-Arrival Synchronization**

Establishes a relationship between $\langle t_n^j \rangle$ and $\left\langle t_n^k \right\rangle$, $j \neq k$.

(b) **Departure-Departure Synchronization**

A relationship between $\langle \delta_n^j \rangle$ and $\left\langle \delta_n^k \right\rangle$ $(j \neq k)$ is determined.

Example 3.4.1 (Fork-Join Network)

The fork-join principle is of value in several computer and (flexible) manufacturing applications. To understand what is meant by a K-dimensional fork-join queue, let us consider $K \in \mathbb{N}$ parallel $FIFO$ queues constituting a system to which arrivals occur as batches of size $B \in \mathbb{N}_K$ (often $B = K$). Upon arrival such a batch is split, so that each of the B customers is allocated to a (different) single server.

The situation could be viewed in a slightly different way, as the simultaneous arrival of customers to B (often all) of the K queues. In other words, we have an arrival-arrival synchronization constraint — called a *fork* primitive — of the form $t_n^j = t_n^k \;\forall\, j,k \in \mathbb{N}_B$ for each fixed n. There is no guarantee, however, that the B customers complete their service at the same time, but we do require that a departure-departure constraint, called a *join* primitive, be satisfied. Specifically we demand that $\delta_n^j = \delta_n^k \;\forall\, j,k \in \mathbb{N}_B$ for each fixed n. Often, a buffer is used to hold those customers which have finished their service before others from the same group. An alternative method is to block the server until $t = \max_j \delta_n^j$ is reached, but this approach does not allow the server to attend to a customer from the next batch before t.

Theoretically-speaking, a fork primitive can occur in a network without a join primitive, or conversely. As mentioned initially, though, there are several instances where they occur together. A production line (for cars, television sets, fridges, ...) often requires the manufacture of components before final assembly. In such a case, the servers would be the machines which manufacture the K (say) different components. When an order is placed for an item, the K parts arrive and are allocated to the appropriate machines in accordance with the fork principle. After service, the join concept comes into operation: this allows for all components to be simultaneously available for assembly.

It is often convenient to describe a fork-join queueing network in terms of a digraph (Appendix A) in which nodes represent tasks and arcs represent precedence relations between tasks. When more than one edge emanates from a node, then we have a fork node; a join node is characterised by several edges entering such a node. The usefulness of such a representation (aside from motivating the terms *fork* and *join*) is evident in parallel computer processes. A batch customer may, for example, be a computer program which comprises several subroutines, each of which may be executed simultaneously in a different processor (the fork primitive). Before further program execution is possible, the results from the subroutines must be gathered together (the join primitive). Programming languages like ADA make use of this fork-join construct.

More extensive detail on this and related topics is provided in Baccelli & Makowski [5, §V], Baccelli *et al.* [6], Baccelli *et al.* [7] and Nelson & Tantawi [49].

(c) **Service-Service Synchronization**

Restriction placed on the connection between $\langle \sigma_n^j \rangle$ and $\left\langle \sigma_n^k \right\rangle$, $j \neq k$.

(d) **Departure-Arrival Synchronization**

Relationship established between $\langle \delta_n^j \rangle$ and $\left\langle t_n^k \right\rangle$.

Example 3.4.2 (Queues in Series)

A system of J $FIFO$ queues in tandem usually has the constraint that $\delta_n^j = t_n^{j+1} \; \forall \, j \in \mathbb{N}_{J-1}$.

(e) **Service-Arrival Synchronization**

$\langle \sigma_n^j \rangle$ and $\left\langle t_n^k \right\rangle$ satisfy some constraint.

Example 3.4.3 (Resequencing in (Packet-Switched) Networks)

Recall that in §3.2 mention was made of the fact that the packets of a message have to be reorganised into the correct sequence at their destination node in a packet-switched network.

A resequencing buffer stores those packets which were received out of order, that is, until such time as the earlier-transmitted packets have been processed (served). This translates into the fact that, with j

representing the destination node, $t^j_m < \sigma^j_n \; \forall\, m < n$ (service-arrival synchronization).

In the next section greater attention will be given to the concept of resequencing in networks.

3.5 THE RESEQUENCING TECHNIQUE

Baccelli *et al.* [3], and Baccelli & Makowski [5] have proposed a general (and very useful) resequencing model, of which we shall soon consider a special case. A simple diagram illustrates the generic model.

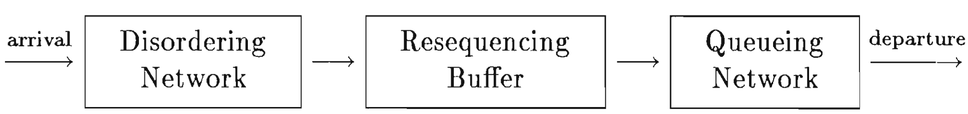


Fig. 3.1 General Resequencing Model

Customers y arrive to the first network in a specific order. Suppose, in particular, that the nth customer y_n arrives at time t_n and later departs from that network at time δ_n. The network is said to be *disordering* for, although $t_j < t_n \; \forall\, j \in \mathbb{N}_{n-1}$, the possibility exists that $\delta_n < \delta_j$ for one or more indices j from $\mathbb{N}_{n-1}$.

Before customers may enter the queueing network proper, their original order of arrival must be restored.[1] This is achieved by the *resequencing buffer* in which $\forall\, n\ y_n$ is delayed (if necessary) until y_{n-1} has left the buffer. Ultimately this means that the order of departure from the resequencing buffer must coincide with the order of arrival to the disordering network.

A particular resequencing system of interest is that in which the queueing network reduces to a single $FCFS$ queue. The intention is to formulate recursive equations for a number of quantities measuring the effectiveness (mainly delay) of the network system. In order to do so, a fairly simple notation is developed. Superscripted $^{(d)}$,$^{(r)}$ and $^{(q)}$ represent the disordering network, resequencing buffer and $FCFS$ queue *(excluding the actual service)* respectively. As before, y_n will denote the nth customer (call, message, ...). Also (much as before):

$t_n^{(\cdot)}$	time of arrival to "$\cdot$" of the nth customer
$\tau_n^{(\cdot)}$	$t_{n+1}^{(\cdot)} - t_n^{(\cdot)}$, interarrival times in "$\cdot$"
$w_n^{(\cdot)}$	delay/waiting time of the nth customer in "$\cdot$"
$W_n^{(\cdot,*)}$	combined delay in "$\cdot$" and "$*$"
s_n	service time in the $FCFS$ queue
δ_n	departure time from the $FCFS$ queue, after service
T_n	$\delta_n - t_n^{(d)}$ end-to-end system delay

[1]This establishes a connection with Venkatasubramanyan & Sanders [58] in which order restoration in flexible manufacturing systems (rather than computer-communication networks) is considered.

Lemma 3.5.1 (Recursive Equations for the Resequencing Model) For the resequencing model described previously, the following equations hold:

$$w_{n+1}^{(r)} = \left[w_n^{(r)} - \left(w_{n+1}^{(d)} - w_n^{(d)}\right) - \tau_n^{(d)}\right]^+ \tag{3.5.1}$$

$$w_{n+1}^{(q)} = \left[w_n^{(q)} + s_n - \tau_n^{(q)}\right]^+ \tag{3.5.2}$$

$$W_{n+1}^{(d,r)} = \max\left[w_{n+1}^{(d)}, W_n^{(d,r)} - \tau_n^{(d)}\right] \tag{3.5.3}$$

$$W_{n+1}^{(r,q)} = \left[W_n^{(r,q)} + s_n - \left(w_{n+1}^{(d)} - w_n^{(d)}\right) - \tau_n^{(d)}\right]^+ \tag{3.5.4}$$

$$W_{n+1}^{(d,r,q)} = \max\left[w_{n+1}^{(d)}, W_n^{(d,r,q)} + s_n - \tau_n^{(d)}\right] \tag{3.5.5}$$

$$T_{n+1} = w_{n+1}^{(d)} + s_{n+1} + \left[T_n - w_{n+1}^{(d)} - \tau_n^{(d)}\right]^+ \tag{3.5.6}$$

<u>PROOF</u>

Before we verify equations (3.5.1)–(3.5.6), some remarks are in order. Firstly, a recursive equation for $w_{n+1}^{(d)}$ does not appear in (3.5.1)–(3.5.6) simply because there is insufficient information on the operation of the disordering network. Secondly, the following equations (self-evident though they may be) will prove useful later:

$$t_n^{(r)} = t_n^{(d)} + w_n^{(d)} \tag{3.5.7a}$$

$$t_n^{(q)} = t_n^{(r)} + w_n^{(r)} = t_n^{(d)} + W_n^{(d,r)} \tag{3.5.7b}$$

We now proceed to prove (3.5.1)–(3.5.6), but not necessarily in that order.

(a) Consider equation (3.5.1), which may be rewritten as

$$
\begin{aligned}
w_{n+1}^{(r)} &= \left[w_n^{(r)} - \left[\left(t_{n+1}^{(d)} + w_{n+1}^{(d)}\right) - \left(t_n^{(d)} + w_n^{(d)}\right)\right]\right]^+ \\
\Rightarrow \quad w_{n+1}^{(r)} &= \left[w_n^{(r)} - \left(t_{n+1}^{(r)} - t_n^{(r)}\right)\right]^+
\end{aligned}
\tag{3.5.1$'$}
$$

where (3.5.7a) is used in the last step. It suffices to verify (3.5.1′).

Case 1: $t_{n+1}^{(r)} \leq t_n^{(r)}$

y_{n+1} needs to wait in the resequencing buffer until the arrival of y_n that itself waits for a time interval of $w_n^{(r)}$. Thus $w_{n+1}^{(r)} = w_n^{(r)} + \left(t_n^{(r)} - t_{n+1}^{(r)}\right) \geq 0$ so that (3.5.1′) is valid in this case.

Case 2: $t_{n+1}^{(r)} > t_n^{(r)}$

When y_{n+1} reaches the resequencing buffer

(i) y_n will already have left the buffer if $t_{n+1}^{(r)} - t_n^{(r)} > w_n^{(r)}$ so that y_{n+1} need not wait there

(ii) a part of $w_n^{(r)}$, namely $t_{n+1}^{(r)} - t_n^{(r)}$, has already elapsed if $t_{n+1}^{(r)} - t_n^{(r)} \leq w_n^{(r)}$ so that y_{n+1} need only wait in the resequencing buffer for a time of $w_n^{(r)} - \left(t_{n+1}^{(r)} - t_n^{(r)}\right) \geq 0$

(3.5.1′) remains valid in either instance.

(b) As a corollary to (a), we have

$$\begin{aligned} W_{n+1}^{(d,r)} &= w_{n+1}^{(d)} + w_{n+1}^{(r)} \\ &= w_{n+1}^{(d)} + \max\left[0, w_n^{(r)} - \left(w_{n+1}^{(d)} - w_n^{(d)}\right) - \tau_n^{(d)}\right] \\ &= \max\left[w_{n+1}^{(d)}, w_n^{(d)} + w_n^{(r)} - \tau_n^{(d)}\right] \\ &= \max\left[w_{n+1}^{(d)}, W_n^{(d,r)} - \tau_n^{(d)}\right] \end{aligned}$$

which is (3.5.3).

(c) A fairly similar argument to that used in (a) may be applied to verify (3.5.2), the Lindley equation, key to much of this thesis. The queue discipline is first-come-first-served so that the $(n+1)$th customer must wait for the nth customer (and $t_{n+1}^{(q)} > t_n^{(q)}$).
Now when y_n arrives at time t_n, he experiences an end-to-end delay of $w_n^{(q)} + s_n$. As a result, before y_{n+1} (arriving at t_{n+1}) can enter service, he must wait a time of $w_n^{(q)} + s_n - \left(t_{n+1}^{(q)} - t_n^{(q)}\right)$. This expression is, of course, the portion of the sojourn time of y_n which remains when y_{n+1} arrives. There is the very distinct possibility, however, that $w_n^{(q)} + s_n < \tau_n^{(q)}$. In that case y_{n+1} need not wait at all: y_n has already left. A combination of these arguments leads to (3.5.2).

(d) The *FIFO* discipline does not hold in general for messages arriving to the resequencing buffer. The Lindley principle deduced in (c), however, will be valid for $W_{n+1}^{(r,q)}$. To appreciate this, note that $-\tau_n^{(r)} = t_n^{(r)} - t_{n+1}^{(r)} > 0$ if y_{n+1} arrives before y_n at the resequencing

buffer, so this would represent an increase *not* a reduction in the waiting time of y_{n+1}. (That increase mirrors the fact that y_{n+1} can be sent through to the *FIFO* queue only when y_n has arrived after a time of $-\tau_n^{(r)}$ has elapsed.)
Appropriate adaptation of (3.5.2) then yields

$$W_{n+1}^{(r,q)} = \left[W_n^{(r,q)} + s_n - \tau_n^{(r)}\right]^+$$

but

$$\begin{aligned}\tau_n^{(r)} &= t_{n+1}^{(r)} - t_n^{(r)} \\ &= w_{n+1}^{(d)} - w_n^{(d)} + \tau_n^{(d)} \quad \text{an application of (3.5.7a)}\end{aligned}$$

from which (3.5.4) is deduced.

(e) It is obvious that

$$\begin{aligned}T_{n+1} &= w_{n+1}^{(d)} + W_{n+1}^{(r,q)} + s_{n+1} \qquad (3.5.8) \\ &\overset{(d)}{=\!=} w_{n+1}^{(d)} + s_{n+1} + \left[W_n^{(r,q)} + s_n - \left(w_{n+1}^{(d)} - w_n^{(d)}\right) - \tau_n^{(d)}\right]^+ \\ &\overset{(3.5.8)}{=\!=\!=} w_{n+1}^{(d)} + s_{n+1} + \left[T_n - w_n^{(d)} - \left(w_{n+1}^{(d)} - w_n^{(d)}\right) - \tau_n^{(d)}\right]^+ \\ &= w_{n+1}^{(d)} + s_{n+1} + \left[T_n - w_{n+1}^{(d)} - \tau_n^{(d)}\right]^+\end{aligned}$$

which is (3.5.6).

(f) Equation (3.5.5) may be derived as a corollary to (e).

$$\begin{aligned} W_{n+1}^{(d,r,q)} &= T_{n+1} - s_{n+1} \qquad (3.5.9) \\ &\overset{(3.5.6)}{=\!=} w_{n+1}^{(d)} + \max\left[0, T_n - w_{n+1}^{(d)} - \tau_n^{(d)}\right] \\ &= \max\left[w_{n+1}^{(d)}, T_n - \tau_n^{(d)}\right] \\ &\overset{(3.5.9)}{=\!=} \max\left[w_{n+1}^{(d)}, W_n^{(d,r,q)} + s_n - \tau_n^{(d)}\right] \end{aligned}$$

□

3.6 CONCLUSION

This chapter was never intended to be a handbook for an electronic engineer or communications/computer technologist. The purpose it *does* fulfil is two-fold. First, Chapter 3 serves as a valuable introduction to the terminology and ideas which will be exploited in subsequent chapters in a more mathematical environment. In the second place, the discussion provided demonstrates that the mathematics and probability in this field (of queueing and network systems) are applied in real-world situations. It is due to these very applications and, indeed, the rapid advance of technology that the field of queueing theory has blossomed in recent decades.

CHAPTER 4

ON A CLASS OF NETWORKS AND DISCIPLINES

4.1 INTRODUCTION AND PRELIMINARIES

In Chapters 5, 6 and 8, there is a need to establish notation and terminology for models based on a symmetric starlike switching network. A degree of complication arises in that the network obviously will operate differently according to whether a circuit- or message-switching rule (§3.2) applies. Uniform notation, free of the switching rule involved, will, therefore, be of initial concern.

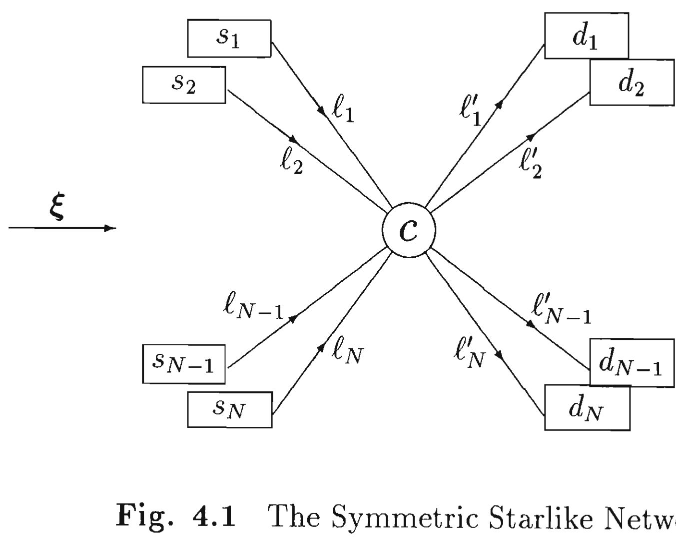


Fig. 4.1 The Symmetric Starlike Network

Call the central node c. Denote the collection of N peripheral sources by

$\mathcal{S} = \{s_i, i \in \mathbb{N}_N\}$ and that of the N peripheral destinations (addresses or receivers, depending on the context) by $\mathcal{D} = \{d_j, j \in \mathbb{N}_N\}$. When there is no ambiguity, we may replace s_i with i, and similarly with d_j. For the arcs $s_i c$ and cd_j, write ℓ_i and ℓ'_j respectively, where the former shall be called the ith *source line* and the latter, the jth *destination line*. In the context of graph theory (Appendix A), our directed star graph $\mathcal{G}$ is of the form $K_{1,2N}$ (Definition A.4.3 and Remark A.4.1) with a vertex set $V(\mathcal{G}) = \mathcal{S} \cup \{c\} \cup \mathcal{D}$ and an arc set $\mathcal{A} = \{\ell_i, i \in \mathbb{N}_N\} \cup \left\{\ell'_j, j \in \mathbb{N}_N\right\}$. Often it will be convenient to refer to a process (of service, waiting, ...) associated with ℓ_i ($i \in \mathbb{N}_N$) as a *first stage/phase process.* Naturally enough, *second stage* will correspond to the process associated with ℓ'_j. This summarises the structure of our network.

Next we need to consider the nature of the source input. A general arrival stream $\boldsymbol{\xi}$ decomposes into N IID stationary Poisson random marked flows $\{\xi_i, i \in \mathbb{N}_N\}$, the *source flows*, each of which is associated with a corresponding source s_i (and line ℓ_i). The common Poisson intensity is given by $\lambda > 0$.

The (conditionally) IID marks of each RMPP (Appendix B) ξ_i are given by pairs $(\boldsymbol{S}, d)$ where $\boldsymbol{S}$ is a vector of service times and $d \in \mathbb{N}_N$, naturally enough, refers to the destination. Where the issue of service is concerned, some distinction will arise between the circuit- and message-switching cases. As is evident from §3.2, a call in our circuit-switched star network may be processed only once the path (circuit) $s_i cd$ has been established. In this case the vector $\boldsymbol{S}$ reduces to $S(\in \mathbb{R}_\oplus) \sim F$. When the star network is based on the message-switching principle, then $\boldsymbol{S} = \left(S^{[1]}, S^{[2]}\right) (\in \mathbb{R}^2_\oplus) \sim F$ (say) with marginals $F^{[1]}$ and $F^{[2]}$. Here, of course, $S^{[1]}$ and $S^{[2]}$ represent the service times on the lines ℓ_i and ℓ'_d respectively. For either switching rule, $d \sim U_N$, the discrete uniform (equiprobability) distribution on $\mathbb{N}_N$ with $P[d = j] = \frac{1}{N}\ \forall\, j \in \mathbb{N}_N$. It is typically assumed that $\boldsymbol{S}$ (or S) and d are mutually independent. Consequently, the marginal distribution of the mark value is given by the product $F \times U_N$.

An analogous process of interest is given by the family $\boldsymbol{\xi}' = \left\{\xi'_j, j \in \mathbb{N}_N\right\}$ of IID stationary Poisson RMPPs, each of which represents a destination flow associated with ℓ'_j and d_j. Marks are given by pairs $(\boldsymbol{S}, s)$ where $\boldsymbol{S}$ is as before and $s \in \mathbb{N}_N$ is the inverse address, i.e., the (number of the) source at which the particular call (or message) in question arrived. The probability and independence conditions which applied to $(\boldsymbol{S}, d)$ are equally valid here. Taking these facts into consideration, it is evident that, while the criteria for their "decompositions" differ, the RMPP families $\boldsymbol{\xi}$ and $\boldsymbol{\xi}'$ are equivalent ways of describing the network. More especially, it is obvious that any ξ'_j has the same intensity λ as does ξ_i.

Often the notation $\boldsymbol{\omega} = \{\omega_i, i \in \mathbb{N}_N\}$, $\boldsymbol{\omega}' = \left\{\omega'_j, j \in \mathbb{N}_N\right\}$ will be used to represent the sample realisations of the corresponding ensemble flows $\boldsymbol{\xi}, \boldsymbol{\xi}'$ respectively. The terminology used for the "objects" of the flows will differ according to the situation. *Customer, unit, demand* or *request* are reasonably neutral terms but there may be a desire for a more modern approach relevant to communications applications. *Call* will often be more appropriate for a circuit-switched network, *message* or *program* for a message-switched system but this need not be a hard-and-fast rule.

$y = [t; (\boldsymbol{S}, d)]$ and $y = [t; (S, d)]$ represent messages and calls respectively. To emphasise that $d = j$ (say), we may write $y \in \omega_j'$, and to indicate that y has source i, it is often convenient to write $y \in \omega_i$. Otherwise, we may always consider $y = [t; (\boldsymbol{S}, s, d)]$. In either case, it will be convenient to be able to access individual components. This will be achieved in a logical way. For instance

$$\hat{y} = \left[\hat{t}; \left(\hat{S}, \hat{s}, \hat{d}\right)\right] \Longrightarrow$$

$$t(y) = \hat{t} \quad S(y) = \hat{S} \quad s(y) = \hat{s} \text{ and } d(y) = \hat{d}$$

4.2 PROTOCOLS AND DISCIPLINES

A variety of protocols and synchronization constraints (Baccelli *et al.* [4], Berezner *et al.* [9], Karpelevich *et al.* [35] and Rybko & Mikhailov [53]) have been considered for the second stage in the starlike network. It will become evident, however, that their application is not restricted to this situation alone. For instance, one such protocol is used in the mth stage of a "generalised starlike network" (Chapter 6) and for an even more general network in Chapter 7.

(a) $\boldsymbol{FAFS}$: **F**irst **A**rrived in the Network
First **S**erved in the Second Stage

(i) **Message-Switched Network**

A message y with source i and destination j arrives at the network at time t. Before transmission along the line ℓ_i, y must wait for a period $w^{[1]}$ for all messages $\tilde{y} \in \omega_i$ with $t(\tilde{y}) < t$ to leave this first stage, according to the usual $FCFS$ discipline. The actual transmission time is given by $S^{[1]}$ so that y has a sojourn time in the first phase of $T^{[1]} = w^{[1]} + S^{[1]}$, whereafter it relinquishes the line ℓ_i.

The second stage of the waiting process begins immediately thereafter. All messages $\breve{y} \in \omega'_j$ with $t(\breve{y}) < t$ must complete their transmission along ℓ'_j before y itself can be transmitted. In particular, y will wait a time of $w^{[2]}$ and then be served for an interval of duration $S^{[2]}$; $T^{[2]} = w^{[2]} + S^{[2]}$ is then the second-stage sojourn time. Theoretical motivation, however, renders the quantities $W = T^{[1]} + w^{[2]}$ and the *response time* $T = W + S^{[2]}$ of greater interest than $w^{[2]}$ and $T^{[2]}$ respectively. The vector $\boldsymbol{w} = \left(w^{[1]}, W\right)$ will also be of interest.

(ii) **Circuit-Switched Network**

A call y with source i and destination j arrives at the network at time t. Before it may gain access to the line ℓ_i, y must wait for a period of $w^{[1]} = T^{[1]}$ for all calls $\tilde{y} \in \omega_i$ with $t(\tilde{y}) < t$ to leave the network. (Naturally, some or all of these calls (if there are any) may already have departed before t.) This $FCFS$-type discipline is in compliance with the circuit-switching principle considered in §3.2. This rule also stipulates that, once y is allocated ℓ_i, it retains that line until it leaves the network.

An additional waiting period of $w^{[2]}$ (measured from $t + w^{[1]}$) ensues before y has access to ℓ'_j. During that time, all calls $\breve{y} \in \omega'_j$ with $t(\breve{y}) < t$ leave the network. (This allows for greater flexibility than in §3.2 where rejection of y is immediate if ℓ'_j is not available.) So y waits for an interval $W = T^{[1]} + w^{[2]}$ before it is processed for a time of S (counted from $t + W$). Ultimately y departs at $t + T$ after releasing the lines ℓ_i and ℓ'_j, and experiencing an end-to-end delay given by $T = W + S$.

(b) ***FAAFS***: **F**irst **A**rrived in the Network and **A**llocated to Service in the Second Stage, **F**irst **S**erved in the Second Stage

(i) **Message-Switched Network**

At time t message y, having inverse address i and destination j, arrives at the starlike network. It is delayed for $w^{[1]}$ during which time all messages $\tilde{y} \in \omega_i$ with $t(\tilde{y}) < t$ complete their first phase of processing in accordance with the $FCFS$ idea. Then ℓ_i is available to y, enabling it to be transmitted for $S^{[1]}$. The total first-stage delay is then given by $T^{[1]} = w^{[1]} + S^{[1]}$, after which y releases ℓ_i.

Message y begins its second stage immediately thereafter. If the server is free, y need not wait at all. Otherwise, some message $\breve{y} \in \omega_j'$ must be in service. In case $t(\breve{y}) > t$, two possibilities, determined by the priority rule, emerge:

(1) Under a *nonpreemptive* service rule, y proceeds to the head of the second-stage queue and waits for a time of $w^{[2]}$ while $\breve{y}$ and possibly other higher time-priority messages (arriving in the interim) complete their service

but

(2) under a *preemptive (resume)* rule, y interrupts the service of $\breve{y}$, which is forced to rejoin the queue.

The third situation which may arise is that $\breve{y}$ arrived to the network before y. Then y is allocated its position in the queue (much as in $FAFS$) according to $t = t(y)$ (messages with earlier times of arrival to the network situated closer to the head of the queue) and will wait for a period of $w^{[2]}$.

At the end of this period $w^{[2]}$ (which we have seen may be zero), y is processed for time $S^{[2]}$ and then leaves the network. As before, $T^{[2]} = w^{[2]} + S^{[2]}$, $W = T^{[1]} + w^{[2]}$ and $T = T^{[1]} + T^{[2]}$ represent various delay quantities of interest.

(ii) **Circuit-Switched Network**

A call y with source i and destination j arrives at the network at time t. y is then required to wait for a time of $w^{[1]} = T^{[1]}$ to be allocated ℓ_i while all other calls $\tilde{y} \in \omega_i$ with $t(\tilde{y}) < t$ leave the network. Once y has been allocated ℓ_i, it retains that line and waits to be allocated its second channel, ℓ'_j, in order to complete the circuit $s_i c d_j$.

As in (i), y need not wait to get the line ℓ'_j if the server in question is free or if the preemptive (resume) rule is used and the call being processed has $t(\breve{y}) > t$. (y may have to wait at a later stage if its own service is preempted.) In the latter case, the service of $\breve{y}$ is interrupted. Otherwise y *will* wait for some nonzero time $w^{[2]}$ according to the same basic principles outlined for the $FAAFS$ message-switching scenario. The service time is given by S, $W = T^{[1]} + w^{[2]}$ and $T = W + S$.

(c) ***FCFS***: **F**irst **C**ome to the Second Stage
First **S**erved in the Second Stage

This discipline is fairly straightforward and does not require detailed explanation.

A customer y having source i and destination j arrives at the network at time t. In accordance with the switching rule in question and the ordinary $FCFS$ discipline, it is allocated ℓ_i after time $w^{[1]}$. After an additional period it gains access to line ℓ'_j, again on the basis of the switching rule and $FCFS$ discipline. Thereafter it is ready for its (second-stage) service.

It is beneficial to view the various protocols in terms of the general resequencing model first encountered in detail in §3.5. We adopt (and adapt) notation from that section as well as from §3.4. Not only does this scheme allow us to compare the disciplines but it also provides a mechanism for defining them in a precise and concise way.

Initially we focus our attention on why the starlike network lends itself to disordering. For convenience, we consider a sample realisation ω_j' of ξ_j'. The subscript $_n$, as in y_n, will (for now) be indicative of the fact that y_n is the nth customer from ω_j' to arrive externally to the starlike network. The sequence of customers $\langle y_m, m \in \mathbb{N}\rangle$ from ω_j' may be partitioned into sequences of the form $\langle y_{i_n}\rangle$ where $y_{i_n} = y_m$ is the mth customer from ω_j' to arrive exogeneously to the network but also the nth customer from ω_i ($i \in \mathbb{N}_N$) headed for destination j.

The queue associated with ω_i ($i \in \mathbb{N}_N$) operates according to the $FCFS$ discipline. It follows that $t_{i_n}^{(d)} < t_{i_{n+1}}^{(d)}$ and $\delta_{i_n}^{(d)} < \delta_{i_{n+1}}^{(d)}$ determine the times of arrivals and departures from the disordering network for each $i \in \mathbb{N}_N$. It is feasible, in fact, very likely that there are instances in which, for customers $y_{i_n}, y_{k_l} \in \omega_j'$, $t_{i_n}^{(d)} < t_{k_l}^{(d)}$ but $\delta_{i_n}^{(d)} > \delta_{k_l}^{(d)}$. This situation may arise because the ith and kth queues operate independently and because the waiting times are usually random variables. Consequently it is quite possible to have $w_{i_n}^{(d)} > w_{k_l}^{(d)} + t_{k_l}^{(d)} - t_{i_n}^{(d)}$, which leads to disordering.

The second-stage protocols we have considered compensate differently for the discrepancy between the sequence of first and second phase arrival times. These protocols obviously have different degrees of efficiency (in terms of response/waiting times and fairness to customers), as we shall see shortly.

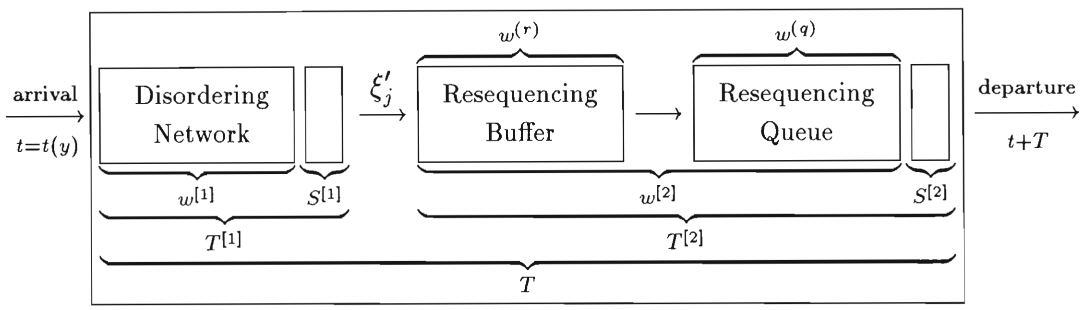


Fig 4.2(a) Resequencing Model

Message-Switched Starlike Network

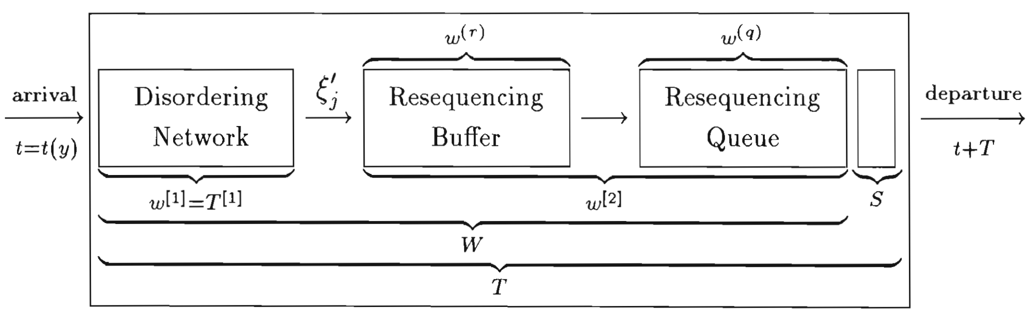


Fig. 4.2(b) Resequencing Model

Circuit-Switched Starlike Network

In both the message- and circuit-switched cases

$$w^{(d)} = T^{[1]} \quad W^{(r,q)} = w^{[2]} \ \text{ and } \ W^{(d,r,q)} = W$$

Generally, it will not be possible to interpret the disordering network, resequencing buffer and queue as being in series. Thus, although Figures 4.2(a) and 4.2(b) may give the impression of structures in tandem, a customer in the second stage still has an impact on the mechanism of the first. Finally, a protocol is regarded as work-conserving if it may be seen to be work-conserving within the confines of the switching rule. We stress though that, by design, a circuit-switching discipline is "less work-conserving" than a message-switching one because of the fact that, in the former case, access to a full circuit is required. For instance, consider two calls y and $\tilde{y}$ in a circuit-switched starlike network. Suppose that $y \in \omega_i \cap \omega'_j$, $\tilde{y} \in \omega_i \cap \omega'_k$ and y is being processed on the circuit $\ell_i \ell'_j$, but that ℓ'_k is free. $\tilde{y}$ is then prevented from gaining access to ℓ'_k because ℓ_i is occupied (simultaneously with ℓ'_j).

We are now able to present (summary) Table 4.1.

	(a) ***FAFS***	(b) ***FAAFS***	(c) ***FCFS***
Time-sequence determining order of Second-Stage Service	$\langle t_n \rangle$	$\langle t_n \rangle$ and $\left\langle t_n + T_n^{[1]} \right\rangle$	$\left\langle t_n + T_n^{[1]} \right\rangle$
Role of Resequencing Buffer	y_n delayed for $w_n^{(r)} \geq 0$ not work-conserving	trivial "identity resequencing" $w_n^{(r)} = 0$ work-conserving	trivial "identity resequencing" $w_n^{(r)} = 0$ work-conserving
Transmission from Buffer to Queue	y_n forwarded at $t_n + T_n^{[1]} + w_n^{(r)}$ only when $\langle y_k, k \in \mathbb{N}_{n-1} \rangle$ have been forwarded	y_n instantaneously forwarded at $t_n + T_n^{[1]}$	y_n instantaneously forwarded at $t_n + T_n^{[1]}$
Second-Stage Queue	trivial "identity resequencing" ——— *FCFS*	resequencing: new arrival from resequencing buffer assigned to position in queue based on $\langle t_n \rangle$ ——— *FCFS* IF arrivals stopped at some point	trivial "identity resequencing" ——— *FCFS*
Second-Stage Service	nonpreemptive	nonpreemptive or preemptive	nonpreemptive
System Evaluation	fair inefficient	fair efficient optimal scheme	unfair efficient

Table 4.1

Comparison of Second-Stage Protocols based on the Resequencing Model

The final row-block of Table 4.1, entitled *System Evaluation*, although based in part on theoretical arguments, was mostly intuition-based before a simulation study *for the "M/M" case* was undertaken. The reader is encouraged to refer to Appendix D where the technical (and theoretical) details pertaining to the computer program and the simulation procedure are provided. In the present section, we will be concerned mostly with presenting and briefly analysing the results of the simulation, nonetheless, we initially do make some relevant technical remarks.

(a) Figures 4.3 and 4.4 are graphical representations of the data contained in Tables 4.2(a) and 4.3(a)–(b) respectively. Tables 4.2(b) and 4.3(c) provide supplementary $FAFS$ data.

(b) Although the program (Appendix D) allows for greater flexibility, an equal and fairly large number, $N = 200$, of sources and destinations was assumed for our starlike network model (with $m = 2$ stages).

(c) The main objective was to find the mean sojourn time (denoted by[1] ET in Figures 4.3 and 4.4, and below) of a message/call in the network. The figures given are (grand) averages based on three separate trials of half a million (arriving) requests each. All requests leaving the first stage (99–100% of the total) had their $T^{[1]}$ incorporated into $ET^{[1]}$ (the Stage 1 figure in the tables); all those leaving the network (again, 99%+ of those arriving) contributed to $ET^{[2]}$ (the Stage 2 value). The total, ET, was determined by $ET^{[1]} + ET^{[2]}$ which, together with rounding[2], has led to some apparent discrepancies (in the totals) evident in the

[1] *not* ET (which is the actual expected value), although $ET \approx$ET

[2] to three decimal places

tables.

(d) Each of the $N = 200$ arrival streams is Poisson with rate λ. In the message-switching case, with $i = 1, 2$, service time $S^{[i]} \sim \text{Exp}\left((\mu^{[i]})^{-1}\right)$ and $\rho^{[i]}$ (or $\rho.i$) $= \frac{\lambda}{\mu^{[i]}} = \lambda$. (With the exception of the results of Table 4.2(b), we standardise: $\mu^{[1]} = 1 = \mu^{[2]}$, although the program is far more flexible than this.) For the circuit-switched network, $S \sim \text{Exp}(\mu^{-1})$ with $\mu = 1$ so that $\rho = \lambda$. We have not yet presented the theory motivating our choice for the range of values of the traffic intensity. This issue will, however, be clarified in later chapters.

(e) Abbreviations $FAFS$ and $FCFS$ are as before while $FAAFS_N$ (or $FAAFS\_N$) and $FAAFS_P$ (or $FAAFS\_P$) distinguish the nonpreemptive and preemptive forms (respectively) of the $FAAFS$ discipline.

(f) In addition to comparisons between the four protocols, one may also compare the observed averages with the expected $FCFS$ values in the extreme right-hand columns of Table 4.2(a) and Tables 4.3(a)–(c), and bracketed in Table 4.2(b).

  (i) In the case of message-switching $\text{E}T^{[i]}_{FCFS} = \frac{1}{\mu^{[i]} - \lambda} = \frac{1}{\mu^{[i]}(1-\rho^{[i]})}$, a well-known result for the ordinary $M/M/1$ queue.

  (ii) For the circuit-switched networks, we use a heuristically-determined value which, as is evident from the simulation results, works very well. We transform the entire network into $M(\lambda_{\text{sys}})/M(\mu_{\text{sys}})/1$ queues where $\lambda_{\text{sys}} = \lambda$ and $\mu_{\text{sys}} = \mu - \lambda$, i.e., we treat the stage 2 sojourn time as the system service time. Consequently

$$\rho_{\text{sys}} = \frac{\lambda}{\mu-\lambda} \quad \text{E}T = \frac{1}{\mu_{\text{sys}} - \lambda_{\text{sys}}} = \frac{1}{\mu - 2\lambda} \quad \text{E}T^{[1]} = \text{E}T - \text{E}T^{[2]} = \frac{\lambda}{(\mu-\lambda)(\mu-2\lambda)}$$

$\rho^{[1]} = \rho^{[2]}$	Stage	$FAFS$	$FAAFS_N$	$FAAFS_P$	$FCFS$	$ET^{[i]}_{FCFS}$
	1	1.111	1.110	1.111	1.111	1.111
0.1	2	1.204	1.112	1.109	1.110	1.111
	Total	2.314	2.223	2.219	2.221	2.222
	1	1.249	1.249	1.249	1.249	1.250
0.2	2	1.491	1.249	1.248	1.250	1.250
	Total	2.740	2.498	2.497	2.499	2.500
	1	1.427	1.430	1.432	1.427	1.429
0.3	2	1.903	1.428	1.429	1.425	1.429
	Total	3.329	2.859	2.860	2.852	2.857
	1	1.664	1.668	1.662	1.661	1.667
0.4	2	2.522	1.673	1.660	1.660	1.667
	Total	4.185	3.342	3.322	3.321	3.333
	1	1.989	2.004	2.012	1.991	2.000
0.5	2	3.511	1.993	1.993	2.001	2.000
	Total	5.500	3.997	4.005	3.992	4.000
	1	2.499	2.502	2.501	2.495	2.500
0.6	2	5.272	2.501	2.508	2.480	2.500
	Total	7.771	5.003	5.009	4.976	5.000
	1	3.327	3.325	3.326	3.321	3.333
0.7	2	8.689	3.333	3.317	3.320	3.333
	Total	12.016	6.658	6.643	6.641	6.667
	1	4.945	5.013	4.943	4.918	5.000
0.8	2	15.775	5.007	4.933	4.914	5.000
	Total	20.720	10.019	9.876	9.832	10.000
	1	9.617	9.513	9.542	9.564	10.000
0.9	2	38.710	9.455	9.575	9.451	10.000
	Total	48.327	18.968	19.116	19.014	20.000

Table 4.2(a)

Mean waiting times for a message-switched starlike network

$\rho^{[2]}$ \ $\rho^{[1]}$	**Stage**	0.375 (0.400)	0.500 (0.667)	0.750 (2.000)
0.375 (0.400)	1	0.400	0.665	1.975
	2	0.586	0.850	3.060
	Total	0.986	1.516	5.035
0.500 (0.667)	1	0.399	0.669	1.998
	2	0.870	1.178	3.706
	Total	1.269	1.846	5.704
0.750 (2.000)	1	0.399	0.667	2.001
	2	2.234	2.632	5.784
	Total	2.633	3.299	7.785

Table 4.2(b)

Mean waiting times for a message-switched starlike network

Special values for $FAFS$ discipline

$\rho = \frac{\lambda}{\mu}$	Stage	$FAFS$	$FAAFS_N$	$FAAFS_P$	$FCFS$	$\mathbf{E}T^{[i]}_{FCFS}$
0.025	1	0.027	0.027	0.027	0.027	0.027
	2	1.027	1.025	1.026	1.026	1.026
	Total	1.055	1.052	1.053	1.053	1.053
0.050	1	0.059	0.059	0.059	0.059	0.058
	2	1.058	1.052	1.053	1.052	1.053
	Total	1.117	1.110	1.112	1.110	1.111
0.075	1	0.097	0.094	0.095	0.095	0.095
	2	1.092	1.080	1.079	1.081	1.081
	Total	1.189	1.175	1.174	1.176	1.176
0.100	1	0.146	0.138	0.138	0.139	0.139
	2	1.137	1.110	1.110	1.112	1.111
	Total	1.284	1.248	1.249	1.251	1.250
0.125	1	0.212	0.189	0.189	0.192	0.190
	2	1.196	1.141	1.142	1.143	1.143
	Total	1.408	1.330	1.330	1.335	1.333
0.150	1	0.309	0.250	0.249	0.251	0.252
	2	1.277	1.174	1.176	1.176	1.176
	Total	1.586	1.425	1.425	1.427	1.429
0.175	1	0.495	0.323	0.318	0.326	0.326
	2	1.425	1.211	1.210	1.211	1.212
	Total	1.920	1.535	1.527	1.537	1.538
0.200	1	1.814	0.415	0.404	0.415	0.417
	2	2.131	1.251	1.251	1.250	1.250
	Total	3.945	1.666	1.655	1.665	1.667

Table 4.3(a)

Mean waiting times for a symmetric circuit-switched starlike network

$\rho = \frac{\lambda}{\mu}$	Stage	$FAAFS_N$	$FAAFS_P$	$FCFS$	$\mathbf{E}T^{[i]}_{FCFS}$
	1	0.654	0.624	0.663	0.667
0.250	2	1.333	1.332	1.330	1.333
	Total	1.987	1.956	1.993	2.000
	1	1.032	0.961	1.056	1.071
0.300	2	1.428	1.424	1.422	1.429
	Total	2.460	2.384	2.478	2.500
	1	1.648	1.507	1.753	1.795
0.350	2	1.529	1.534	1.529	1.538
	Total	3.178	3.040	3.282	3.333
	1	2.855	2.430	3.168	3.333
0.400	2	1.661	1.654	1.651	1.667
	Total	4.516	4.084	4.820	5.000
	1	5.646	4.539	7.776	8.182
0.450	2	1.806	1.808	1.804	1.818
	Total	7.452	6.348	9.580	10.000

Table 4.3(b)

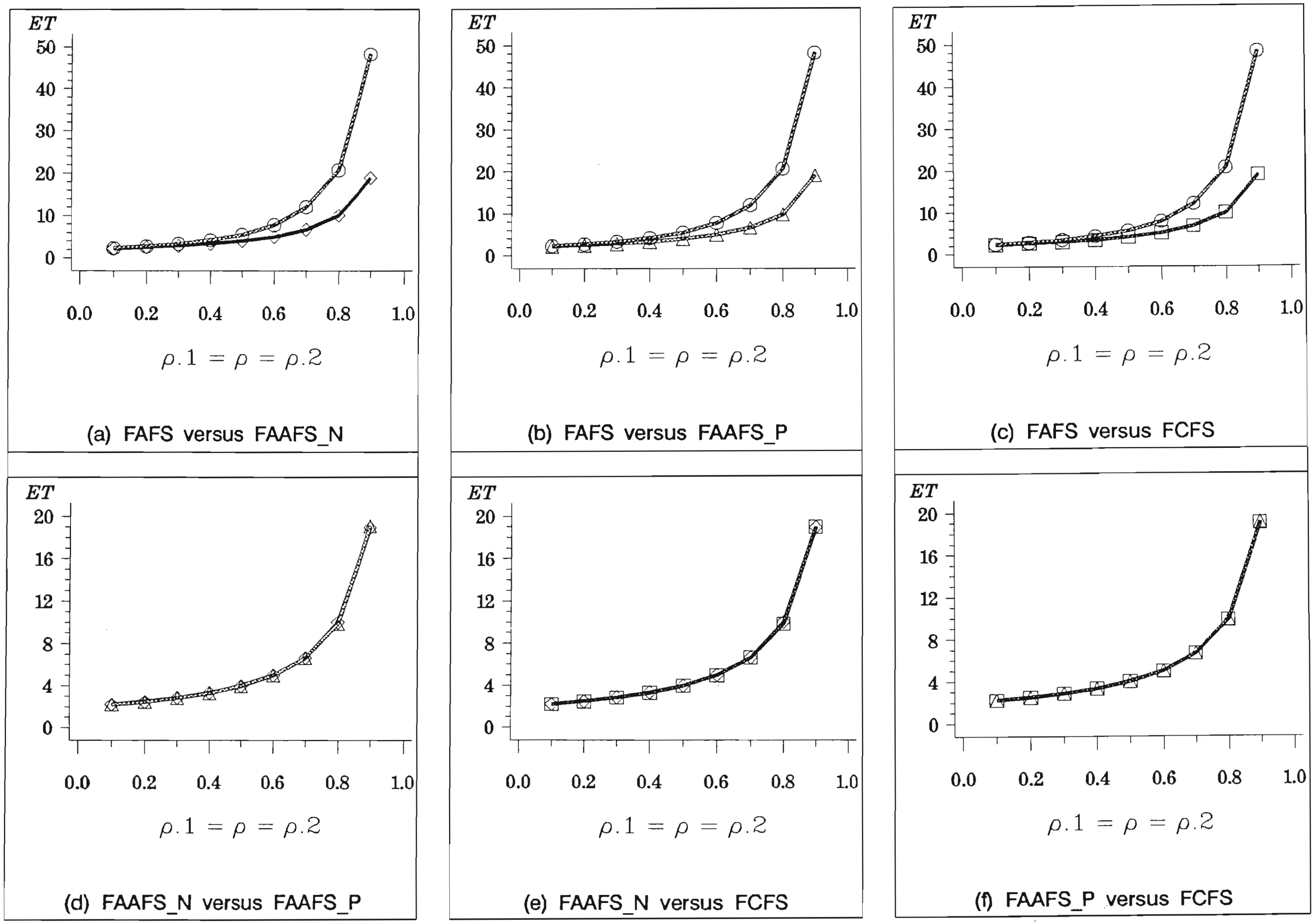
Mean waiting times for a symmetric circuit-switched starlike network

$\rho = \frac{\lambda}{\mu}$	Stage	$FAFS$	$\mathbf{E}T^{[i]}_{FCFS}$
	1	0.010	0.010
0.010	2	1.009	1.010
	Total	1.020	1.020
	1	0.446	0.304
0.168	2	1.388	1.202
	Total	1.835	1.506
	1	0.814	0.375
0.189	2	1.644	1.233
	Total	2.458	1.608

Table 4.3(c)

Mean waiting times for a symmetric circuit-switched starlike network

Special values for $FAFS$ discipline



(a) FAFS versus FAAFS_N (b) FAFS versus FAAFS_P (c) FAFS versus FCFS

(d) FAAFS_N versus FAAFS_P (e) FAAFS_N versus FCFS (f) FAAFS_P versus FCFS

Fig. 4.3 Simulation Study: Message-Switched Starlike Network

Key: ○ $FAFS$ ◇ $FAAFS_N$ △ $FAAFS_P$ □ $FCFS$

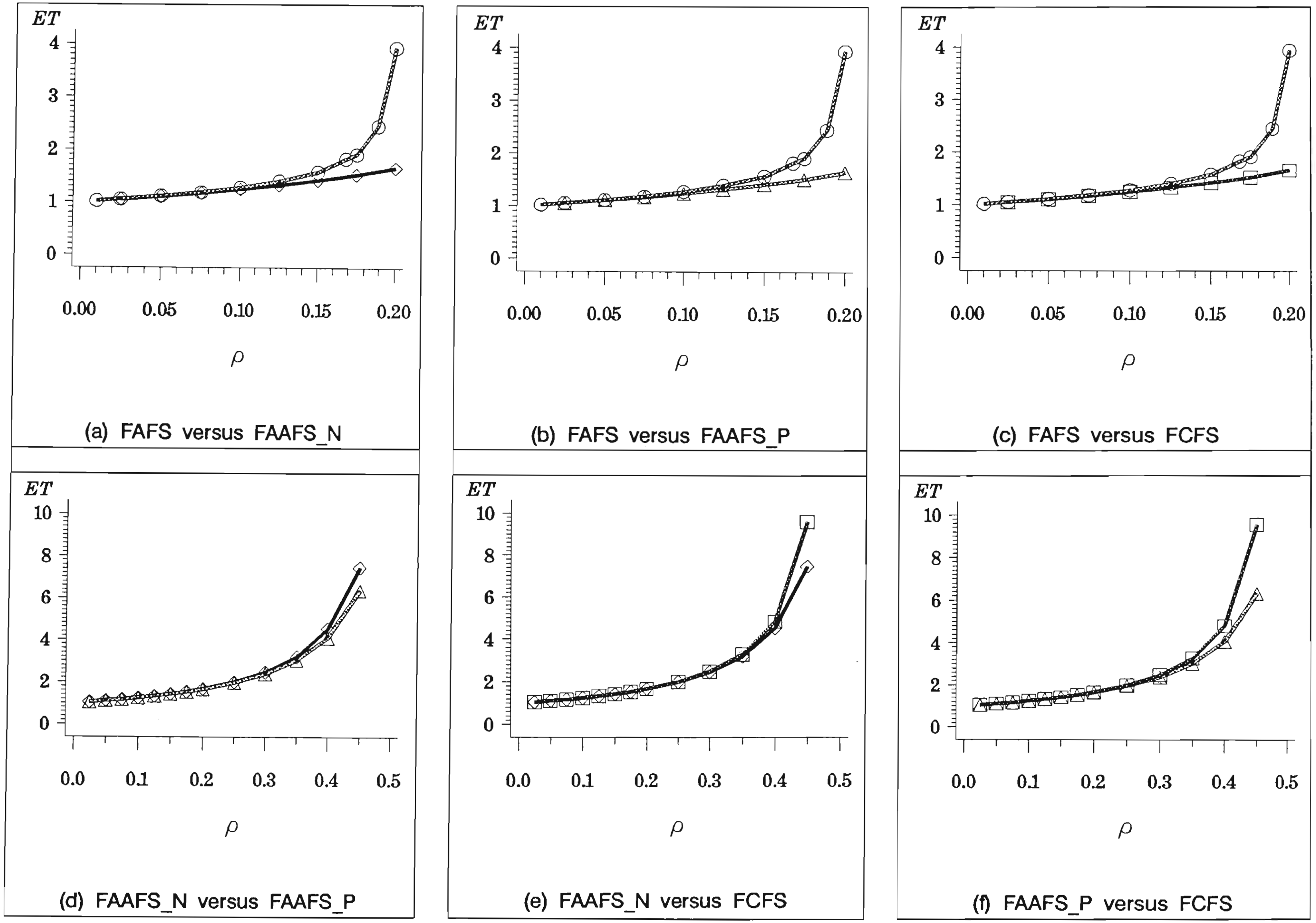


Fig. 4.4 Simulation Study: Circuit-Switched Starlike Network

Key: ○ $FAFS$ ◇ $FAAFS_N$ △ $FAAFS_P$ □ $FCFS$

We now summarise our findings (based on the tabular and graphical evidence) for the two network types.

(a) **Message-Switched Starlike Network**

$$\begin{array}{ccccccc}
ET^{[1]}_{FAFS} & \approx & ET^{[1]}_{FCFS} & \approx & ET^{[1]}_{FAAFS_N} & \approx & ET^{[1]}_{FAAFS_P} \\
ET^{[2]}_{FAFS} & > & ET^{[2]}_{FCFS} & \approx & ET^{[2]}_{FAAFS_N} & \approx & ET^{[2]}_{FAAFS_P} \\
\hline
ET_{FAFS} & > & ET_{FCFS} & \approx & ET_{FAAFS_N} & \approx & ET_{FAAFS_P} \\
\hline
\end{array} \qquad (4.2.1)$$

The difference $ET^{[2]}_{FAFS} - ET^{[2]}_{FCFS}$ (hence $ET_{FAFS} - ET_{FCFS}$) increases — more important, so does the ratio of these quantities — as the value of the traffic intensity increases. For values of $\rho^{[1]} = \rho^{[2]}$ near 1, ET_{FAFS} is 2 to $2\frac{1}{2}$ times greater than ET_{FCFS}. This represents a marked difference in the efficiency of the two protocols. A similar pattern is evident from Table 4.2(b).

(b) **Circuit-Switched Starlike Network**

$$\begin{array}{ccccccc}
ET^{[1]}_{FAFS} & \geq & ET^{[1]}_{FCFS} & \geq & ET^{[1]}_{FAAFS_N} & \geq & ET^{[1]}_{FAAFS_P} \\
ET^{[2]}_{FAFS} & > & ET^{[2]}_{FCFS} & \approx & ET^{[2]}_{FAAFS_N} & \approx & ET^{[2]}_{FAAFS_P} \\
\hline
ET_{FAFS} & > & ET_{FCFS} & \geq & ET_{FAAFS_N} & \geq & ET_{FAAFS_P} \\
\hline
\end{array} \qquad (4.2.2)$$

Predictably the differences outlined above become more marked with an increase in ρ. A similarity between the trends in (4.2.2) (above) and (4.2.1) is the discrepancy which exists between ET_{FAFS} and the sojourn times for the other three protocols. Some points of contrast are that

(i) the differences here emerge in the first and second stages whereas, in the message-switching scenario, they arise in the second phase

(ii) the difference between ET_{FAFS} and ET_{FCFS}, *at first glance*, appears to be less dramatic than in (a)

(iii) as a general rule, ET_{FCFS}, ET_{FAAFS_N} and ET_{FAAFS_P} do not coincide

Before "jumping to conclusions", though, we issue some words of caution. As regards (iii), we should mention that the data presented here, together with other results collected, imply reasonably slow convergence of the $FAAFS$ average waiting times to $\mathrm{E}T^{[i]}_{FCFS}$, especially when the traffic intensity is high. With several million more calls arriving, the distinction between ET_{FAAFS_N}, ET_{FAAFS_P} and ET_{FCFS} becomes less noticeable.[3]

Furthermore, as intimated in technical remark (d), stability of the various networks is obviously dependent on the traffic intensity. While the message-switched networks are "similar" to the queues of Part I in that the joint condition $\rho^{[1]} < 1$ and $\rho^{[2]} < 1$ is necessary and sufficient for the existence of a stationary regime, an analogous remark cannot be made for the circuit-switched starlike network. Heuristic and theoretical arguments suggest a uniform bound of $\rho < \frac{1}{2}$ for the $FAAFS$

[3]Different rates of convergence (as well as random errors) are also responsible for the minor discrepancies reflected in the data.

and $FCFS$ protocols, and a mere $3 - \sqrt{8} \approx 0.1716$ (see (6.2.11)) as the corresponding bound for the $FAFS$ discipline (with $N \to \infty$). (It is clear that values larger than $3 - \sqrt{8}$ will suffice with a finite number of channels. In our case, with $N = 200$, the network appears to remain stable even for $\rho = 0.2$.)

Our discussion in the above paragraph serves a dual purpose. In the first instance, it highlights the fact that in (the point of contrast) (ii) the relative values of the traffic intensity in the message-switching and circuit-switching cases should not be ignored. Secondly, and perhaps more significantly, the efficiency of the (circuit-switched) $FAFS$ system is already in question because of its comparatively limited region of stability.

It would appear that our conclusions (before simulation) on the relative efficiency of the various disciplines in Table 4.1 are, for the most part, justified (but we should always bear in mind that the simulation study was restricted to the exponential case). Subsequent investigations on the (sample) variances and covariances of the sojourn times (obtained by using a slightly adapted form of the program given in Appendix D) indicate, though, that the $FCFS$ discipline is not as "unfair" as was initially believed. On the other hand, recent comments might suggest that the $FAAFS$ circuit-switched protocols are slightly more efficient than $FCFS$.

As a final word on the matter, we propose that the preemptive $FAAFS$ protocol (though perhaps not always practical or feasible) is optimal.

4.3 FURTHER NOTATION AND TERMINOLOGY

In the case of this chapter and later chapters, there will be theoretical motivation for extending or reducing the marks of the flows considered in §§ 4.1 and 4.2. For instance, $\bar{\boldsymbol{\omega}} = \{\bar{\omega}_i, i \in \mathbb{N}_N\}$ is a realisation of the family $\boldsymbol{\eta}$ of extended RMPPs having extended marks $(\boldsymbol{S}, d, \boldsymbol{w})$ (or (S, d, W), depending on the situation).

Of related interest is the idea of an *extended message* $\bar{y} = [t; (\boldsymbol{S}, d, \boldsymbol{w})]$ or call $\bar{y} = [t; (S, d, W)]$. Individual components of $\bar{y}$ are accessed much as for y and $\hat{y}$. For example

$$\bar{y} = \left[\hat{t}; \left(\left(\widehat{S}^{[1]}, \widehat{S}^{[2]}\right), \hat{d}, \left(\widehat{w}^{[1]}, \widehat{W}\right)\right)\right] \implies$$

$$\begin{aligned}
t(\bar{y}) &= \hat{t} & & & & \\
\boldsymbol{S}(\bar{y}) &= \left(\widehat{S}^{[1]}, \widehat{S}^{[2]}\right) & S^{[1]}(\bar{y}) &= \widehat{S}^{[1]} & S^{[2]}(\bar{y}) &= \widehat{S}^{[2]} \\
d(\bar{y}) &= \hat{d} & & & & \\
\boldsymbol{w}(\bar{y}) &= \left(\widehat{w}^{[1]}, \widehat{W}\right) & w^{[1]}(\bar{y}) &= \widehat{w}^{[1]} & W(\bar{y}) &= \widehat{W}
\end{aligned}$$

Definition 4.3.1 (Source Predecessors and Successors)
Let $y = [t; (\boldsymbol{S}, s, d, \boldsymbol{w})]$ and $y' = [t'; (\boldsymbol{S}', s, d', \boldsymbol{w}')]$ be customers with $t' < t$ and $t' = \max[t(\tilde{y}) : t(\tilde{y}) < t \ \wedge \ s(\tilde{y}) = s]$. Then y' is said to be the *(immediate) source predecessor* of y, and y is referred to as the *(immediate) source successor* of y'.

More especially, for the class $\mathcal{Y}_i = \{y \in \omega_i\}$ (of messages (or calls) with common source $i \in \mathbb{N}_N$), we may define a function $p_s: \mathcal{Y}_i \to \mathcal{Y}_i$ given by

$$p_s^n(y) = \begin{cases} p_s\left[p_s^{n-1}(y)\right] & n \in \mathbb{Z}_+ \\ y & n = 0 \\ p_s^{-1}\left[p_s^{n+1}(y)\right] & n \in \mathbb{Z}_- \end{cases}$$

Thus $p_s(y) = y'$ and $p_s^{-1}(y') = y$.
It is convenient to refer to $p_s^n(y)$ $(n \in \mathbb{Z}_+)$ as the nth *source predecessor* of y and $p_s^n(y)$ $(n \in \mathbb{Z}_-, m = -n)$ as the mth *source successor* of y. (When $n = 1$ or $m = 1$, the word *first* may be suppressed.)

Definition 4.3.2 ((*FAFS*) Destination Predecessors and Successors) Let $y = [t; (\boldsymbol{S}, s, d, \boldsymbol{w})]$ and $y'' = [t''; (\boldsymbol{S}'', s'', d, \boldsymbol{w}'')]$ be customers with $t'' < t$ and $t'' = \max[t(\tilde{y}) : t(\tilde{y}) < t \wedge d(\tilde{y}) = d]$. Then y'' is said to be the *(immediate) destination predecessor* of y, and y is referred to as the *(immediate) destination successor* of y''.

More especially, for the class $\mathcal{Y}'_j = \{y \in \omega'_j\}$ (of messages (or calls) with common destination $j \in \mathbb{N}_N$), we may define a function $p_d : \mathcal{Y}'_j \to \mathcal{Y}'_j$ given by

$$p_d^n(y) = \begin{cases} p_d\left[p_d^{n-1}(y)\right] & n \in \mathbb{Z}_+ \\ y & n = 0 \\ p_d^{-1}\left[p_d^{n+1}(y)\right] & n \in \mathbb{Z}_- \end{cases}$$

Thus $p_d(y) = y''$ and $p_d^{-1}(y'') = y$.
It is convenient to refer to $p_d^n(y)$ $(n \in \mathbb{Z}_+)$ as the nth *destination predecessor* of y and $p_d^n(y)$ $(n \in \mathbb{Z}_-, m = -n)$ as the mth *destination successor* of y. (Again when $n = 1$ or $m = 1$, the word *first* may be suppressed.)

Remark 4.3.1 (Predecessors and Successors)
Predecessors and successors are uniquely determined by Definitions 4.3.1 and 4.3.2 (certainly in the $FAFS$ starlike network).

Definition 4.3.3 (Indices)

Suppose $y \in \omega_i \cap \omega'_j$ for a network initiated at (finite) time t_0. The *total index* of y, $n(y)$ determines the relative position at which it arrived to the network, i.e.

$$\begin{cases} n(y) = 1 & \Leftrightarrow \quad t(y) = \min[t(\tilde{y}) > t_0 : \tilde{y} \in \boldsymbol{\omega}] \\ n(y) = k & \Leftrightarrow \quad t(y) = \min[t(\tilde{y}) > t(\bar{y}) : \tilde{y} \in \boldsymbol{\omega}, n(\bar{y}) = k-1] \quad k \in \overline{\mathbb{N}}_1 \end{cases}$$

(if these exist). In terms of earlier notation, $n(y) = k$ iff $y = y_k$ or $t(y) = t_k$.

The *source* and *destination indices* of y, $n_s(y)$ and $n_d(y)$, similarly determine the position of y in $\omega_{s(y)} = \omega_i$ and $\omega'_{d(y)} = \omega'_j$ respectively.

Remark 4.3.2

Let y be as in Definition 4.3.3. It is obvious that for $p_s^k(y)$ and $p_d^l(y)$ $(k, l \in \mathbb{N})$ to be defined, we require $k < n_s(y)$ and $l < n_d(y)$ respectively.

The predecessors and successors of Definition 4.3.2 are less meaningful in the context of the $FAAFS$ discipline. We therefore define an additional concept which has greatest relevance for a nonpreemptive $FAAFS$ circuit-switched network.

Definition 4.3.4 ((*FAAFS*) Delayers)

A call $y' \in \omega_i \cup \omega_j'$ is said to be the *delayer* of $y \in \omega_i \cap \omega_j'$, written $y' = \Delta(y)$, iff $t(y') + T(y') = t(y) + W(y)$.

Remark 4.3.3 (Delayers)

For a network initiated at time t_0, $\Delta(y)$ need not exist, but if it does, then

(a) $\Delta(y) = p_s(y)$ iff $w^{[1]}(y) > 0$ *and* $w^{[2]}(y) = 0$

(b) $\Delta(y) = p_d^n(y)$ for some $n \in \mathbb{Z} \setminus \{0\}$ if $w^{[2]}(y) > 0$

On the other hand, $\Delta(y)$ is not defined if both $w^{[1]}(y)$ and $w^{[2]}(y)$ are zero.

It is convenient to define another quantity (again, of particular relevance in a nonpreemptive *FAAFS* circuit-switched network.)

Definition 4.3.5 ((*FAAFS*) Quasidelayers)

If, for a call y, $\Delta(y)$ exists, then the *quasidelayer* of y, denoted by $\Delta_Q(y)$ is given by

$$\Delta_Q(y) = \begin{cases} \Delta(y) & t(\Delta(y)) < t \\ p_s(y) & t(\Delta(y)) > t \end{cases} \tag{4.3.1}$$

Remark 4.3.4 (Delayers and Quasidelayers)

(a) In contrast to (*FAFS*) predecessors, $\Delta(y)$ may arrive to the *FAAFS* network *after* y. On the other hand, $\Delta_Q(y)$ must, by definition, have arrived *before* y.

(b) As usual, $\Delta^n(y)$ and $\Delta_Q^n(y)$ $(n \in \mathbb{Z}_\oplus)$ may be considered.

CHAPTER 5

THE EXTENDED LINDLEY EQUATION

5.1 INTRODUCTION

The preliminaries for this chapter were mostly dealt with in Chapter 4, but relevant details also appear elsewhere in Part II. We consider a message-switched $FAFS$ discipline on a symmetric starlike network. In the process of this investigation, a Lindley-Loynes equation, which extends that of Chapter 1, will emerge.

In addition to the assumptions that each of the N arrival streams is Poisson (rate λ) and that $S^{[i]} \sim F^{[i]}$ $(i = 1, 2)$, $\boldsymbol{S} \sim F$, we will require:

$$\mathrm{E}e^{aS^{[1]}} < \infty \text{ for some } a > 0 \qquad \mathrm{E}S^{[2]} < \infty \tag{5.1.1a}$$

$$\rho^{[1]} = \lambda \cdot \mathrm{E}S^{[1]} < 1 \qquad \rho^{[2]} = \lambda \cdot \mathrm{E}S^{[2]} < 1 \tag{5.1.1b}$$

where (5.1.1b) gives bounds which are uniform in N.

Not only will the RMPP families $\{\xi_i\}$ (or $\{\xi_j'\}$) and $\{\eta_i\}$ (or $\{\eta_j'\}$) be of interest; so too will the truncated flows $\boldsymbol{\xi}(\vec{t_0}) = \left\{\xi_i(\vec{t_0}) = \xi_i I_{[t_0,\infty)}, i \in \mathbb{N}_N\right\}$ (a restriction of $\boldsymbol{\xi}$ to the interval $[t_0, \infty)$). An initial workload W^0 is incorporated thus: $\boldsymbol{\eta}(\vec{t_0}, W^0) = \left\{\eta_i(\vec{t_0}, W^0), i \in \mathbb{N}_n\right\}$. The number of nodes in each stage may be indicated by a bracketed superscript (e.g., $\eta_i^{(N)}$).

5.2 FORMULATION OF THE GOVERNING EQUATIONS

It is obvious that for a message $y = [t;(\boldsymbol{S}, d, \boldsymbol{w})]$ with $p_s(y) = y'$ and $p_d(y) = y''$ (to simplify the notation)

$$w^{[1]}(y) \quad = \quad \left[w^{[1]}(y') + u^{[1]}(y')\right]^+ \tag{5.2.1}$$

where $u^{[1]}(y') = S^{[1]}(y') - \tau(y')$, $\tau(y') = t - t(y')$. After all, (5.2.1) is (1.1.1) (or (3.5.2)) in a more general form.

A similar relationship may be derived for $w^{[2]}(y)$:

$$w^{[2]}(y) \quad = \quad \left[w^{[2]}(y'') + S^{[2]}(y'') - [t^{[2]}(y) - t^{[2]}(y'')]\right]^+$$

where $t^{[2]}(\cdot) = t(\cdot) + T^{[1]}(\cdot)$, the time of arrival to the second stage (hence the time of departure from stage 1).

It follows that

$$\begin{aligned}
& W(y) \\
= \ & T^{[1]}(y) + w^{[2]}(y) \\
= \ & T^{[1]}(y) + \max\left[0, w^{[2]}(y'') + S^{[2]}(y'') - [(t(y) + T^{[1]}(y)) - (t(y'') + T^{[1]}(y''))]\right] \\
= \ & \max\left[T^{[1]}(y), w^{[2]}(y'') + S^{[2]}(y'') + T^{[1]}(y'') - [t(y) - t(y'')]\right] \\
= \ & \max\left[w^{[1]}(y) + S^{[1]}(y), W(y'') + S^{[2]}(y'') - [t(y) - t(y'')]\right]
\end{aligned}$$

which is also reasonable from an intuitive viewpoint.

Our conclusions may be summarised by:

$$w^{[1]}(y) = \left[w^{[1]}(p_s(y)) + S^{[1]}(p_s(y)) - [t(y) - t(p_s(y))]\right]^+ \qquad (5.2.2a)$$

$$W(y) = \max\left[T^{[1]}(y), W(p_d(y)) + u(p_d(y))\right] \qquad (5.2.2b)$$

where $u(p_d(y)) = S^{[2]}(p_d(y)) - [t(y) - t(p_d(y))]$.

Since each of the N stage-1 queues is of the standard $M/GI/1/\infty/FCFS$ type, and (as indicated before) (5.2.2a) is (1.1.1) in "disguised form", it will be of greater interest to analyse the equation (5.2.2b). This analysis will dominate the rest of the chapter, especially sections 5.3 and 5.5.

5.3 EXISTENCE AND UNIQUENESS OF A STATIONARY REGIME

If condition (5.1.1b) is satisfied, then we may conclude that

(a) there exists a unique stationary family $\boldsymbol{\eta}$ which provides the unique stationary solution of the system of equations (5.2.2a) and (5.2.2b) associated with $\boldsymbol{\xi}$

(b) for any initial condition W^0

$$\eta_i(\vec{t_0}, W^0) \xrightarrow[t_0 \to -\infty]{d} \eta_i \qquad (5.3.1)$$

As in Baccelli *et al.* [4], we state and prove a different result (based on the language of discrete time) on the existence and uniqueness of a stationary regime. A number of preliminary results are needed before we are able to proceed to the main theorem of this section.

Lemma 5.3.1

Let $\langle x_n \rangle$ be a stationary ergodic sequence. Then

(a) $\mathrm{E}|x_n| < \infty \Rightarrow \dfrac{x_n}{n} \underset{\infty}{\overset{n}{\rightarrow}} 0$ a.e.

(b) $\langle x_n \rangle$ an IID sequence and

$$P\left[\sup_{n \in \mathbb{N}} \frac{x_n}{n} < \infty\right] > 0$$

$\Longrightarrow \mathrm{E}x_n^+ < \infty$

PROOF

(a) Let $n \in \mathbb{N}$. Note that

$$\begin{aligned}
\sum_{k=1}^{n}(x_k - x_{k-1}) &= \sum_{k=1}^{n} x_k - \sum_{k=0}^{n-1} x_k \\
&= x_n - x_0 \\
\Rightarrow \qquad \frac{x_n}{n} - \frac{x_0}{n} &= \tfrac{1}{n}\sum_{k=1}^{n}(x_k - x_{k-1})
\end{aligned}$$

Bearing in mind that $\mathrm{E}|x_n| < \infty$, we may now apply the strong law of large numbers for stationary ergodic sequences (Theorem 0.2.2).

$$\begin{aligned}
\lim_{n\to\infty}\left(\frac{x_n}{n}-\frac{x_0}{n}\right) &= \lim_{n\to\infty}\tfrac{1}{n}\sum_{k=1}^{n}(x_k-x_{k-1}) \\
&= \mathrm{E}(x_1-x_0) \\
&= 0 \qquad\qquad \text{since } \mathrm{E}x_1=\mathrm{E}x_0 \\
\Longrightarrow \qquad \lim_{n\to\infty}\frac{x_n}{n} &= \lim_{n\to\infty}\frac{x_0}{n} \\
&= 0 \text{ a.e.}
\end{aligned}$$

(b) Let $\langle x_n\rangle$ be an IID sequence and $P\left[\sup_{n\in\mathbb{N}}\frac{x_n}{n}<\infty\right]>0$ but suppose, to the contrary, that $\int_0^\infty P[x_1\geq x]dx=\mathrm{E}x_1^+=\mathrm{E}x_n^+=\infty$. Further let $M\geq 1$, a finite number. With $q=\int_0^M P[x_1\geq y]dy$:

$$\begin{aligned}
\mathrm{E}x_1^+ &= \int_0^\infty P[x_1\geq y]dy \\
&= \sum_{n=0}^{\infty}\int_{nM}^{(n+1)M} P[x_1\geq y]dy \\
&= \sum_{n=1}^{\infty}\int_{nM}^{(n+1)M} P[x_1\geq y]dy+\int_0^M P[x_1\geq y]dy \\
&= \sum_{n=1}^{\infty}M\int_n^{n+1}P[x_1\geq Mx]dx+q \qquad\qquad y=Mx \\
&\leq \sum_{n=1}^{\infty}M[(n+1)-n]\max_{x\in[n,n+1]}P[x_1\geq Mx]+q \quad (=) \\
&\leq M\sum_{n=1}^{\infty}P[x_1\geq Mn]+q
\end{aligned}$$

Since $M \geq 1$ and is finite, $q < \infty$ and $\mathrm{E}x_1^+ = \infty$, we have

$$\sum_{n=1}^{\infty} P[x_1 \geq nM] = \infty \tag{5.3.2}$$

Next observe that $\forall\, M \geq 1$

$$\begin{aligned}
0 &\leq P\left[\sup_{n\in\mathbb{N}} \frac{x_n}{n} < M\right] \\
&\leq P\left[\bigcap_{n=1}^{\infty}\left[\frac{x_n}{n} < M\right]\right] && (=) \\
&\leq \prod_{n=1}^{\infty} P[x_n < nM] && (=) \quad \text{by the independence property} \\
&\leq \prod_{n=1}^{\infty} P[x_1 < nM] && (=) \quad \text{since identically distributed} \\
&\leq \prod_{n=1}^{\infty} [1 - P[x_1 \geq nM]] \\
&\leq \prod_{n=1}^{\infty} e^{-P[x_1 \geq nM]} && (=) \quad \text{since} \begin{cases} 1 - x < e^{-x} & x \in (0,1] \\ 1 - x = e^{-x} & x = 0 \end{cases} \\
&\leq e^{-\sum_{n=1}^{\infty} P[x_1 \geq nM]}
\end{aligned}$$

so from (5.3.2) we conclude that $P\left[\sup_{n\in\mathbb{N}} \frac{x_n}{n} < \infty\right] = 0$ which contradicts the assumption $P\left[\sup_{n\in\mathbb{N}} \frac{x_n}{n} < \infty\right] > 0$. Hence $\mathrm{E}x_1^+ < \infty$.

$\square$

Our attention now turns to the "discretized" form of (5.2.2b). Some of the ideas, terminology and notation (with adaptations) filter through from Part I.

The (prestationary) extended Lindley-Loynes equation assumes the following form:

$$W_{n+1} \quad = \quad \max\,[T^{[1]}_{n+1}, W_n + u_n] \tag{5.3.3}$$

where $\left\langle (u_n, T^{[1]}_{n+1}), n \in \mathbb{Z} \right\rangle$ is a stationary ergodic sequence with $(u_n, T^{[1]}_{n+1})$ from $\mathbb{R} \times \mathbb{R}_{\oplus}$. When $T^{[1]}_n = 0 \ \forall n \in \mathbb{Z}$ (5.3.3) reduces to the familiar Lindley equation (1.1.1). It is for this reason particularly that the term *extended* has been chosen.

Lemma 5.3.2

If $\langle W_n, n \in \mathbb{Z} \rangle$ satisfies (5.3.3), then subject to

$$W_{n-r}(n-r, W^0) = W^0$$

we have

$$W_m(n-r, W^0) = \max\left[T^{[1]}_m, \max_{n-r<k<m}\left(T^{[1]}_k + \sum_{j=k}^{m-1} u_j\right), W^0 + \sum_{j=n-r}^{m-1} u_j\right] \quad m > n-r \tag{5.3.4}$$

PROOF

We proceed by induction on m:

$$\begin{aligned}
\text{LHS} &= W_{n-r+1}(n-r, W^0) \\
&= \max\,[T^{[1]}_{n-r+1}, W_{n-r}(n-r, W^0) + u_{n-r}] \\
&= \max\,[T^{[1]}_{n-r+1}, W^0 + u_{n-r}] \\
&= \max\left[T^{[1]}_{n-r+1}, \max_{n-r<k<n-r+1}\left(T^{[1]}_k + \sum_{j=k}^{n-r} u_j\right), W^0 + \sum_{j=n-r}^{n-r} u_j\right] \\
&= \text{RHS}
\end{aligned}$$

where we used the fact that $k \in (n-r, n-r+1)_{\mathbb{Z}} = \emptyset$

Suppose that (5.3.4) is valid for some $m > n-r$. Then:

$$\begin{aligned}
& W_{m+1}(n-r, W^0) \\
=\ & \max[T^{[1]}_{m+1}, W_m(n-r, W^0) + u_m) \\
=\ & \max\left\{T^{[1]}_{m+1}, \max\left[T^{[1]}_m, \max_{n-r<k<m}\left(T^{[1]}_k + \sum_{j=k}^{m-1} u_j\right), W^0 + \sum_{j=n-r}^{m-1} u_j\right] + u_m\right\} \\
=\ & \max\left\{T^{[1]}_{m+1}, \max\left[T^{[1]}_m + u_m, \max_{n-r<k<m}\left(T^{[1]}_k + \sum_{j=k}^{m} u_j\right), W^0 + \sum_{j=n-r}^{m} u_j\right]\right\} \\
=\ & \max\left\{T^{[1]}_{m+1}, \max\left[\max_{n-r<k<m+1}\left(T^{[1]}_k + \sum_{j=k}^{m} u_j\right), W^0 + \sum_{j=n-r}^{m} u_j\right]\right\}
\end{aligned}$$

as required.

□

Corollary 5.3.2*a*

(a) $W_m(n-r, W^0)$ for fixed m, n and r $(m \geq n-r)$ is a nondecreasing function of W^0

(b) $W_m(n-r, 0)$ for fixed m and n $(m > n-r)$ is a nondecreasing function of r

<u>PROOF</u>

(a) Write

$$z_m = \max\left[T_m^{[1]}, \max_{n-r<k<m}\left(T_k^{[1]} + \sum_{j=k}^{m-1} u_j\right)\right]$$

Then

$$W_m(n-r, W^0) = \max\left[z_m, W^0 + \sum_{j=n-r}^{m-1} u_j\right]$$

Suppose now that $W_{n-r}(n-r, \overline{W}^0) = \overline{W}^0 > W^0 = W_{n-r}(n-r, W^0)$. We prove the result for all possible cases.

Case 1 : $$W^0 + \sum_{j=n-r}^{m-1} u_j < \overline{W}^0 + \sum_{j=n-r}^{m-1} u_j \leq z_m$$

$$\begin{aligned} W_m(n-r, \overline{W}^0) &= z_m \\ &= W_m(n-r, W^0) \end{aligned}$$

$$Case \quad 2: \qquad z_m \leq W^0 + \sum_{j=n-r}^{m-1} u_j < \overline{W}^0 + \sum_{j=n-r}^{m-1} u_j$$

$$\begin{aligned} W_m(n-r, \overline{W}^0) &= \overline{W}^0 + \sum_{j=n-r}^{m-1} u_j \\ &> W^0 + \sum_{j=n-r}^{m-1} u_j \qquad (=) \\ &> W_m(n-r, W^0) \end{aligned}$$

$$Case \quad 3: \qquad W^0 + \sum_{j=n-r}^{m-1} u_j < z_m \leq \overline{W}^0 + \sum_{j=n-r}^{m-1} u_j$$

$$\begin{aligned} W_m(n-r, \overline{W}^0) &= \overline{W}^0 + \sum_{j=n-r}^{m-1} u_j \\ &\geq z_m \qquad (=) \\ &\geq W_m(n-r, W^0) \end{aligned}$$

Thus $W_m(n-r, \overline{W}^0) \geq W_m(n-r, W^0)$, as claimed.

(b) Suppose $\overline{r} > r$ and recall that $\forall\, m > n-r$

$$W_m(n-r, 0) = \max \left[T_m^{[1]}, \max_{n-r<k<m} \left(T_k^{[1]} + \sum_{j=k}^{m-1} u_j \right), \sum_{j=n-r}^{m-1} u_j \right]$$

Now

$$(n-r, m)_{\mathbb{Z}} \subset (n-\overline{r}, m)_{\mathbb{Z}}$$

Thus

$$
\begin{aligned}
&\max_{n-\bar{r}<k<m}\left(T_k^{[1]}+\sum_{j=k}^{m-1}u_j\right)\\
=\;&\max\left[\max_{\substack{k\in\\(n-r,m)}}\left(T_k^{[1]}+\sum_{j=k}^{m-1}u_j\right),T_{n-r}^{[1]}+\sum_{j=n-r}^{m-1}u_j,\;\max_{\substack{k\in\\(n-\bar{r},n-r)}}\left(T_k^{[1]}+\sum_{j=k}^{m-1}u_j\right)\right]\\
\geq\;&\max\left[\max_{n-r<k<m}\left(T_k^{[1]}+\sum_{j=k}^{m-1}u_j\right),\sum_{j=n-r}^{m-1}u_j\right]
\end{aligned}
$$

where we use the fact that $T_{n-r}^{[1]}+\sum_{j=n-r}^{m-1}u_j \geq \sum_{j=n-r}^{m-1}u_j$ since $T_{n-r}^{[1]}\geq 0$.

Hence the result.

□

Lemma 5.3.3

Fix $c \in \mathbb{R}$. Then $\forall\, a \in \mathbb{R}$ and (continuous) random variables X and Y, the following holds

(a) $P[X+Y<c] \leq P[X<c-a]+P[Y<a]$

(b) $P[X+Y>c] \leq P[X>c-a]+P[Y>a]$

PROOF

(a)

$$\begin{aligned}
\neg\,(\{X<c-a\}\vee\{Y<a\}) &\Rightarrow (\{X\geq c-a\}\wedge\{Y\geq a\}) \\
&\Rightarrow (X+Y\geq c) \\
&\Rightarrow \neg(X+Y<c)
\end{aligned}$$

$$\Longrightarrow \quad [X+Y<c] \subseteq [\{X<c-a\}\cup\{Y<a\}]$$

$$\begin{aligned}
\Rightarrow \quad P[X+Y<c] &\leq P[\{X<c-a\}\cup\{Y<a\}] \\
&\leq P[X<c-a]+P[Y<a]
\end{aligned}$$

(b) In a similar way to (a)

$$\begin{aligned}
\neg\,(\{X>c-a\}\vee\{Y>a\}) &\Rightarrow (X+Y\leq c) \\
\Rightarrow \quad [X+Y>c] &\subseteq [\{X>c-a\}\cup\{Y>a\}] \\
\Rightarrow \quad P[X+Y>c] &\leq P[X>c-a]+P[Y>a]
\end{aligned}$$

$\square$

Theorem 5.3.1

Consider the stationary ergodic sequence $\left\langle (u_n, T_{n+1}^{[1]}) \right\rangle$ associated with (5.3.3).

(a) Suppose the sequence satisfies

$$\mathrm{E}u_0 < 0 \quad \text{and} \quad \mathrm{E}T_0^{[1]} < \infty \tag{5.3.5}$$

Then

(i) there exists a stationary solution $\langle \underline{m}_n, n \in \mathbb{Z} \rangle$ of (5.3.3)

(ii) $\langle \underline{m}_n \rangle$ is the unique stationary solution, and $\forall\, W^0 \geq 0$ and any finite set $J \subset \mathbb{Z}$

$$\lim_{n^0 \to -\infty} \sup\left[L(\langle \underline{m}_n, n \in J \rangle), L(\langle W_n(n^0, W^0), n \in J \rangle)\right] = 0 \tag{5.3.6}$$

(b) Assume now that $\left\langle T_n^{[1]}, n \in \mathbb{Z} \right\rangle$ is an IID sequence, distributed independently of $\langle u_n, n \in \mathbb{Z} \rangle$, and that a (finite) stationary solution sequence for (5.3.3), $\langle \underline{W}_n, n \in \mathbb{Z} \rangle$, exists. Then

(i) $\mathrm{E}u_0 \leq 0$, and if $\langle u_n \rangle$ is an IID sequence of nonzero random variables, we have $\mathrm{E}u_0 < 0$

(ii) under the original conditions $\mathrm{E}T_0^{[1]} < \infty$

PROOF

From (5.3.4) we obtain

$$\begin{aligned} W_n(n-r,W^0) &= \max\left[T_n^{[1]}, \max_{n-r<k<n}\left(T_k^{[1]}+\sum_{j=k}^{n-1}u_j\right), W^0+\sum_{j=n-r}^{n-1}u_j\right] \\ &= \max\left[T_n^{[1]}, \max_{k\in\mathbb{N}_{r-1}}\left(T_{n-k}^{[1]}+\sum_{j=n-k}^{n-1}u_j\right), W^0+\sum_{j=n-r}^{n-1}u_j\right] \end{aligned}$$

Next

$$\begin{aligned} &\lim_{r\to\infty}\left[W^0+\sum_{j=n-r}^{n-1}u_j\right] \\ &= W^0+\lim_{r\to\infty} r\left(\frac{1}{r}\sum_{j=1}^{r}u_{n-j}\right) \\ &< 0 && (\leq) \text{ a.e.} \\ &< T_n^{[1]} && T_n^{[1]}\in\mathbb{R}_\oplus \qquad (5.3.7) \end{aligned}$$

(or, more precisely, $\lim_{r\to\infty} P\left[W^0+\sum_{j=n-r}^{n-1}u_j>0\right]=0$) where we use (5.3.5) in conjunction with the ergodic theorem (Theorem 0.2.2) and the fact that $W^0\in\mathbb{R}_\oplus$ (i.e., finite). As a result

$$\begin{aligned} \lim_{r\to\infty} W_n(n-r,W^0) &= \lim_{r\to\infty}\max\left[T_n^{[1]}, \max_{k\in\mathbb{N}_{r-1}}\left(T_{n-k}^{[1]}+\sum_{j=n-k}^{n-1}u_j\right)\right] \\ &= \max\left[T_n^{[1]}, \sup_{k\in\mathbb{N}}\left(T_{n-k}^{[1]}+\sum_{j=n-k}^{n-1}u_j\right)\right] \end{aligned}$$

(a) (i) The stationarity of the above limit is a consequence of the stationarity of $\left\langle (u_n, T^{[1]}_{n+1}) \right\rangle$. We next observe that, a.e.

$$\lim_{k\to\infty}\left(T^{[1]}_{n-k}+\sum_{j=1}^{k}u_{n-j}\right) \;=\; \lim_{k\to\infty}\left[k\left(\frac{T^{[1]}_{n-k}}{k}+\frac{1}{k}\sum_{j=1}^{k}u_{n-j}\right)\right] \;<\; 0$$

by applying Lemma 5.3.1
(since we have $T^{[1]}_n \in \mathbb{R}_\oplus \Rightarrow \mathrm{E}|T^{[1]}_n| = \mathrm{E}T^{[1]}_n < \infty$ by (5.3.5)) and the strong law in a similar way to that which leads to (5.3.7).

Thus $\sup\limits_{k\in\mathbb{N}}\left(T^{[1]}_{n-k}+\sum\limits_{j=n-k}^{n-1}u_j\right)$ is attained for finite k w.p.1 and is, therefore, itself finite. Hence

$$\underline{m}_n = \max\left[T^{[1]}_n, \sup_{k\in\mathbb{N}}\left(T^{[1]}_{n-k}+\sum_{j=n-k}^{n-1}u_j\right)\right] \tag{5.3.8}$$

yields an a.e. finite stationary solution.

(ii) Most of this part has already been proved. We merely note that some of what we have already concluded could be summarised as

$$P[W_n(n-r, W^0) > W_n(n-r, 0)] \underset{\infty}{\overset{r}{\to}} 0$$

(b) (i) We are given that an a.e. finite stationary solution $\langle \underline{W}_n \rangle$ exists. Now

$$\underline{m}_n = \max\left[T^{[1]}_n, \sup_{k\in\mathbb{N}}\left(T^{[1]}_{n-k}+\sum_{j=n-k}^{n-1}u_j\right)\right] \le \underline{W}_n \;\forall\, n\in\mathbb{Z} \text{ a.s.}$$

because it is the minimal solution (obtained as $\underline{m}_n = \lim\limits_{r\to\infty} W_n(n-r, 0)$).

Hence $\underline{m}_n$ must itself be a.e. finite, which implies that

$$\sup_{k\in\mathbb{N}}\left(T^{[1]}_{n-k}+\sum_{j=n-k}^{n-1}u_j\right) \text{ is a.e. finite.}$$

Suppose, to the contrary, that $\mathrm{E}u_0 > 0$. Then

$\lim_{k\to\infty}\sum_{j=1}^{k}u_{n-j} = \lim_{k\to\infty} k\left[\frac{1}{k}\sum_{j=1}^{k}u_{n-j}\right] = \infty$ which contradicts the fact that the supremum is finite. Thus we must have $\mathrm{E}u_0 \le 0$.

If we make the additional assumption that $\langle u_n\rangle$ is an IID sequence of nonzero r.v.s, then a suitable adaptation of Lindley [45] (especially pages 280–283) gives $\mathrm{E}u_0 < 0$.

(ii) Let $M \in \mathbb{N}$. We have noted that

$$\begin{aligned}
\underline{W}_n \geq \underline{m}_n &= \max\left[T_n^{[1]}, \sup_{k\in\mathbb{N}}\left(T_{n-k}^{[1]} + \sum_{j=n-k}^{n-1} u_j\right)\right] \\
\Rightarrow \qquad \underline{W}_n &\geq \sup_{k\in\mathbb{N}}\left(T_{n-k}^{[1]} + \sum_{j=n-k}^{n-1} u_j\right) \\
\Rightarrow \quad [\underline{W}_n < M] &\subseteq \left[\sup_{k\in\mathbb{N}}\left(T_{n-k}^{[1]} + \sum_{j=n-k}^{n-1} u_j\right) < M\right] \\
\Rightarrow P[\underline{W}_n < M] &\leq P\left[\sup_{k\in\mathbb{N}}\left(T_{n-k}^{[1]} + \sum_{j=n-k}^{n-1} u_j\right) - M < 0\right] \\
&\leq P\left[\sup_{k\geq M}\left(T_{n-k}^{[1]} + \sum_{j=n-k}^{n-1} u_j - M\right) < 0\right] \\
&\leq P\left[\sup_{k\geq M}\left(T_{n-k}^{[1]} + \sum_{j=n-k}^{n-1} u_j - k\right) < 0\right] \qquad (=) \\
&\leq P\left[\sup_{k\geq M}\left(\frac{T_{n-k}^{[1]}}{k} + \frac{1}{k}\sum_{j=n-k}^{n-1} u_j - 1\right) < 0\right]
\end{aligned} \tag{5.3.9}$$

In the third-last inequality we used the fact that

$$\begin{aligned}
\sup_{k\geq M} x_k &\leq \sup_{k\in\mathbb{N}} x_k \\
\Rightarrow \left[\sup_{k\geq M} x_k < 0\right] &\supseteq \left[\sup_{k\in\mathbb{N}} x_k < 0\right]
\end{aligned}$$

Thus

$$
\begin{aligned}
& P[\underline{W}_n < M] \\
\leq\ & P\left[\sup_{k\geq M}\left(\frac{T^{[1]}_{n-k}}{k} + \frac{1}{k}\sum_{j=n-k}^{n-1} u_j\right) < 1\right] && \text{from (5.3.9)} \\
\leq\ & P\left[\sup_{k\geq M}\frac{T^{[1]}_{n-k}}{k} + \inf_{k\geq M}\frac{1}{k}\sum_{j=n-k}^{n-1} u_j < 1\right] && \text{Lemma 0.2.1} \\
\leq\ & P\left[\sup_{k\geq M}\frac{T^{[1]}_{n-k}}{k} < 1-a\right] + P\left[\inf_{k\geq M}\frac{1}{k}\sum_{j=n-k}^{n-1} u_j < a\right] && \text{Lemma 5.3.3}
\end{aligned}
$$

where the last inequality holds $\forall a \in \mathbb{R}$, in particular for $a \ll \mathrm{E}u_n$. Continuing:

$$
\begin{aligned}
& 1 \\
=\ & P[\underline{W}_n < \infty] \\
=\ & \lim_{M\to\infty} P[\underline{W}_n < M] \\
\leq\ & \lim_{M\to\infty} P\left[\sup_{k\geq M}\frac{T^{[1]}_{n-k}}{k} < 1-a\right] + \lim_{M\to\infty} P\left[\inf_{k\geq M}\frac{1}{k}\sum_{j=n-k}^{n-1} u_j < a\right] \quad (=) \\
\leq\ & \lim_{M\to\infty} P\left[\sup_{k\geq M}\frac{T^{[1]}_{n-k}}{k} < 1-a\right] + 0
\end{aligned}
$$

where we use the Strong Law of Large Numbers in the final step.

$$\text{Hence} \quad \lim_{M\to\infty} P\left[\sup_{k\geq M}\frac{T^{[1]}_{n-k}}{k} < 1-a\right] = 1$$

$$\Longrightarrow \forall\,\epsilon>0\ \exists M_0 \text{ s.t. } M>M_0 \Rightarrow P\left[\sup_{k\geq M}\frac{T^{[1]}_{n-k}}{k} > 1-a\right] < \frac{\epsilon}{2} \tag{5.3.10}$$

We may also find sufficiently large $K = 1-a$ such that

$$\forall\,\epsilon>0 \quad P\left[\sup_{k\leq M}\frac{T^{[1]}_{n-k}}{k} > K\right] < \frac{\epsilon}{2} \tag{5.3.11}$$

(for some finite $M > M_0$). Moreover

$$\begin{aligned} P\left[\sup_{k\in\mathbb{N}}\frac{T^{[1]}_{n-k}}{k} > 1-a\right] \quad \leq \quad & P\left[\sup_{k\in\mathbb{N}_M}\frac{T^{[1]}_{n-k}}{k} > 1-a\right] + \\ & P\left[\sup_{k\in\overline{\mathbb{N}}_{M-1}}\frac{T^{[1]}_{n-k}}{k} > 1-a\right] \end{aligned} \tag{5.3.12}$$

A combination of (5.3.10), (5.3.11) and (5.3.12) leads to

$$\begin{aligned} & \lim_{K\to\infty} P\left[\sup_{k\in\mathbb{N}}\frac{T^{[1]}_{n-k}}{k} > K\right] && \leq \ \epsilon \\ \Longrightarrow \quad & P\left[\sup_{k\in\mathbb{N}}\frac{T^{[1]}_{n-k}}{k} < \infty\right] && \geq \ 1-\epsilon \quad \forall\,\epsilon>0 \\ \Longrightarrow \quad & \sup_{k\in\mathbb{N}}\frac{T^{[1]}_{n-k}}{k} < \infty && \text{a.e.} \\ \Longrightarrow \quad & P\left[\sup_{k\in\mathbb{N}}\frac{T^{[1]}_{n-k}}{k} < \infty\right] && > \ 0 \\ \Longrightarrow \quad & \mathrm{E}T^{[1]}_n \ < \ \infty && \begin{array}{l}\text{by Lemma 5.3.1}\\ \text{since } T^{[1]}_n = (T^{[1]}_n)^+\end{array} \end{aligned}$$

□

Remark 5.3.1

(a) Part (b) of Theorem 5.3.1 is effectively a converse of part (a) of that theorem.

(b) (5.3.8) is another example of a strong solution (cf. Remark 2.5.1(d) and Definition 2.5.1). In fact, this chapter is connected in several (obvious and more subtle) ways to Part I. For instance $W_n(n-r, W^0)$ and $\boldsymbol{w}_n(\boldsymbol{w}, \varphi)$ are related concepts.

(c) In the notation of Part I, $\mathrm{E}u_n = \mathrm{E}(s_n - \tau_n)$ so that the stipulations of Theorem 5.3.1 (e.g., (5.3.5)) relate to those in (5.1.1b).

5.4 APPROXIMATION THEOREMS

A common approach for determining network characteristics is to use the Poisson (or independence) approximation, i.e., the flows of customers are taken to be Poisson (and assumptions on the independence of different service channels are made). The Poisson approximation may be paralleled with the mean-field approximation of statistical mechanics. A key feature is the effect of letting the number of nodes $N \to \infty$ (hence a mean-field *limit*).

Dobrushin & Sukhov [29] and Brown & Pollett [19] contain the first rigorous proofs of convergence to a so-called limiting mean-field picture. Although the latter article is not concerned with starlike networks, but instead with a slight generalisation of Jackson networks (sometimes termed migration processes), the concept of a Poisson approximation does feature prominently.

The usual (and intuitive) notational conventions for the RMPPs are applied in the sequel. A subscript for the channel is suppressed because of the symmetry properties.

Theorem 5.4.1

Suppose (5.1.1b) is satisfied. Then for any bounded time interval $\mathcal{T} \subset \mathbb{R}$, we have the following result (which "involves" the variation metric)

$$\lim_{t_0 \to -\infty,\, N \to \infty} \sup \left| L(\eta'^{(N)}(\vec{t_0}) I_{\mathcal{T}}) - L(\eta'^{(\infty)} I_{\mathcal{T}}) \right| = 0$$

<u>Proof</u>

Recall that, under the conditions of Theorem 5.3.1, the minimal solution ((5.3.8)) to (5.3.3) is of the form

$$\begin{aligned} W_n &= T_n^{[1]} + \max \left[0, \sup_{k \in \mathbb{N}} \left(T_{n-k}^{[1]} - T_n^{[1]} + \sum_{j=n-k}^{n-1} (S_j^{[2]} - (t_{j+1} - t_j)) \right) \right] \\ \Rightarrow \quad w_n^{[2]} &= \sup_{k \in \mathbb{N}} \left[\sum_{j=n-k}^{n-1} S_j^{[2]} + (T_{n-k}^{[1]} - T_n^{[1]}) - (t_n - t_{n-k}) \right]^+ \\ \Rightarrow \quad w_n^{[2]} &= \sup_{m < n} \left[\sum_{j=m}^{n-1} S_j^{[2]} + (T_m^{[1]} - T_n^{[1]}) - (t_n - t_m) \right]^+ \end{aligned} \tag{5.4.1}$$

We wish, however, to formulate the continuous analogue of (5.4.1). This is not a difficult task. Before proceeding, we simplify the notation by writing $y \in \omega_j'$ where ω_j' is a realisation of an RMPP with marks $(\boldsymbol{S}, \boldsymbol{w})$. The destination of message y is known to be j; its source is not known (nor of real interest at this stage).

Also

$$\begin{array}{rcl} y^F & = & y^F\left(t', t; \omega_j'\right) \text{ is the first message of } \omega_j' \\ & & \text{to arrive in the network during the interval } [t', t] \\ y^L & = & y^L\left(t', t; \omega_j'\right) \text{ is the last message of } \omega_j' \\ & & \text{to arrive in the network during the interval } [t', t] \end{array}$$

Then

$$v_{t(y)}\left(\omega_j'\right) = \sup_{t' < t(y)} \left[\sum_{\tilde{y} \in \omega_j' :\, t(\tilde{y}) \in [t', t(y))} S^{[2]}\left(\tilde{y}\right) + T^{[1]}\left(y^F\right) - T^{[1]}\left(y^L\right) - \left(t(y) - t'\right) \right]^+ \tag{5.4.2}$$

is the desired expression which can be generalised from $t(y)$ to arbitrary t, as follows:

$$v_t\left(\omega_j'\right) = \sup_{t' < t} \left[\sum_{\tilde{y} \in \omega_j' :\, t(\tilde{y}) \in\ [t', t)} S^{[2]}(\tilde{y}) + T^{[1]}\left(y^F\right) - T^{[1]}\left(y^L\right) - (t - t') \right]^+ \tag{5.4.3}$$

Clearly we have the following relationship[1]

$$W(y) = T^{[1]}(y) + v_{t(y)}(\omega_j')$$

[1] $v_{t(y)}(\omega_j') = w^{[2]}(y)$

where, from now on, we will again suppress the subscript $(_j)$. Define

$$\overleftarrow{z}_t(\omega') = \sup[t' < t : v_{t'}(\omega') = 0] \tag{5.4.4a}$$

$$\overrightarrow{z}_t(\omega') = \inf[t'' > t : v_{t''}(\omega') = 0] \tag{5.4.4b}$$

so that, roughly speaking, $\overleftarrow{z}_t(\omega')$ is the time of the last renewal in ω' prior to t and, similarly, $\overrightarrow{z}_t(\omega')$ is the time of the next renewal in ω' after t. For instance, suppose $\overleftarrow{z}_t(\omega') = t_1$ and $\overrightarrow{z}_t(\omega') = t_2$. If $v_t(\omega') = 0$ (a renewal at t), then $\overrightarrow{z}_{t_1}(\omega') = t = \overleftarrow{z}_{t_2}(\omega')$. Otherwise $\overrightarrow{z}_{t_1}(\omega') = t_2$ or, equivalently, $\overleftarrow{z}_{t_2}(\omega') = t_1$.

Return to the main argument of the proof. Write ω'_* for the realisation of the flow with marks $(\boldsymbol{S}, w^{[1]})$ obtained from ω' by deleting component W.

Now suppose that for some pair of realisations ω' and $\overline{\omega}'$ (and the corresponding ω'_* and $\overline{\omega}'_*$) the following properties are satisfied:

$$\min[\overleftarrow{z}_t(\omega'), \overleftarrow{z}_t(\overline{\omega}')] > t - s \tag{5.4.5a}$$

$$\max[\overleftarrow{z}_t(\omega'), \overleftarrow{z}_t(\overline{\omega}')] < \bar{t} < t \tag{5.4.5b}$$

$$\omega'_* I_{(t-s,t]} = \overline{\omega}'_* I_{(t-s,t]} \tag{5.4.5c}$$

where $t \in \mathbb{R}, s \in \mathbb{R}_+$.

We conclude that

$$t - s < \left\{ \begin{array}{l} \sup[t' < t : v_{t'}(\omega') = 0] \\ \sup[t'' < t : v_{t''}(\overline{\omega}') = 0] \end{array} \right\} < \bar{t} < t \tag{5.4.6}$$

and from (5.4.5c)

$$\begin{aligned} v_{\tilde{t}}(\omega') &= v_{\tilde{t}}(\overline{\omega}') \quad \forall\, \tilde{t} \in (\bar{t}, t] \\ \Rightarrow \; W(y, \omega') &= W(y, \overline{\omega}') \quad \forall\, y \in \omega' I_{(\bar{t},t]} = \overline{\omega}' I_{(\bar{t},t]} \end{aligned}$$

The remainder of the proof hinges mostly on a cumbersome sequence of inequalities. For this reason the details (most of which may be found in Baccelli *et al.* [4]) are omitted.

□

As an extension to Theorem 5.4.1, we have the following result.

Theorem 5.4.2

Suppose (5.1.1b) is satisfied. Then for any bounded time interval $\mathcal{T} \subset \mathbb{R}$ and any finite set $\mathcal{J} \subset \mathbb{N}$, we have

$$\lim_{t_0 \to -\infty,\, N \to \infty} \sup \left| L\left\{ \eta_j'^{(N)}(\vec{t_0}) I_{\mathcal{T}}; j \in \mathcal{J} \right\} - L\left\{ \eta_j'^{(\infty)} I_{\mathcal{T}}; j \in \mathcal{J} \right\} \right| = 0$$

5.5 THE EXPONENTIAL CASE

Consider the following stationary-regime stochastic equation which is analogous to (5.3.3):

$$W \sim \max\left[T, W + S - \tau\right] \tag{5.5.1}$$

The four random variables on the right-hand side, W, T ($\sim T^{[1]}$), S ($\sim S^{[2]}$) and τ, are nonnegative and assumed to be mutually independent. The distributions of the latter three are known. In particular $S \sim F^{[2]}$ and $T \sim \text{Exp}(\lambda^{-1})$. What is unknown is the distribution of the waiting time W.

If we let $X \sim W + S - \tau$, then (5.5.1) may be rewritten as

$$X + \tau - S \sim \max(X, T)$$

or

$$X \sim S - \tau + \max(X, T) \tag{5.5.2}$$

Much as before, random variables X, T, S and τ on the right of equation (5.5.2) may be assumed to be independent.

In this section, further conditions are stipulated for (5.5.2). In addition to the assumption that $\tau \sim \text{Exp}(\lambda^{-1})$, we suppose that $S \sim \text{Exp}(\mu^{-1})$. G shall represent the distribution function (and g, the density) of the difference $u = S - \tau$, while the distribution functions of T and X are denoted by $\mathcal{F}_1$ and $\mathcal{F}$ respectively.

We wish to rewrite (5.5.2) i.t.o. these distribution functions. In order to do so, note that

$$\begin{aligned} F_{\max(X,T)}(r) &= P[\max(X,T) \leq r] \\ &= P[\{X \leq r\} \cap \{T \leq r\}] \\ &= P[X \leq r].P[T \leq r] \qquad \text{independence of } X \text{ and } T \\ &= \mathcal{F}(r)\mathcal{F}_1(r) \\ F_{\max(X,T)}(r) &= (\mathcal{F}\mathcal{F}_1)(r) \end{aligned} \tag{5.5.3}$$

The independence of the various random variables also leads to the following:

$$\begin{aligned}
\mathcal{F}(x) &= F_X(x) \\
&= F_{u+\max(X,T)}(x) && \text{from (5.5.2)} \\
&= F_u * F_{\max(X,T)}(x) \\
&= G * (\mathcal{F}\mathcal{F}_1)(x) && \text{from (5.5.3)} \\
&= \int_{-\infty}^{\infty} \mathcal{F}(x-y)\mathcal{F}_1(x-y)dG(y) \\
\text{or} \quad \mathcal{F}(x) &= \int_0^{\infty} \mathcal{F}(r)\mathcal{F}_1(r)g(x-r)dr && (5.5.4)
\end{aligned}$$

since T is a nonnegative random variable.

Clearly it is necessary to determine the distribution and density functions, G and g respectively, of $u = S - \tau$.

From a diagram it is clear that $P[S - \tau \leq x]$ has to be considered separately for the cases $x \geq 0$ and $x < 0$:

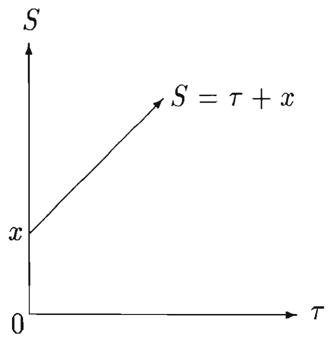


Fig. 5.1(a)
Graph of $S = \tau + x$
$x \geq 0$

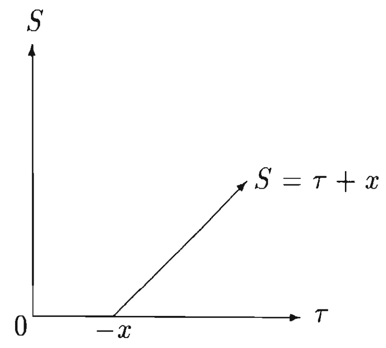


Fig. 5.1(b)
Graph of $S = \tau + x$
$x < 0$

$$Case \quad 1 \quad : \quad x \geq 0$$

$$\begin{aligned}
G(x) &= P[S-\tau \leq x] \\
&= P[S \leq \tau + x] \\
&= \int_{\tau=0}^{\infty} \lambda e^{-\lambda\tau} \int_{S=0}^{\tau+x} \mu e^{-\mu S} dS d\tau \\
&= \int_{\tau=0}^{\infty} \lambda e^{-\lambda\tau} [1 - e^{-\mu(\tau+x)}] d\tau \\
&= 1 - \tfrac{\lambda}{\lambda+\mu} e^{-\mu x}
\end{aligned}$$

$$\begin{aligned}
\Longrightarrow \quad g(x) &= G'(x) \\
&= \tfrac{\lambda\mu}{\lambda+\mu} e^{-\mu x}
\end{aligned}$$

$$Case \quad 2 \quad : \quad x < 0$$

$$\begin{aligned}
G(x) &= P[S \leq \tau + x] \\
&= \int_{\tau=-x}^{\infty} \lambda e^{-\lambda\tau} \int_{S=0}^{\tau+x} \mu e^{-\mu S} dS d\tau \\
&= \int_{\tau=-x}^{\infty} \lambda e^{-\lambda\tau} [1 - e^{-\mu(\tau+x)}] d\tau \\
&= e^{\lambda x} - \tfrac{\lambda}{\lambda+\mu} e^{\lambda x} \\
&= \tfrac{\mu}{\lambda+\mu} e^{\lambda x}
\end{aligned}$$

$$\Longrightarrow \quad g(x) = \tfrac{\lambda\mu}{\lambda+\mu} e^{\lambda x}$$

Thus

$$g(x) = \begin{cases} \tfrac{\lambda\mu}{\lambda+\mu} e^{-\mu x} & x \geq 0 \\ \tfrac{\lambda\mu}{\lambda+\mu} e^{\lambda x} & x < 0 \end{cases}$$

$$\lim_{x\to 0+} g(x) = \left\{ \begin{array}{c} \frac{\lambda\mu}{\lambda+\mu} \\ g(0) \end{array} \right\} = \lim_{x\to 0-} g(x)$$

so, in fact

$$g(x) = \begin{cases} \frac{\lambda\mu}{\lambda+\mu} e^{-\mu x} & x \geq 0 \\ \frac{\lambda\mu}{\lambda+\mu} e^{\lambda x} & x \leq 0 \end{cases} \tag{5.5.5}$$

Moreover $g(x)$ has a derivative on $\mathbb{R} \setminus \{0\}$:

$$g'(x) = \begin{cases} -\mu g(x) & x > 0 \\ \lambda g(x) & x < 0 \end{cases} \tag{5.5.6}$$

and a second derivative on $\mathbb{R} \setminus \{0\}$:

$$g''(x) = \begin{cases} \mu^2 g(x) & x > 0 \\ \lambda^2 g(x) & x < 0 \end{cases} \tag{5.5.7}$$

Since g is continuous on $\mathbb{R}$ with piecewise continuous derivative g', we have that

$$\mathcal{F}'(x) = \int_0^\infty \mathcal{F}(r)\mathcal{F}_1(r) g'(x-r)dr \tag{5.5.8}$$

The jump of g' at 0 (Apostol [1, *p.* 93]) is given by

$$\begin{aligned} J &= g'(0+) - g'(0-) \\ &= -(\lambda+\mu)g(0) \\ &= -\lambda\mu \end{aligned} \tag{5.5.9}$$

Thus

$$\begin{aligned} \mathcal{F}_+''(x) &= \int_0^\infty \mathcal{F}(r)\mathcal{F}_1(r) g''(x-r)dr + J(\mathcal{F}\mathcal{F}_1)(x+0) \\ &= \int_0^\infty \mathcal{F}(r)\mathcal{F}_1(r) g''(x-r)dr - \lambda\mu\mathcal{F}(x)\mathcal{F}_1(x+0) \end{aligned} \tag{5.5.10}$$

For $x \neq 0$, however,

$$\begin{aligned} & \lambda\mu g(x) + (\lambda-\mu)g'(x) \\ = & \left\{ \begin{array}{lcll} \lambda\mu g(x) - \mu(\lambda-\mu)g(x) & = & \mu^2 g(x) & x > 0 \\ \lambda\mu g(x) + \lambda(\lambda-\mu)g(x) & = & \lambda^2 g(x) & x < 0 \end{array} \right\} \\ = & g''(x) \end{aligned} \tag{5.5.11}$$

As a result, by combining (5.5.10) and (5.5.11),

$$\begin{aligned} \mathcal{F}''_+(x) + \lambda\mu\mathcal{F}(x)\mathcal{F}_1(x+0) &= \int_0^\infty \mathcal{F}(r)\mathcal{F}_1(r)[\lambda\mu g(x-r) + (\lambda-\mu)g'(x-r)]dr \\ &= \lambda\mu\mathcal{F}(x) + (\lambda-\mu)\mathcal{F}'(x) \end{aligned}$$

so that the right-hand second derivative is determined by

$$\mathcal{F}''_+(x) + (\mu-\lambda)\mathcal{F}'(x) = \lambda\mu\mathcal{F}(x)[1-\mathcal{F}_1(x+0)] \tag{5.5.12a}$$

and similarly for the left-hand second derivative:

$$\mathcal{F}''_-(x) + (\mu-\lambda)\mathcal{F}'(x) = \lambda\mu\mathcal{F}(x)[1-\mathcal{F}_1(x-0)] \tag{5.5.12b}$$

We note that, if $\mathcal{F}_1$ is continuous at x, then $\mathcal{F}_1(x+0) = \mathcal{F}_1(x) = \mathcal{F}_1(x-0)$. From this, it is easily deduced that $\mathcal{F}''_+(x)$ and $\mathcal{F}''_-(x)$, of (5.5.12a) and (5.5.12b) respectively, coincide so that $\mathcal{F}''(x)$ exists and is given by

$$\mathcal{F}''(x) + (\mu-\lambda)\mathcal{F}'(x) = \lambda\mu\mathcal{F}(x)\overline{\mathcal{F}_1}(x) \tag{5.5.12}$$

The results used in this section are from the theory of functionals and generalized functions, especially the *Heaviside unit-step function* $H(x) = \left\{ \begin{array}{ll} 1 & x > 0 \\ 0 & x < 0 \end{array} \right.$

and its generalised derivative, the *Dirac delta function* $\delta(x)$. The material required for this section may be found in Butkov [20, Chapter 6], Dettman [27, *pp.* 235–242] and Kolmogorov & Fomin [44, §21].

We continue with some additional properties of interest. Substitution into (5.5.4) yields

$$\begin{aligned} \mathcal{F}(0) &= \int_0^\infty \mathcal{F}(r)\mathcal{F}_1(r)g(-r)dr \\ &= \int_0^\infty \mathcal{F}(r)\mathcal{F}_1(r)\tfrac{\lambda\mu}{\lambda+\mu}e^{-\lambda r}dr \end{aligned} \tag{5.5.13}$$

using (5.5.5) in accordance with the fact that $-r \in (-\infty, 0]$.

If $x < 0$, then (5.5.4) reduces to

$$\begin{aligned} \mathcal{F}(x) &= \int_0^\infty \mathcal{F}(r)\mathcal{F}_1(r)\tfrac{\lambda\mu}{\lambda+\mu}e^{\lambda(x-r)}dr \\ &= e^{\lambda x}\int_0^\infty \mathcal{F}(r)\mathcal{F}_1(r)\tfrac{\lambda\mu}{\lambda+\mu}e^{-\lambda r}dr \\ &= e^{\lambda x}\mathcal{F}(0) \qquad \text{from (5.5.13)} \end{aligned} \tag{5.5.14}$$

Letting $x = 0$ in (5.5.8) gives

$$\begin{aligned} \mathcal{F}'(0) &= \int_0^\infty \mathcal{F}(r)\mathcal{F}_1(r)g'(-r)dr \\ &= \lambda\int_0^\infty \mathcal{F}(r)\mathcal{F}_1(r)g(-r)dr \quad \text{from (5.5.6)} \\ &= \lambda\mathcal{F}(0) \qquad \text{from (5.5.13)} \end{aligned} \tag{5.5.15}$$

Example 5.5.1 (A Special Case)

Suppose $T \sim \mathrm{Exp}(\nu^{-1})$ so that

$$\mathcal{F}_1(x) = 1 - e^{-\nu x} \qquad x \geq 0 \tag{5.5.16}$$

In this special case (5.5.12) becomes

$$\mathcal{F}''(x) + (\mu - \lambda)\mathcal{F}'(x) - \lambda\mu e^{-\nu x}\mathcal{F}(x) = 0 \quad x > 0 \tag{5.5.17}$$

($\mathcal{F}_1(x)$ is continuous $\forall\, x > 0$).

The following transformation reduces (5.5.17) to a more suitable form:

$$t = t(x) = \frac{2\sqrt{\lambda\mu}}{\nu} e^{-\frac{\nu}{2}x} \tag{5.5.18a}$$

$$\mathcal{F} = t^{\alpha} U \qquad \alpha = \frac{\mu - \lambda}{\nu} \tag{5.5.18b}$$

Then, representing differentiation w.r.t. t with the usual fluxional notation, from (5.5.18a), we obtain:

$$\begin{aligned} \tfrac{\nu^2}{4}t^2 &= \lambda\mu e^{-\nu x} \quad \text{(a term in (5.5.17))} \\ t'(x) &= -\tfrac{\nu}{2}t \end{aligned}$$

Also

$$
\begin{aligned}
\dot{\mathcal{F}}\left(t(x)\right) &= \alpha t^{\alpha-1}U + t^{\alpha}\dot{U} \\
&= t^{\alpha-1}\left[\alpha U + t\dot{U}\right] \\
\mathcal{F}'\left(t(x)\right) &= \frac{d\mathcal{F}\left(t(x)\right)}{dx} \\
&= \dot{\mathcal{F}}\left(t(x)\right)t'(x) \\
&= -\tfrac{\nu}{2}t^{\alpha}\left[\alpha U + t\dot{U}\right] \\
\mathcal{F}''\left(t(x)\right) &= \frac{d\mathcal{F}'\left(t(x)\right)}{dx} \\
&= \frac{d\mathcal{F}'\left(t(x)\right)}{dt}t'(x) \\
&= \tfrac{\nu^2}{4}t\left\{\alpha t^{\alpha-1}\left[\alpha U + t\dot{U}\right] + t^{\alpha}\left[(\alpha+1)\dot{U} + t\ddot{U}\right]\right\} \\
&= \tfrac{\nu^2}{4}t^{\alpha}\left[\alpha^2 U + (2\alpha+1)t\dot{U} + t^2\ddot{U}\right]
\end{aligned}
$$

So

$$\mathcal{F} = t^{\alpha+2}\tfrac{1}{t^2}U \tag{5.5.19a}$$

$$\mathcal{F}' = -\tfrac{\nu}{2}t^{\alpha+2}\left[\alpha\tfrac{1}{t^2}U + \tfrac{1}{t}\dot{U}\right] \tag{5.5.19b}$$

$$\mathcal{F}'' = \tfrac{\nu^2}{4}t^{\alpha+2}\left[\alpha^2\tfrac{1}{t^2}U + (2\alpha+1)\tfrac{1}{t}\dot{U} + \ddot{U}\right] \tag{5.5.19c}$$

Substitution of (5.5.19a)–(5.5.19c) into (5.5.17) (on multiplying through by $\frac{4}{\nu^2 t^{\alpha+2}}$) yields

$$\alpha^2 \tfrac{1}{t^2} U + (2\alpha+1)\tfrac{1}{t}\dot{U} + \ddot{U} - \tfrac{2(\mu-\lambda)\alpha}{\nu}\tfrac{1}{t^2}U - \tfrac{2(\mu-\lambda)}{\nu}\tfrac{1}{t}\dot{U} - U = 0$$

$$\begin{aligned} \Rightarrow\ & \ddot{U} + (2\alpha+1-2\alpha)\tfrac{1}{t}\dot{U} - \left(1 - \tfrac{\alpha^2}{t^2} + \tfrac{2\alpha^2}{t^2}\right)U = 0 \ \text{ from (5.5.18b)} \\ \Rightarrow\ & \ddot{U} + \tfrac{1}{t}\dot{U} - \left(1 + \tfrac{\alpha^2}{t^2}\right)U = 0 \end{aligned} \tag{5.5.20}$$

Bessel's modified differential equation (5.5.20) (Bronshtein & Semendyayev [18, *pp.* 410–413]) has the following solution:

$$U(t) = \begin{cases} c_1 I_\alpha(t) + c_2 K_\alpha(t) & \alpha \in \mathbb{R} \\ c_1 I_\alpha(t) + c_2 I_{-\alpha}(t) & \alpha \in \mathbb{R} \setminus \mathbb{Z} \end{cases} \tag{5.5.21}$$

(since $I_{\pm\alpha}(t)$ and $K_\alpha(t)$ are solutions to (5.5.20) but
I_α and $I_{-\alpha}$ are linearly dependent for $\alpha \in \mathbb{Z}$;
I_α and K_α are linearly independent $\forall\, \alpha \in \mathbb{R}$)

where

$$\begin{aligned} I_\alpha(t) &= i^{-\alpha} J_\alpha(it) \\ &= \sum_{k=0}^{\infty} \frac{1}{k!\,\Gamma(\alpha+k+1)} \left(\frac{t}{2}\right)^{2k+\alpha} \qquad \alpha \in \mathbb{R} \end{aligned}$$

is the *modified Bessel function of the first kind of order* α,

$$K_\alpha(t) = \begin{cases} \dfrac{\pi\,[I_{-\alpha}(t) - I_\alpha(t)]}{2\sin(\alpha\pi)} & \alpha \in \mathbb{R} \setminus \mathbb{Z} \\[2ex] \lim\limits_{p\to\alpha} \dfrac{\pi\,[I_{-p}(t) - I_p(t)]}{2\sin(p\pi)} & \alpha \in \mathbb{Z} \end{cases}$$

is the *modified Bessel function of the second kind of order* α (also known as the *Macdonald function*), and

c_1, c_2 are determined by the constraints $\mathcal{F}(-\infty) = 0$ ($\mathcal{F}(0) = 0$), $\mathcal{F}(\infty) = 1$

Remark 5.5.1

(5.5.16) is of particular interest when $\nu = \mu - \lambda$ (which is obviously positive because $\frac{\lambda}{\mu} < 1$). In this instance, the $M/GI/1/\infty$ queue is nothing more than the familiar $M/M/1$ system.

Using (5.5.18b), $\alpha = \frac{\mu-\lambda}{\nu} = 1$ ($\Rightarrow \mathcal{F} = tU$) so that (5.5.21) becomes

$$U(t) = c_1 I_1(t) + c_2 K_1(t)$$

Example 5.5.2 (Basic Lindley Equation)

If $P[T = 0] = 1$, then $T \sim H(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases}$ In this case (5.5.1) reduces to $W \sim [W + S - \tau]^+$ w.p.1 and (5.5.12) becomes

$$\mathcal{F}''(x) + (\mu - \lambda)\mathcal{F}'(x) = \lambda\mu\mathcal{F}(x)\overline{H}(x)$$

or

$$\mathcal{F}''(x) = \begin{cases} (\lambda - \mu)\mathcal{F}'(x) & x \geq 0 \\ (\lambda - \mu)\mathcal{F}'(x) + \lambda\mu\mathcal{F}(x) & x < 0 \end{cases} \tag{5.5.22}$$

We shall solve the differential equations given in (5.5.22) (but with $x = 0$ excluded) in §6.4.

CHAPTER 6

HIGHER-ORDER LINDLEY EQUATIONS

6.1 INTRODUCTION

In the previous chapter, a so-called extended Lindley equation was shown to govern the operation of a $FAFS$ message-switched starlike network. In this chapter, we will be concerned with the circuit-switched analogue of that discipline. Consideration of a "generalised starlike network" with m (often greater than 2) stages will lead to an mth-order Lindley equation.

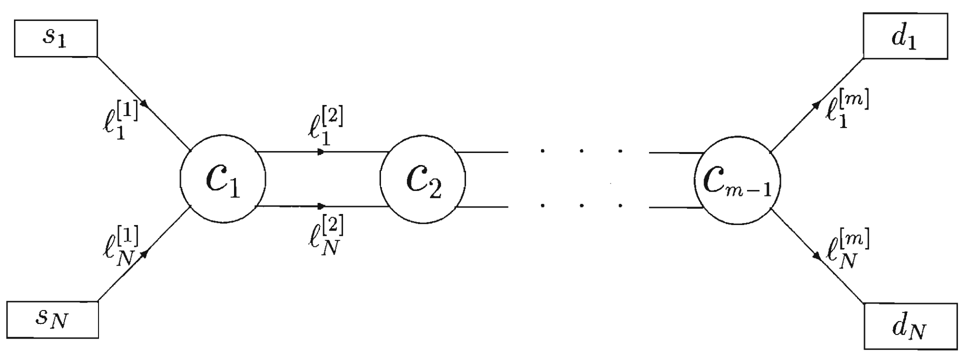


Fig. 6.1

Generalisation of the Starlike Network

Figure 6.1 above illustrates the situation. Minor adjustments and generalisations of notation used in Figure 4.1 are made. $\ell_i^{[j]}$ denotes the ith line of the jth stage ($i \in \mathbb{N}_N$ and $j \in \mathbb{N}_m$) and c_k is the hub linking stages k and $k+1$, so $\ell_{\cdot}^{[j]} = c_{j-1}c_j$ ($j \neq 1$). Sources and destinations are labelled as before.

In Chapter 4, the circuit-switching $FAFS$ second-stage protocol was introduced on a starlike network. In the current chapter, we will be concerned mostly with general $m \geq 2$. The $FAFS$ discipline operates in the final (mth) stage where transmission occurs. Otherwise, in all m stages, the general circuit-switching principle of waiting for all those calls with earlier arrival times, *and* which have a path intersecting with one's own, to leave the network, applies.

6.2 CONSTRUCTION OF THE mth-ORDER LINDLEY EQUATION, AND ITS SOLUTIONS

It is convenient to extend and generalise the concepts of source and destination predecessors and successors (Definitions 4.3.1 and 4.3.2) to our m-stage network.

To do so, it is necessary to realise that, when $m > 2$, the "marked" path through the network is determined not merely by a source (line) and destination (line) but by m lines: $(i_1, i_2, \ldots, i_m) = \boldsymbol{i}$, where i_j is the line in stage j ($\in \mathbb{N}_m$) and $i_j \in \mathbb{N}_N$.

Consequently $y^{(j)} = [t'; (S', (i'_1, \ldots, i_j, \ldots, i'_m), W')]$ is said to be the *(immediate) jth-stage predecessor* of $y = [t; (S, (i_1, \ldots, i_j, \ldots, i_m), W)]$ (symbolically $y^{(j)} = p_j(y)$) iff

$$t' < t \text{ and } t' = \max\left[t(\tilde{y}) : t(\tilde{y}) < t \wedge i_j(\tilde{y}) = i_j\right]$$

Obviously the definition may be extended to include the identity $y = p_j^0(y)$, nth jth-stage predecessors $p_j^n(y)$ and successors $p_j^{-n}(y)$ $(j \in \mathbb{N}_m, n \in \mathbb{N})$. Much as in Definition 4.3.3, we may consider *total index* $n(y)$ and *jth-stage index* $n_j(y)$.

According to the synchronization constraints for this circuit-switching model, y may gain access to line i_j only once $y^{(j)}$ has left the network. Let $w^{[1]}(y)$ denote the wait y experiences from the time it arrives to the time it gains its first-stage line i_1. Furthermore, write $w^{[j]}(y)$ $(j \neq 1)$ for the waiting time of y in the jth stage (i.e., the length of the interval from the time y gets the line i_{j-1} to that time when it pretends on i_j). Also write $W\left(y^{(j)}\right)$ for the network waiting time (excluding service) of predecessor $y^{(j)}$, $\tau\left(y^{(j)}\right) = t - t\left(y^{(j)}\right)$ and $u\left(y^{(j)}\right) = S\left(y^{(j)}\right) - \tau\left(y^{(j)}\right)$.

Then, applying the idea of Chapter 1

$$\begin{aligned}
w^{[j]}(y) &= \left[W\left(y^{(j)}\right) + u\left(y^{(j)}\right) - \sum_{i=1}^{j-1} w^{[i]}(y)\right]^+ \\
\Longrightarrow \quad W(y) &= \sum_{j=1}^{m} w^{[j]}(y) \\
&= \sum_{j=1}^{m-1} w^{[j]}(y) + \left[W\left(y^{(m)}\right) + u\left(y^{(m)}\right) - \sum_{i=1}^{m-1} w^{[i]}(y)\right]^+ \\
&= \max\left[\sum_{i=1}^{m-1} w^{[i]}(y), W\left(y^{(m)}\right) + u\left(y^{(m)}\right)\right] \\
&= \max_{j\in\mathbb{N}_m}\left[W\left(y^{(j)}\right) + u\left(y^{(j)}\right)\right]^+
\end{aligned}$$

using induction, initial condition $w^{[1]}(y) = \left[W\left(y^{(1)}\right) + u\left(y^{(1)}\right)\right]^+$ and Remark 2.3.1. From a practical point of view the result is fairly intuitive since we know that call y may be transmitted only once the last of its m predecessors have left the network. Obviously the possibility exists that $p_j(y) = p_k(y)$ $((j \neq k) \in \mathbb{N}_m)$, which then implies that y and $y^{(j)} = y^{(k)}$ share lines i_j *and* i_k. We will return to this situation later in the chapter.

For now we summarise our conclusions with the equation

$$W(y) = \max_{j\in\mathbb{N}_m}\left[W\left(y^{(j)}\right) + u\left(y^{(j)}\right)\right]^+ \tag{6.2.1}$$

At this stage it is important to realise that we are not making any claims regarding stationarity or uniqueness (nor are we specifying the finiteness or otherwise of N) but rather formulating a (non-stationary) Lindley equation much like (1.1.1). Less significantly, it is evident that, for now, (6.2.1) is formulated i.t.o. random variables rather than their distributions.

More especially, we aim to prove the existence of a stationary regime with waiting times governed by the stochastic equation (which is of particular interest when $N \to \infty$), given by

$$W \sim \max_{j \in \mathbb{N}_m} \left[W^{(j)} + u^{(j)}\right]^+ \tag{6.2.2}$$

Naturally enough, (6.2.2) bears a strong resemblance to (6.2.1).

In accordance with previous notation and assumptions, $u^{(j)} = S^{(j)} - \tau^{(j)}$ where IID service times $S^{(j)} \sim F$, $j \in \mathbb{N}_m$ and IID interarrival times $\tau^{(j)} \sim \text{Exp}(\lambda^{-1})$, $j \in \mathbb{N}_m$ (although a more general distribution is possible). $\left\langle S^{(j)},\ j \in \mathbb{N}_m\right\rangle$, $\left\langle \tau^{(j)},\ j \in \mathbb{N}_m\right\rangle$ and $\left\langle W^{(j)},\ j \in \mathbb{N}_m\right\rangle$ are assumed to be mutually independent sequences. In addition $\left\langle W, W^{(j)},\ j \in \mathbb{N}_m\right\rangle$ is supposed to be a sequence of identically distributed random variables.

There is a natural generalisation of the source-flow family $\boldsymbol{\xi}$ and destination-flow family $\boldsymbol{\xi}'$ of Chapters 4 and 5. Any call with line i_j on its path-vector $\boldsymbol{i}$ will form part of the jth-stage flow $\xi_{i_j} \in \boldsymbol{\xi}^{[j]}$ ($i_j \in \mathbb{N}_N$, $j \in \mathbb{N}_m$). By assumption, and the symmetry of the model, $\boldsymbol{\xi}^{[j]}$ is a collection of IID Poisson flows each of which has rate λ. The marks of ξ_{i_j} are given by $(\boldsymbol{i}_{(j)}, S)$ where $\boldsymbol{i}_{(j)}$ is that $(m-1)$-dimensional vector derived from $\boldsymbol{i}$ by extracting the jth component. The inclusion of i_j would be redundant because $\boldsymbol{\xi}^{[j]}$ specifies the jth line. Realisations ω_{i_j} will include arrival times of calls in that flow.

Mention was made in Chapter 4 of extended RMPPs, and the idea was exploited in Chapter 5. As was the case there, the waiting time is included in the "mark" of $\eta_{i_j} \in \boldsymbol{\eta}^{[j]}$. Sample realisations, which include times, are determined by $\tilde{\omega}_{i_j} \in \tilde{\boldsymbol{\omega}}^{[j]}$.

The concept of RMPPs may be taken even further. Each of the N^m different paths $(i_1, \ldots, i_m)$ may be associated with a flow $\xi(\boldsymbol{i})$ (strictly $\omega(\boldsymbol{i})$ if we are referring to sample realisations) of calls having that path. The fact that $\boldsymbol{i}$ determines $\xi(\boldsymbol{i})$ means that it really need not be included in the marks. Since the total arrival rate to the network is $N\lambda$, symmetry and IID assumptions imply that any one stream $\xi(\boldsymbol{i})$ has rate $\frac{N\lambda}{N^m} = \frac{\lambda}{N^{m-1}}$. Extended flows $\eta(\boldsymbol{i})$ are naturally defined.

The formalities of the preceding paragraphs, with the rather unfortunate notation that accompanies them, are necessary. In order not to let notation detract from the basic argument, however, the first theorem is formulated for the case $m = 2$ and the family $\boldsymbol{\xi}$ (or $\boldsymbol{\xi}'$) and $\boldsymbol{\eta}$ (or $\boldsymbol{\eta}'$) of flows. The statement and proof for more general values of m is essentially the same. In any case, the proof of the theorem relies on the results on the following pages.

Lemma 6.2.1 (Extended Version of Chebyshev's Inequality)
Suppose X is a random variable with distribution function F. If, for some $a > 0$, $\mathrm{E}e^{aX}$ exists, then $P[X > 0] <\mathrm{E}e^{aX}$.

<u>PROOF</u>

Write $A = [X > 0]$. Then we prove the result for a continuous X.

$$
\begin{aligned}
\mathrm{E}e^{aX} &= \int e^{ax} dF(x) \\
&= \int_A e^{ax} dF(x) + \int_{\overline{A}} e^{ax} dF(x) \\
&\geq \int_A e^{ax} dF(x) \\
&\overset{a>0}{>} \int_A dF(x) & (=) \\
&> P[A] & (=) \\
&> P[X > 0]
\end{aligned}
$$

□

Corollary 6.2.1*a*

Suppose X is a random variable with distribution function F, and $c \in \mathbb{R}$. If, for some $a > 0$, $\mathrm{E}e^{aX}$ exists, then $P[X > c] < e^{-ac}\mathrm{E}e^{aX}$.

<u>Proof</u>

$$\begin{aligned} P[X > c] &= P[X - c > 0] \\ &< \mathrm{E}e^{a(X-c)} \qquad \text{by Lemma 6.2.1} \end{aligned}$$

and the desired result follows.

□

Corollary 6.2.1*b*

Suppose $\langle X_n, n \in \mathbb{N}_r\rangle$ is an IID sequence of random variables, where $X_n \sim X$, the random variable in Lemma 6.2.1. Write $S_r = \sum_{n=1}^{r} X_n$.

Then, $\forall\, a > 0$ for which $\mathrm{E}e^{aX}$ exists, $P[S_r > 0] < \left(\mathrm{E}e^{aX}\right)^r$.

<u>Proof</u>

From the lemma

$$\begin{aligned} P[S_r > 0] &< \mathrm{E}e^{aS_r} && (=) \\ &< \mathrm{E}e^{a\sum_{n=1}^{r} X_n} && (=) \\ &\overset{\text{IID}}{<} \prod_{n=1}^{r} \mathrm{E}e^{aX} && (=) \\ &< \left(\mathrm{E}e^{aX}\right)^r \end{aligned}$$

□

Lemma 6.2.2

Let $\langle X_n, n \in I\rangle$ be an IID sequence, where $X_n \sim X$ and $|I| = m \in \mathbb{N}$. Then

$$P\left[\max_{n\in I} X_n > c\right] \;\leq\; mP[X > c].$$

PROOF

$$\begin{aligned}
\left[\max_{n\in I} X_n > c\right] \;&\subseteq\; \left[\bigcup_{n\in I}(X_n > c)\right] \\
\Longrightarrow\; P\left[\max_{n\in I} X_n > c\right] \;&\leq\; P\left[\bigcup_{n\in I}(X_n > c)\right] \\
&\leq\; \sum_{n\in I} P[X_n > c] \qquad (=) \\
&\overset{\cdot\mathrm{ID}}{\leq}\; \sum_{n\in I} P[X > c] \qquad (=) \\
&\leq\; mP[X > c]
\end{aligned}$$

□

We are now in a position to prove the theorem.

Theorem 6.2.1

Suppose

$$\inf_{a>0} \mathrm{E}e^{aS}\frac{\lambda}{\lambda+a} < \frac{1}{2} \tag{6.2.3}$$

Then the family $\boldsymbol{\eta} = \{\eta_i, i \in \mathbb{N}_N\}$ of extended RMPPs previously described exists and is unique.

<u>PROOF</u>

Consider the arrival of a (virtual) call $y \in \omega_i \cap \omega_j'$ at time t. We compare the waiting times $W_{t'}, W_{t''}$ of that call if zero initial conditions (system empty and servers idle) are imposed at times t' and t'' ($t'' < t'$) respectively. With these waiting times we associate the extended (truncated) flows $\eta_i(\overrightarrow{t'})$ and $\eta_i(\overrightarrow{t''})$.

Let us consider (realisations of) the collection $\overleftarrow{\boldsymbol{\xi}}_t = \left\{\overleftarrow{\xi}_{t,i}, i \in \mathbb{N}_N\right\}$ of reversed flows, where $\overleftarrow{\xi}_{t,i}$ is the flow ξ_i observed in reverse time from t (to t' and t''). This (reversed) RMPP and its marks obviously have the same distribution properties as ξ_i (Chapter 4).

We are then able to construct a *source-predecessor path*

$$P_s(y) : (y = p_s^0(y)) \; \ldots \; p_s^{n_s(y)-1}(y)$$

and a *destination-predecessor path*

$$P_d(y) : (y = p_d^0(y)) \; \ldots \; p_d^{n_d(y)-1}(y)$$

This information is, however, more effectively reflected in the form of a *predecessor graph* (more accurately, *digraph*) of y. The root r is associated with y while its adjacent vertices correspond to $p_s(y)$ and $p_d(y)$ and their respective service times. As mentioned towards the beginning of this section, $p_s(y)$ and $p_d(y)$ could coincide, in which case the vertices merge (one arc emanantes from y to $p_s(y) = p_d(y)$). A more general possibility is that $s(p_s(y)) = s(p_d(y))$. Yet another situation which can arise is that a single call $\tilde{y}$ is a predecessor for two different calls, y_1 and y_2. In other words, $\tilde{y} = p_s(y_1)$ *and* $\tilde{y} = p_d(y_2)$, in which case two arcs enter $\tilde{y}$.

The arcs $(y, p_s(y))$ and $(y, p_d(y))$ are assigned values $\tau(p_s(y)) = t - t(p_s(y))$ and $\tau(p_d(y)) = t - t(p_d(y))$ respectively (with the obvious adjustment in the case of coincidence). The situation is repeated at the vertices $p_s(y)$ and $p_d(y)$, and so on until a digraph, with dipaths of length n characterised by the n-dimensional vectors $\boldsymbol{\theta} \in \mathcal{M}^n$ ($\mathcal{M} = \{s, d\}$), emerges. $\theta_i = s$ (d) iff the ith vertex on the dipath is the source (destination) predecessor of the $(i-1)$st vertex on that path. Obviously, because of the correspondence between the path and the vector $\boldsymbol{\theta}$, at most 2^n paths of length n are possible. (If $p_s(\tilde{y}) = p_d(\tilde{y})$ then, of course, there are fewer than 2^n such paths, and we adopt the convention that $\theta_i = s$).

To simplify the notation, we may write $r = v_0 = y$, $v_i = p_{\theta_i}(v_{i-1})$ and

$e_i = (v_{i-1}, v_i)$. We may now define

$$\left.\begin{array}{lcl} u_i(\boldsymbol{\theta}) & = & S_{\boldsymbol{\theta}}(v_i) - \tau_{\boldsymbol{\theta}}(e_i) \\ U_n(\boldsymbol{\theta}) & = & \sum_{i=1}^{n} u_i(\boldsymbol{\theta}) \\ U_n & = & \max_{\boldsymbol{\theta} \in \mathcal{M}^n} U_n(\boldsymbol{\theta}) \end{array}\right\} \qquad (6.2.4)$$

and let $A_n = \{$a path of length n with positive delay exists$\}$

For such a path associated with some fixed $\boldsymbol{\theta}$, the vertices and edges correspond, of course, with IID service and interarrival times respectively.

If $\theta_i = s$, then $v_i = p_s(v_{i-1})$ and the calls corresponding to v_i and v_{i-1} are from the same source flow which implies that $\tau_{\boldsymbol{\theta}}(e_i) =_{st} \tau_i^* \sim \text{Exp}(\lambda^{-1})$ (independently of $\tau_{\boldsymbol{\theta}}(e_i)$). On the other hand, if $\theta_i = d$, then $v_i = p_d(v_{i-1})$ and the calls under consideration are from the same destination flow. If we bear in mind the discussion on destination flows earlier in this chapter and in previous chapters, we again conclude that $\tau_{\boldsymbol{\theta}}(e_i)$ is stochastically equivalent to an independent copy of an $\text{Exp}(\lambda^{-1})$ random variable. v_{i-1} must wait for $p_d(v_{i-1})$ to leave the network before v_{i-1} can be served. v_{i-1} and $p_d(v_{i-1})$ are *successive arrivals from the same destination flow* hence the time elapsing between their moments of arrival *to the network* $\sim\text{Exp}(\lambda^{-1})$. (Refer to Definition 4.3.2.)

We may also write $S_{\boldsymbol{\theta}}(v_i) =_{st} S_i^*$, an independent copy of a random variable distributed as F. On letting $u_i^* = S_i^* - \tau_i^* \sim u^*$, we obtain

$$
\begin{aligned}
P(A_n) &= P[U_n > 0] \\
&= P\left[\max_{\boldsymbol{\theta}\in\mathcal{M}^n} U_n(\boldsymbol{\theta}) > 0\right] \\
&\leq |\mathcal{M}|^n P\left[U_n(\boldsymbol{\theta}) > 0\right] \quad (=) \quad \text{by Lemma 6.2.2} \\
&\leq 2^n P\left[\sum_{i=1}^{n} u_i(\boldsymbol{\theta}) > 0\right] \quad (=) \\
&\leq 2^n P\left[\sum_{i=1}^{n} u_i^* > 0\right] \\
&< 2^n \left(\mathrm{E}e^{au^*}\right)^n \qquad \text{by Corollary 6.2.1}b \\
\Longrightarrow \quad P(A_n) &\leq 2^n \left(\inf_{a>0} \mathrm{E}e^{au^*}\right)^n \\
&< (2\cdot\tfrac{1}{2})^n \qquad \text{by (6.2.3)} \\
&\leq q^n \qquad q \in (0,1) \\
\Longrightarrow \quad \sum_n P(A_n) &< \infty \\
\Rightarrow \quad P\left(\limsup_{n\to\infty} A_n\right) &= 0
\end{aligned}
\tag{6.2.5}
$$

Hence $\widetilde{W}(y) = \left[\sup_{n\in\mathbb{N}} U_n\right]^+$ is attained on a path of finite length w.p.1, and thus is itself finite w.p.1. Consequently (w.p.1) $\widetilde{W}(y) = W(y)$, the actual waiting time of the call under consideration. (We will tend to suppress reference to call y in the sequel.)

If $|W_{t'} - W_{t''}| = 0$ (w.p.1) then convergence as $t' \to -\infty$, does not create any problems. So suppose $|W_{t'} - W_{t''}| > 0$ with positive probability. Then the predecessor path on which $W_{t''}$ is "derived" must originate from before t'.

In order to prove the existence of $\boldsymbol{\eta}$, we need to show convergence of the truncated flows $\boldsymbol{\eta}(\vec{t'})$ and $\boldsymbol{\eta}(\vec{t''})$, i.e., we require that

$$\lim_{t' \to -\infty} P\left[|W_{t'} - W_{t''}| > 0\right] = 0 \tag{6.2.6}$$

Let n_0 be the length of the predecessor path associated with $W_{t''}$. Now

$$\begin{aligned} P\left(\bigcup_{k=n_0}^{\infty} A_k\right) &\leq \sum_{k=n_0}^{\infty} P(A_k) \\ &< \sum_{k=n_0}^{\infty} q^k \qquad (=) \text{ by } (6.2.5) \\ &< \frac{q^{n_0}}{1-q} \end{aligned}$$

$$\Longrightarrow \quad \forall\, \epsilon > 0 \ \exists\, N_0 = N_0(\epsilon) \text{ s.t. } n_0 > N_0 \Rightarrow P\left(\bigcup_{k=n_0}^{\infty} A_k\right) < \frac{\epsilon}{2} \tag{6.2.7}$$

which takes account of the case $n_0 \in \overline{\mathbb{N}}_{N_0}$, so we need to investigate $n_0 \in \mathbb{N}_{N_0}$.

For the latter case, let us consider the possibility of the time-extent of the path exceeding some T_0. By virtue of the fact that $\langle \tilde{\tau}_n \rangle$ is a sequence of proper random variables (interarrival times on the path corresponding to

$W_{t''}$), we have

$$\forall\, \epsilon > 0 \;\; \exists\, T_0 > 0 \text{ s.t. } P\left[\sum_{i=1}^{N_0} \widetilde{\tau}_i > T_0\right] < \frac{\epsilon}{2N_0}$$

$$\begin{aligned}
\Rightarrow \;\; P\left[\bigcup_{n_0=1}^{N_0} \left\{\sum_{i=1}^{n_0} \widetilde{\tau}_i > T_0\right\}\right] &\leq \sum_{n_0=1}^{N_0} P\left[\sum_{i=1}^{n_0} \widetilde{\tau}_i > T_0\right] \\
&\leq \sum_{n_0=1}^{N_0} P\left[\sum_{i=1}^{N_0} \widetilde{\tau}_i > T_0\right] \\
&< \frac{\epsilon}{2} \qquad\qquad (6.2.8)
\end{aligned}$$

Thus, on combining (6.2.7) and (6.2.8), we get that $\forall\, \epsilon > 0 \;\; \exists\, T_0 > 0$ s.t.

$$\begin{aligned}
P\left[|W_{t'} - W_{t''}| > 0\right] &\leq P\left[\bigcup_{n_0=1}^{N_0} \left\{\sum_{i=1}^{n_0} \widetilde{\tau}_i > T_0\right\}\right] + P\left(\bigcup_{n_0=N_0+1}^{\infty} A_{n_0}\right) \\
&< \frac{\epsilon}{2} + \frac{\epsilon}{2} \qquad\qquad (=) \\
&< \epsilon
\end{aligned}$$

Letting $t'' < t' < -T_0 < 0$ yields (6.2.6).

As a consequence of the symmetry of the starlike network model and the distribution assumptions about the arrival process to any source, $\boldsymbol{\eta}$ is seen to be time-stationary and symmetric under permutations of its indices. (For the former, consider the expression for $\widetilde{W}(y)$ i.t.o. stationary random variables, as well as the stationarity of the arrival and service processes.)

Also, we have established that the waiting time is derived on a path extending for a finite time and having finite length (w.p.1). This implies uniform control in t' of the memory of $\boldsymbol{\eta}(\overrightarrow{t'})$ which, in turn, means that $\boldsymbol{\eta}$ inherits the property that its waiting times satisfy a system of network equations of the form given in (6.2.1). In other words, under zero initial conditions, if we fix some realization $\boldsymbol{\omega}$, then $\exists T(\boldsymbol{\omega})$ which gives the earliest possible time for the generation of a predecessor path that affects our virtual arrival. As a result, the predecessor paths for $p_s(y)$ and $p_d(y)$ must also terminate 'before' $T(\boldsymbol{\omega})$ is reached in the reverse flow (otherwise y would have a path of greater time-extent, and this would lead to a contradiction). Hence the control in memory.

Thus far we have established the existence of $\boldsymbol{\eta}$ (associated with W), but now we need to consider its uniqueness. In order to prove that there is no other random variable $\widehat{W}$ (or its distribution) which solves the system of network equations of the form (6.2.1), with the associated flow $\hat{\boldsymbol{\eta}}$ having the stationarity and symmetry properties $\boldsymbol{\eta}$ does, then we need to show that the memory of an initial condition W^0 is "lost" (much as in Theorem 5.3.1). Briefly, this is a consequence of the bound

$$P\left[W^0 + \sum_{i=1}^{n}(S_i' - \tau_i') > 0\right] < \epsilon$$

which is a suitable adaptation of the inequality given in Lemma 6.2.1 and its corollaries, for sufficiently large n.

□

Remark 6.2.1

The most general way of expressing the bound analogous to (6.2.3) for any $m \in \mathbb{N}$ and τ not necessarily exponential[1] is

$$\inf_{a>0} \mathrm{E}e^{a(S-\tau)} < \frac{1}{m} \tag{6.2.9}$$

a bound which is *uniform* in N. (Recall the discussion in §4.2 about the effect — on the above quantity and on the traffic intensity — of having N finite.) If $\tau \sim \mathrm{Exp}(\lambda^{-1})$ *and* $S \sim \mathrm{Exp}(\mu^{-1})$, implying that we have $M/M/1/FAFS$ channels, then this leads to

$$\mathrm{E}e^{a(S-\tau)} = \mathrm{E}e^{-a\tau}\mathrm{E}e^{aS} = \psi_\lambda(ia)\psi_\mu(-ia) = \frac{\lambda}{(\lambda+a)}\frac{\mu}{(\mu-a)}$$

where we use the independence of S and τ, and $\psi_\beta(\cdot)$ is the characteristic function of an $\mathrm{Exp}(\beta^{-1})$ random variable. As a result

$$\lambda\mu \inf_{a>0}[(\lambda+a)(\mu-a)]^{-1} < \tfrac{1}{m} \tag{6.2.10}$$

It is easy to show that the infimum in (6.2.10) is equal to $\frac{4}{(\mu+\lambda)^2}$ (for $a = \frac{\mu-\lambda}{2}$ (positive because $\frac{\lambda}{\mu} < 1$)). As a result, we may simplify (6.2.10) as follows

$$\begin{aligned} & \tfrac{4\lambda\mu}{(\mu+\lambda)^2} < \tfrac{1}{m} \\ \Rightarrow\ & \tfrac{4\rho}{(1+\rho)^2} < \tfrac{1}{m} \\ \Rightarrow\ & \rho^2 + 2(1-2m)\rho + 1 > 0 \\ \Rightarrow\ & \left.\begin{array}{lll} \rho & < & (2m-1) - \sqrt{(2m-1)^2-1} \text{ or} \\ \rho & < & (2m-1) - 2\sqrt{m(m-1)} \end{array}\right\} \end{aligned} \tag{6.2.11}$$

since the other (larger) solution is invalid.

[1] Without a Poisson arrival stream, however, we cannot always guarantee the validity of the results of the various theorems in this chapter.

In Theorem 6.2.1, we were concerned with *infinite* limits w.r.t. time for a *finite* number N of channels in each of the $m = 2$ stages. The next theorem, in which the case $N \to \infty$ is considered, is proved for a general $m \geq 2$. We use the idea of an extended call, which was introduced in Chapter 4.

Theorem 6.2.2

Given the extended call $\overline{y} = [\overline{t}; (\overline{S}, \overline{\boldsymbol{i}}, \overline{W})]$, suppose

$$\inf_{a>0} \mathrm{E}e^{aS} \frac{\lambda}{\lambda + a} < \frac{1}{m} \tag{6.2.12}$$

and N tends to infinity. Then $\overline{W} \xrightarrow{d} W_{\min}$ where $W_{\min}$ is the minimal solution of (6.2.2).

<u>PROOF</u>

Consider a rooted ditree $\mathcal{T}$ with root r and branching m (Appendix A). With each edge e of $\mathcal{T}$ we associate an independent copy u_e of $u = S - \tau$ ($S \sim F, \tau \sim \mathrm{Exp}(\lambda^{-1})$). We are interested in (finite) dipaths with r as initial vertex. For each such (fixed) path $\pi \in \Pi_n$ (collection of paths of length n) write

$$U^*(\pi) = \sum_{e \in \pi} u_e$$

and also define

$$U_n^* = \max_{\pi \in \Pi_n} U^*(\pi)$$

$$W^* = \left[\sup_{n \in \mathbb{N}} U_n^*\right]^+ \tag{6.2.13}$$

Analogously to (6.2.5), we obtain

$$\begin{aligned} P[U_n^* > 0] &= P\left[\max_{\pi \in \Pi_n} U^*(\pi) > 0\right] \\ &< m^n \left(\mathrm{E}e^{au}\right)^n \\ \Longrightarrow \qquad P[U_n^* > 0] &\leq m^n \left(\inf_{a>0} \mathrm{E}e^{au}\right)^n \\ &\leq q^n \qquad\qquad q \in (0,1) \\ \Longrightarrow \quad P\left[\limsup_{n\to\infty} [U_n^* > 0]\right] &= 0 \end{aligned} \tag{6.2.14}$$

which implies that W^* is a.e. finite and $\sup_{n\in\mathbb{N}} U_n^*$ is attained a.e. We may also conclude that

$$P\left[\left[\sup_{k\in\mathbb{N}_n} U_k^*\right]^+ < W^*\right] \leq q^n \quad (q \in (0,1)) \tag{6.2.15}$$

Since we have established that W^* is a.e. finite, we may deduce that $L(W^*)$ provides the minimal solution to (6.2.2). This may be demonstrated informally (with a slight abuse of notation) as follows:

$$\begin{aligned}
W^* &\sim \max_{j\in\mathbb{N}_m}\left[\left[\max\sum_{i=1}^{1}u_i^{(j)}\right]^+,\left[\sup_{n\in\mathbb{N}}\max\sum_{i=1}^{n+1}u_i^{(j)}\right]^+\right] \\
&\sim \max_{j\in\mathbb{N}_m}\left[\sup_{n=1}\max\sum_{i=1}^{n}u_i^{(j)},\ \sup_{n\in\overline{\mathbb{N}}_1}\max\sum_{i=1}^{n}u_i^{(j)},0\right] \\
&\sim \max_{j\in\mathbb{N}_m}\left[\sup_{n\in\mathbb{N}}\max\sum_{i=1}^{n}u_i^{(j)}\right]^+ \\
&\sim \max_{j\in\mathbb{N}_m}\left[\sup_{n\in\mathbb{N}}\left[\max\sum_{i=2}^{n}u_i^{(j)}\right]+u_1^{(j)}\right]^+ \\
&\sim \max_{j\in\mathbb{N}_m}\left[\max\left[\sup_{n=1}\sum_{i=2}^{n}u_i^{(j)},\ \sup_{n\in\overline{\mathbb{N}}_1}\sum_{i=2}^{n}u_i^{(j)}\right]+u_1^{(j)}\right]^+ \\
&\sim \max_{j\in\mathbb{N}_m}\left[\max\left[0,\ \sup_{n\in\overline{\mathbb{N}}_1}\sum_{i=2}^{n}u_i^{(j)}\right]+u_1^{(j)}\right]^+ \\
&\sim \max_{j\in\mathbb{N}_m}\left[\left[\sup_{n\in\overline{\mathbb{N}}_1}\sum_{i=2}^{n}u_{i-1}^{(j)}\right]^+ + u_1^{(j)}\right]^+ \\
&\sim \max_{j\in\mathbb{N}_m}\left[\left[\sup_{n\in\overline{\mathbb{N}}_1}\sum_{i=1}^{n-1}u_i^{(j)}\right]^+ + u_1^{(j)}\right]^+ \\
&\sim \max_{j\in\mathbb{N}_m}\left[\left[\sup_{n\in\mathbb{N}}\sum_{i=1}^{n}u_i^{(j)}\right]^+ + u_1^{(j)}\right]^+ \\
&\sim \max_{j\in\mathbb{N}_m}\left[W^* + u^{(j)}\right]^+
\end{aligned}$$

where the IID characteristics of the $\left\langle u_i^{(j)}\right\rangle$ are used.

The minimality is essentially a consequence of a zero initial condition.

Our intention now is to show that the limiting predecessor digraph (i.e., limiting in the sense $N \to \infty$) assumes the form of the ditree $\mathcal{T}$.

It is clear that definitions and calculations analogous to (6.2.13), (6.2.14) and (6.2.15) (or (6.2.4) and (6.2.5)) may be formulated (with the bar notation). In particular we have

$$P\left[\left[\sup_{k\in\mathbb{N}_n} \overline{U}_k\right]^+ < \overline{W}\right] \leq q^n \quad (q \in (0,1)) \tag{6.2.16}$$

Two major differences between the predecessor digraph for finite N and the ditree $\mathcal{T}$ are (as was indicated in Theorem 6.2.1 for $m = 2$, and elsewhere) that in the former case, the following are feasible:

(a) $p_j(y)$ and $p_k(y)$ are from a common flow(s); possibly they even coincide

(b) $\tilde{y}$ is a predecessor for more than one call

If a predecessor digraph is constructed for some call y and we label the level of the kth predecessors ($k \in \mathbb{Z}_\oplus$) of y as the kth level, then the maximum number of (distinct) calls at the kth level is m^k. On "truncation" of the predecessor digraph at the nth level, we would encounter at most $\sum_{k=0}^{n-1} m^k = \frac{m^n-1}{m-1}$ calls as vertices of that truncated graph *having immediate predecessors on the graph* and at most $\frac{m}{m-1}(m^n-1)$ calls with immediate successors on the graph.

The probability of one of (a) or (b) occurring under the given circumstances is thus $O\left(\frac{c_n}{N}\right)$, where c_n is a constant depending on n.

Since $O\left(\frac{c_n}{N}\right) \xrightarrow[\infty]{N} 0$ (implying that the truncated predecessor graph tends to a segment of $\mathcal{T}$) and because of (6.2.15) and (6.2.16), the convergence in distribution of $\overline{W}$ to W^* may be framed i.t.o. $\left[\sup_{k\in\mathbb{N}_n} \overline{U}_k\right]^+ \xrightarrow[N\to\infty]{d} \left[\sup_{k\in\mathbb{N}_n} U_k^*\right]^+$.
Thus we have the desired result.

□

Remark 6.2.2
The "Big Oh" notation which appears in the above proof for sequences is explained in greater detail in the next section for the case of functions.

6.3 MORE ON SOLUTIONS OF THE LINDLEY EQUATION

Later in this section we shall see that the minimal solution given in (6.2.13) is not the unique solution to (6.2.2) ($N \to \infty$). In anticipation of this, we introduce some additional concepts.

We replace those edges $e(\pi)$ at the nth level of our ditree $\mathcal{T}$ by random variables $\kappa_{e(\pi)}$ that serve as initial conditions, and which are assumed to be

IID with distribution ψ independently of $\pi \in \Pi_n$. Write

$$U_n^{[\psi]}(\pi) = \begin{cases} U(\pi) & \text{if } \pi \in \Pi_k, \quad k \in \mathbb{N}_{n-1} \\ U(\pi_{(e)}) + \kappa_{e(\pi)} & \pi \in \Pi_n \end{cases} \tag{6.3.1}$$

where $\pi_{(e)}$ is $\pi \in \Pi_n$ with $e(\pi)$ removed. Next let

$$W_n^{[\psi]} = \left[\max_{\pi \in \bigcup_{k \in \mathbb{N}_n} \Pi_k} U_n^{[\psi]}(\pi) \right]^+ \sim G_n^{[\psi]} \tag{6.3.2}$$

Should the limiting distribution

$$G^{[\psi]} = \lim_{n \to \infty} G_n^{[\psi]} \tag{6.3.3}$$

exist, then that distribution yields a solution to (6.2.2) under assumptions of a general distribution for $\langle \tau^{(j)} \rangle$.

If

$$\inf_{a>0} \mathrm{E}e^{a(S-\tau)} \;\le\; \tfrac{1}{m} \tag{6.3.4}$$

(compare (6.2.9)), then we may define the following quantities:

$$\alpha \;=\; \inf[a > 0 : \mathrm{E}e^{a(S-\tau)} \le \tfrac{1}{m}] \tag{6.3.5a}$$

$$\beta \;=\; \sup[b > 0 : \mathrm{E}e^{b(S-\tau)} \le \tfrac{1}{m} \wedge \neg(\exists\, a \in (\alpha, b) \text{ s.t. } \mathrm{E}e^{a(S-\tau)} > \tfrac{1}{m})] \tag{6.3.5b}$$

If (6.3.4) holds *and* the infimum (in (6.3.4)) is attained, then

$$\mathrm{E}e^{a(S-\tau)} \;\le\; \tfrac{1}{m} \quad \forall\, a \in \begin{cases} [\alpha, \beta] & \text{if } \beta < \infty \\ [\alpha, \infty) & \text{if } \beta = \infty \end{cases} \tag{6.3.6a}$$

Similarly, if (6.2.9) is satisfied, then

$$\mathrm{E}e^{a(S-\tau)} \;<\; \tfrac{1}{m} \quad \forall\, a \in (\alpha, \beta) \tag{6.3.6b}$$

Our next topic of interest pertains to a class of exponential-type distributions.

Let $H(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases}$ which, of course, is the distribution function of $\phi_0(x) = \begin{cases} 1 & x = 0 \\ 0 & x \neq 0 \end{cases}$. Then a succinct way of writing the distribution function of an exponential random variable with mean λ^{-1} is

$$\Psi_\lambda(x) = H(x)\left[1 - e^{-\lambda x}\right] \tag{6.3.7}$$

which may be generalised to the mixture $\text{Exp}(\lambda^{-1}, p)$ with the distribution function

$$\begin{aligned} \Psi_{(\lambda,p)}(x) &= p\Psi_\lambda(x) + (1-p)H(x) \\ &= H(x)\left[p(1 - e^{-\lambda x}) + (1-p)\right] \qquad (6.3.8) \\ \Rightarrow \quad \Psi_{(\lambda,p)}(x) &= H(x)[1 - pe^{-\lambda x}] \qquad\qquad p \in [0,1] \end{aligned}$$

Clearly the $\text{Exp}(\lambda^{-1}, 1)$ and $\text{Exp}(\lambda^{-1})$ distributions coincide, and the distribution corresponding to $\text{Exp}(\lambda^{-1}, 0)$ is simply $H(x)$. Moreover, it is a well-known fact, for example, that the $M(\lambda^{-1})/M(\mu^{-1})/1$ queue has a stationary waiting-time distribution of the type $\text{Exp}((\mu - \lambda)^{-1}, \rho)$ $(\rho < 1)$.

Return to (6.2.2), and let

$$\left.\begin{aligned} X &\sim W + u \\ X^{(j)} &\sim W^{(j)} + u^{(j)} \sim X \end{aligned}\right\} \tag{6.3.9}$$

Then (6.2.2) reduces to

$$X \sim u + \max_{j \in \mathbb{N}_m} [X^{(j)}]^+ \tag{6.3.10}$$

where $\left\langle X^{(j)}, j \in \mathbb{N}_m \right\rangle$ are IID, $\sim \mathcal{F}$ independently of $u \sim K$.

Now

$$\begin{aligned} P\left[[X^{(j)}]^+ \leq x\right] &= P[\max[0, X^{(j)}] \leq x] \\ &= P[\{X^{(j)} \leq x\} \cap \{x \geq 0\}] \\ &= (\mathcal{F}H)(x) \end{aligned}$$

Consequently

$$\begin{aligned} P\left[\max_{j \in \mathbb{N}_m} [X^{(j)}]^+ \leq x\right] &= P\left[\bigcap_{j \in \mathbb{N}_m} \{[X^{(j)}]^+ \leq x\}\right] \\ &= [\mathcal{F}H]^m(x) \qquad \text{by independence} \\ &= (\mathcal{F}^m H)(x) \end{aligned}$$

Hence (6.3.10) may be expressed i.t.o. distribution functions, as follows:

$$\begin{aligned} \mathcal{F}(x) &= K * [(\mathcal{F}H)(x)]^m \\ &= \int_{-\infty}^{\infty} H(x-y)[\mathcal{F}(x-y)]^m dK(y) \end{aligned} \tag{6.3.11}$$

We observe that, since $W \sim \max_{j \in \mathbb{N}_m} [X^{(j)}]^+$, we have

$$L(W) = [\mathcal{F}H]^m \tag{6.3.12}$$

As a consequence of this relationship between $\mathcal{F}$ and the distribution of W, many of the results further on in this section will be framed i.t.o. $\mathcal{F}$. The associated result for W (or its distribution) will (tacitly) be implied.

In the current section, the Landau symbol O appears fairly often. We therefore recall the following definition.

Definition 6.3.1 ("Big Oh")

Let f and g be real functions of a single variable. Then f *is (at most) of the order of* g, written $f = O(g)$, as $x \to \infty$ iff

$$\exists\, h > 0,\ X \text{ s.t. } |f(x)| \leq h|g(x)| \ \ \forall\, x > X \qquad (6.3.13)$$

A similar definition may be formulated for the more general instance of $x \to a$ ($a \in \mathbb{R}^*$) but we will not need it here. In fact, our specific interest is in

$$f(x) \ = \ O\left(e^{-ax}\right) \ \ (x \to \infty) \qquad (6.3.14)$$

In this case it is more convenient to write (6.3.13) as

$$\exists\, h > 0,\ X \text{ s.t. } |f(x)|e^{ax} \leq h \ \ \forall\, x > X \qquad (6.3.15)$$

The proof of the following lemma is easy and is, therefore, omitted.

Lemma 6.3.1 (Properties of $O\left(e^{-ax}\right)$)

Let f be as in (6.3.14). Assume in addition that $c \neq 0$ and $f_i(x) = O\left(e^{-a_i x}\right)$ $(i = 1, 2)$. Then

(a) $cf(x) = O\left(e^{-ax}\right)$

(b) (i) $f_1(x) + f_2(x) = O\left(e^{-a_1 x} + e^{-a_2 x}\right)$

(ii) $f_1(x) + f_2(x) = O\left(e^{-\underline{a}x}\right)$ where $\underline{a} = \min[a_1, a_2]$

(c) $f_1(x)f_2(x) = O\left(e^{-(a_1+a_2)x}\right)$

(d) $f(x) = O\left(e^{-bx}\right)\ \forall\, b \leq a$

Remark 6.3.1

$f(x) = O\left(e^{-ax}\right)\ \ (x \to \infty)$ iff

$$\exists\, h > 0,\ X \text{ and } \phi \text{ s.t. } f(x) = \phi(x)e^{-ax} \text{ where } |\phi(x)| \leq h\ \forall\, x > X \qquad (6.3.16)$$

(6.3.16) is sufficient since $|\phi(x)e^{-ax}|e^{ax} = |\phi(x)| \leq h\ \forall\, x > X$ and necessary because $\exists\, h > 0,\ X$ s.t. $h \geq |f(x)|e^{ax} = |[f(x)e^{ax}]e^{-ax}|e^{ax} \overset{\text{def}}{=} |\phi(x)|\ \ \forall\, x > X$.

We shall shortly make use of the inequalities given in the next lemma.

Lemma 6.3.2

Suppose $r \in \mathbb{N}$, $b \le 1$ and $|b_i| \le 1$ $(i = 1, 2)$. Then

(a) $1 - rb \le (1-b)^r$

(b) $|b_1^r - b_2^r| \le r|b_1 - b_2|$

PROOF

The results are obvious if $b = 0$ or $r = 1$ so in the sequel we do not consider these cases.

(a) Use induction on r to prove this result, *Bernoulli's inequality.* (Refer to Bronshtein & Semendyayev [18, *p.* 218].)

(b)

$$\begin{aligned} |b_1^r - b_2^r| &= \left|(b_1 - b_2)\sum_{k=0}^{r-1} b_1^k\, b_2^{r-1-k}\right| \\ &= |b_1 - b_2| \sum_{k=0}^{r-1} |b_1|^k |b_2|^{r-1-k} \\ &\le |b_1 - b_2| \sum_{k=0}^{r-1} 1 \qquad (=) \\ &\le r|b_1 - b_2| \end{aligned}$$

□

In view of (6.3.11) and forthcoming theory, we consider the following lemma.

Let $b > 0$, $b_i > 0$ $(i = 1, 2)$ and $x \in \mathbb{R}$. Write

$$\begin{aligned} 0 \;\le\; f(b,x) &= \int_{-\infty}^{x} e^{by} dK(y) \\ &= \int_{\mathbb{R}} H(x-y) e^{by} dK(y) \qquad (6.3.17) \\ &\le \mathrm{E} e^{bu} \end{aligned}$$

$$\begin{aligned} 0 \;\le\; \overline{f}(b,x) &= \int_{x}^{\infty} e^{by} dK(y) \\ &= \int_{\mathbb{R}} e^{by} dK(y) - \int_{-\infty}^{x} e^{by} dK(y) \qquad (6.3.18) \\ &= \mathrm{E} e^{bu} - f(b,x) \end{aligned}$$

Lemma 6.3.3

(a)

$$e^{bx}\overline{K}(x) \le \overline{f}(b,x) \qquad (6.3.19)$$

where, as usual,

$$\overline{K}(x) = 1 - K(x) \text{ and } K(x) = \int_{-\infty}^{x} dK(y) = \int_{\mathbb{R}} H(x-y) dK(y)$$

(b) $b_1 < b_2 \Rightarrow e^{-b_1 x}\overline{f}(b_1, x) \le e^{-b_2 x}\overline{f}(b_2, x)$

(c) If $a \in (\alpha, \beta)$ and (6.3.6b) holds, then

(i) $\overline{K}(x) = O\left(e^{-ax}\right)$

(ii) $e^{-\alpha x}\overline{f}(\alpha, x) = O\left(e^{-ax}\right)$

PROOF

(a)

$$\begin{aligned}\overline{f}(b,x) &= \int_x^{\infty} e^{by} dK(y) \\ &\geq e^{bx}\int_x^{\infty} dK(y) \quad (=) \quad e^{by} \geq e^{bx} \text{ for } y \in [x,\infty) \\ &\geq e^{bx}\overline{K}(x)\end{aligned}$$

(b) $e^{-b_2 x}\overline{f}(b_2,x) - e^{-b_1 x}\overline{f}(b_1,x) = \int_x^{\infty} \left[e^{b_2(y-x)} - e^{b_1(y-x)}\right] dK(y) \geq 0$

(c) (i)

$$\begin{aligned}|\overline{K}(x)|e^{ax} &\leq \overline{f}(a,x) && (=) \text{ by } (6.3.19) \\ &\leq \mathrm{E}e^{au} - f(a,x) && \text{by } (6.3.18) \\ &\leq \tfrac{1}{m} && \text{by } (6.3.6b)\end{aligned}$$

(ii)

$$\begin{aligned}\left|e^{-\alpha x}\overline{f}(\alpha,x)\right| e^{ax} &\leq \left|e^{-ax}\overline{f}(a,x)\right| e^{ax} && (=) \text{ by (b)} \\ &\leq \overline{f}(a,x) \\ &\leq \tfrac{1}{m} && \text{as in (i)}\end{aligned}$$

hence the results.

□

The form of (6.3.11) also suggests the appropriateness of using the nonlinear operator

$$\Lambda: \mathcal{F} \mapsto K * [\mathcal{F}H]^m \tag{6.3.20}$$

It will be convenient, therefore, to investigate some of the properties of Λ *on the class $\mathcal{C}$ of (possibly improper) distribution functions.* (Clearly $\Lambda: \mathcal{C} \mapsto \mathcal{C}$.)

Lemma 6.3.4 (Monotonicity of Λ)

The operator Λ is nondecreasing in the following sense:

$$0 \leq [\mathcal{F}H](x) \leq [\mathcal{F}^*H](x) \quad \Rightarrow \quad 0 \leq \Lambda\mathcal{F}(x) \leq \Lambda\mathcal{F}^*(x) \quad x \in \mathbb{R} \tag{6.3.21}$$

<u>PROOF</u>

$$\begin{aligned}
(\Lambda\mathcal{F}^*)(x) &= \int_{\mathbb{R}} [\mathcal{F}^*H]^m(x-y)dK(y) \\
&\geq \int_{\mathbb{R}} [\mathcal{F}H]^m(x-y)dK(y) \quad (=) \quad \text{monotonicity of integrals} \\
&\geq (\Lambda\mathcal{F})(x) \\
&\geq 0
\end{aligned}$$

$\square$

Lemma 6.3.5

If $\mathrm{E}e^{au} < \frac{1}{m}$ $\forall a \in (\alpha, \beta)$ as in (6.3.6b), $h > 0$ and

$$|\mathcal{F}(x) - \mathcal{F}^*(x)|e^{ax} \leq h \tag{6.3.22}$$

Then

$$|\Lambda\mathcal{F}(x) - \Lambda\mathcal{F}^*(x)|e^{ax} \leq bh \tag{6.3.23}$$

where $b = m\mathrm{E}e^{au} < 1$.

<u>Proof</u>

$$
\begin{array}{rll}
 & |[\mathcal{F}(x)]^m - [\mathcal{F}^*(x)]^m| & \\
\leq & m|\mathcal{F}(x) - \mathcal{F}^*(x)| & \text{by Lemma 6.3.2(b)} \\
\leq & mhe^{-ax} & \text{by (6.3.22)} \\
\Longrightarrow & |\Lambda\mathcal{F}(x) - \Lambda\mathcal{F}^*(x)| & \\
= & |K * [\mathcal{F}H]^m(x) - K * [\mathcal{F}^*H]^m(x)| & \\
\leq & K * |[\mathcal{F}(x)]^m - [\mathcal{F}^*(x)]^m|\, H(x) \qquad (=) & \\
\leq & \int_{\mathbb{R}} H(x-y)\, |[\mathcal{F}(x-y)]^m - [\mathcal{F}^*(x-y)]^m|\, dK(y) & \\
\leq & mhe^{-ax} \int_{-\infty}^{\infty} e^{ay} H(x-y) dK(y) & \text{by above} \\
\leq & mhe^{-ax}\mathrm{E}e^{au} & \text{by (6.3.17)} \\
\Longrightarrow & |\Lambda\mathcal{F}(x) - \Lambda\mathcal{F}^*(x)|e^{ax} & \\
\leq & (m\mathrm{E}e^{au})\, h &
\end{array}
$$

□

Remark 6.3.2

In Lemma 6.3.5, we are able to apply Lemma 6.3.2(b) because $\mathcal{F}$ and $\mathcal{F}^*$ are distribution functions.

Corollary 6.3.5*a*

If $\mathrm{E}e^{au} < \frac{1}{m}\ \forall a \in (\alpha, \beta)$ and

$$\Lambda\mathcal{F}(x) \quad = \quad \mathcal{F}(x) + O\left(e^{-ax}\right) \quad (x \to \infty) \tag{6.3.24}$$

then

(a) the sequence $\langle \mathcal{F}_n = \Lambda^n \mathcal{F}, n \in \mathbb{Z}_\oplus \rangle$ converges pointwise (to some distribution function $\tilde{\mathcal{F}}$)

(b)
$$\tilde{\mathcal{F}}(x) \quad = \quad \mathcal{F}(x) + O\left(e^{-ax}\right) \quad (x \to \infty) \tag{6.3.25}$$

<u>PROOF</u>

By virtue of (6.3.15) and (6.3.24)

$$\begin{aligned} &\exists\, h > 0, X \text{ s.t. } && |\Lambda\mathcal{F}(x) - \mathcal{F}(x)| e^{ax} \leq h \quad \forall\, x > X \\ &\Rightarrow && \left|\Lambda^{j+1}\mathcal{F}(x) - \Lambda^j \mathcal{F}(x)\right| e^{ax} \leq b^j h \quad \forall\, x > X \quad (j \in \mathbb{Z}_\oplus) \end{aligned} \tag{6.3.26}$$

by an inductive extension of Lemma 6.3.5.

(a) Next let $k, l \in \mathbb{N}$ and suppose w.l.g. that $k \leq l$. Then for $x > X$

$$\begin{aligned}
|\mathcal{F}_l(x) - \mathcal{F}_k(x)|e^{ax} &= \left|\sum_{j=k}^{l-1}[\Lambda^{j+1}\mathcal{F}(x) - \Lambda^j\mathcal{F}(x)]\right| e^{ax} \\
&\leq \sum_{j=k}^{l-1}|\Lambda^{j+1}\mathcal{F}(x) - \Lambda^j\mathcal{F}(x)|e^{ax} \\
&\leq h\sum_{j=k}^{l-1} b^j \qquad (=) \qquad \text{by (6.3.26)} \\
&\leq \frac{h(b^k - b^l)}{1-b}
\end{aligned}$$

and apply Cauchy's convergence criterion (for each fixed x).

(b) For $x > X$

$$\begin{aligned}
|\tilde{\mathcal{F}}(x) - \mathcal{F}(x)|e^{ax} &= \left|\sum_{j=0}^{\infty}[\mathcal{F}_{j+1}(x) - \mathcal{F}_j(x)]\right| e^{ax} \\
&\leq h\sum_{j=0}^{\infty} b^j \qquad (=) \\
&\leq \frac{h}{1-b}
\end{aligned}$$

$$\Longrightarrow \qquad \tilde{\mathcal{F}}(x) = \mathcal{F}(x) + O\left(e^{-ax}\right) \qquad \text{by (6.3.15)}$$

□

Lemma 6.3.6

Suppose that the distribution function $\mathcal{F}$ may be represented as

$$\mathcal{F}(x) \;=\; 1 - ce^{-\alpha x} + O\left(e^{-ax}\right) \quad (x \to \infty) \tag{6.3.27}$$

where $c \geq 0$ and $0 < \alpha < a < \min[2\alpha, \beta]$. Then

$$\Lambda\mathcal{F}(x) \;=\; \mathcal{F}(x) + O\left(e^{-ax}\right) \tag{6.3.28}$$

<u>Proof</u>

$$\begin{aligned}
\mathcal{F}(x) &= 1 + O\left(e^{-\alpha x}\right) + O\left(e^{-ax}\right) \\
&= 1 + O\left(e^{-\alpha x}\right) && \text{by Lemma 6.3.1} \\
\Longrightarrow \quad [\mathcal{F}(x)]^m &= 1 + mO\left(e^{-\alpha x}\right) + \sum_{j=2}^{m} \binom{m}{j} \left[O\left(e^{-\alpha x}\right)\right]^j \\
&= 1 - mce^{-\alpha x} + \sum_{j=2}^{m} \binom{m}{j} O\left(e^{-j\alpha x}\right) && \text{by Lemma 6.3.1(c) \& (6.3.27)} \\
&= 1 - mce^{-\alpha x} + O\left(e^{-ax}\right)
\end{aligned} \tag{6.3.29}$$

where, in the last step, we use the restriction $a < 2\alpha$ in conjunction with Lemma 6.3.1(d), as well as parts (a) and (b) of that lemma.

(6.3.29), together with Remark 6.3.1, implies that $\exists\, h' > 0, X$ and Φ s.t. $[\mathcal{F}(x)]^m = 1 - mce^{-\alpha x} + \Phi(x)e^{-ax}$ where $|\Phi(x)| \leq h'\ \forall x > X$ but $\mathcal{F}(x) \in [0, 1]$

and $a > 0$ so $|\Phi(x)| \leq h \; \forall \, x > 0$ (where h is some positive number). Using the notation of (6.3.17), (6.3.18) and Lemma 6.3.3, we have

$$\begin{aligned}
& \Lambda\mathcal{F}(x) \\
= & \int_{-\infty}^{x} [\mathcal{F}(x-y)]^m dK(y) \\
= & \int_{-\infty}^{x} \left[1 - mce^{-\alpha(x-y)} + \Phi(x-y)e^{-a(x-y)}\right] dK(y) && \text{by (6.3.29)} \\
= & K(x) - mce^{-\alpha x} f(\alpha, x) + e^{-ax} \int_{-\infty}^{x} \Phi(x-y) e^{ay} dK(y) \\
= & [1 - \overline{K}(x)] - mce^{-\alpha x} \left[\mathrm{E}e^{\alpha u} - \overline{f}(\alpha, x)\right] + O\left(e^{-ax}\right) \qquad (=) \\
= & 1 - mce^{-\alpha x} \left(\tfrac{1}{m}\right) + \left[mce^{-\alpha x}\overline{f}(\alpha, x) - \overline{K}(x) + O\left(e^{-ax}\right)\right] && \text{by (6.3.5a)} \\
= & 1 - ce^{-\alpha x} + O\left(e^{-ax}\right)
\end{aligned}$$

where, in the last step, we use Lemmata 6.3.3(c) and 6.3.1.

$\square$

In the course of proving the next result (Theorem 6.3.1) we will make use of Siegel's Theorem which appears in Devaney [28, §3.4], where a proof may be found. (Siegel's Theorem is also stated and proved (in a different form) in Arnold [2, Chapter 5].) For the sake of completeness, the statement of this theorem is given here.

Let $P(z)$ be an analytic function satisfying $P(0) = 0$ and $P'(0) = \lambda$ with $0 < |\lambda| < 1$. Then there is a neighbourhood U of 0 and an analytic map $H: U \to \mathbb{C}$ such that

$$P \circ H(z) = H \circ L(z)$$

where $L(z) = \lambda z$.

Theorem 6.3.1

(a) Suppose (6.3.4) holds.

(i) (6.2.13) provides the minimal solution to (6.2.2).

(ii) $\alpha < \infty \Rightarrow \forall\, \nu \in [\alpha, \beta]$ (refer (6.3.5)) and $\forall\, p \in [\frac{1}{m}, 1]$, $\exists$ limiting distributions $G^{[\Psi_\nu]}$ and $G^{[\Psi_{(\nu,p)}]}$ giving solutions to (6.2.2).

(b) Assume now that (6.2.9) is satisfied. Then

(i) the supremum on the right of (6.2.13) is attained w.p.1.

(ii) For each fixed $c \in \mathbb{R}_\oplus$, let $\mathcal{F}(c)$ denote the class of distribution functions of the form

$$\mathcal{F}(x) \quad = \quad 1 - ce^{-\alpha x} + O\left(e^{-ax}\right) \quad (x \to \infty) \tag{6.3.30}$$

where $\alpha < a \;(= a(\mathcal{F})) \; < \min[2\alpha, \beta]$

(1) For fixed c, $\mathcal{F}(c)$ contains exactly one solution, $G^{(c)}$ so that the limiting distribution $G^{[\psi]}$ in (6.3.3) exists and coincides with $G^{(c)}$.

(2) $\left\{G^{(c)}; c \geq 0\right\}$ forms a linear strictly-ordered family of distributions and $G^{(0)}$ coincides with G, the minimal solution.

(c) Finally, suppose

$$\inf_{a>0} \mathrm{E}e^{a(S-\tau)} > \frac{1}{m} \tag{6.3.31}$$

If $P[u > 0] > 0$, then (6.2.2) has no solution.

PROOF

At this stage, we make an observation that if $P[u \leq 0] = 1$, then (6.3.4) becomes unnecessary. That a solution to (6.2.2) exists is intuitive (consider this equation or (6.2.13)). In view of this fact, when we wish to prove the sufficiency of (6.3.4), we need only consider cases for which $P[u > 0] > 0$ The infimum on the left-hand side of (6.3.4) is then attained.

(a) (i) If we use a method similar to that employed before, i.e., edges of $\mathcal{T}$ are associated with random variables (in this case u and X), then (6.3.20) summarises the single-step approach to the root of the tree. (6.2.13) corresponds to the limit (as $n \to \infty$) in the iterative procedure

$$\mathcal{F}_{n+1} = \Lambda \mathcal{F}_n \quad n \in \mathbb{Z}_{\oplus} \tag{6.3.32}$$

subject to $\mathcal{F}_0 = H$. Fixed points of Λ are, therefore, of interest.

In this case of $\mathcal{F}_0 = H$, we have

$$
\begin{aligned}
(\mathcal{F}_1 H)(x) &= [\Lambda \mathcal{F}_0(x)][H(x)] \\
&= \begin{cases} \Lambda H(x) = \int_{\mathbb{R}} H(x-y)dK(y) & \\ \qquad\quad = K(x) \le 1 = H(x) & x \ge 0 \\ \qquad\ 0 = H(x) & x < 0 \end{cases}
\end{aligned}
$$

$$
\begin{aligned}
\Longrightarrow \quad 0 &\le (\mathcal{F}_1 H)(x) \le (\mathcal{F}_0 H)(x) \\
\Longrightarrow \quad 0 &\le \mathcal{F}_n(x) \quad\ \le \mathcal{F}_0(x) = H(x) \le 1
\end{aligned}
\tag{6.3.33}
$$

by an (inductive) extension of (6.3.21). It follows that for each x, the sequence $\langle \mathcal{F}_n(x), n \in \mathbb{Z}_{\oplus} \rangle$ is monotonic and bounded so that the (pointwise) limit $\mathcal{F}(x) = \lim_{n\to\infty} \mathcal{F}_n(x) \in [0,1]$ exists, and inherits the monotonic property (in x) of the sequence.

If $\mathcal{F}$ is to be a proper distribution function and hence a minimal solution of (6.3.10), then (since (6.3.33) gives us $\lim_{x\to-\infty} \mathcal{F}(x) = 0$) it is sufficient that $\lim_{x\to\infty} \mathcal{F}(x) = 1$, a fact we prove in (ii).

(ii) If, in the iterative procedure, $\mathcal{F}_0(x) = \mathcal{F}_0^{(\nu,p)}(x) = \Psi_{(\nu,p)}(x)$, $\nu \in [\alpha, \beta]$, then

$$x < 0 \Rightarrow \mathcal{F}_0^{(\nu,p)}(x) = 0 \le \mathcal{F}_1^{(\nu,p)}(x)$$

For $x \geq 0$, we need to remember the jump discontinuity of $1-p$:

$$\begin{aligned}
& \mathcal{F}_1^{(\nu,p)}(x) \\
= \;& \Lambda \mathcal{F}_0^{(\nu,p)}(x) \\
= \;& K * \left[\mathcal{F}_0^{(\nu,p)}(x)\right]^m \\
= \;& \int_{\mathbb{R}} H(x-y)[1-pe^{-\nu(x-y)}]^m dK(y) + (1-p)^m K(x) \\
\geq \;& \int_{-\infty}^{x} [1-mpe^{-\nu(x-y)}] dK(y) + (1-p)^m K(x) && (=) \text{ Lem. 6.3.2} \\
\geq \;& \int_{-\infty}^{x} dK(y) - mpe^{-\nu x} f(\nu, x) + (1-p)^m K(x) && (=) \\
\geq \;& K(x) + mpe^{-\nu x}[\overline{f}(\nu, x) - \mathrm{E}(e^{\nu u})] + (1-p)^m K(x) \\
\geq \;& K(x) + mpe^{-\nu x}\left[e^{\nu x}\overline{K}(x) - \tfrac{1}{m}\right] + (1-p)^m K(x) && \text{by (6.3.19);}\ \nu \in [\alpha, \beta] \\
\geq \;& K(x) + mp\overline{K}(x) - pe^{-\nu x} + (1-p)^m K(x)
\end{aligned}$$

Consequently

$$\begin{aligned}
& \mathcal{F}_1^{(\nu,p)}(x) - \mathcal{F}_0^{(\nu,p)}(x) \\
\geq \;& K(x) + mp\overline{K}(x) - pe^{-\nu x} - (1 - pe^{-\nu x}) + (1-p)^m K(x) && (=) \\
\geq \;& \overline{K}(x)(mp-1) + (1-p)^m K(x)
\end{aligned}$$

Now the second term is nonnegative because each of its factors are; if $p \geq \frac{1}{m}$, then the first term is also nonnegative. Thus $p \in \mathcal{P} = [\frac{1}{m}, 1]$ is a sufficient, but not a necessary, condition for $\mathcal{F}_1^{(\nu,p)}(x) \geq \mathcal{F}_0^{(\nu,p)}(x)$.

It is interesting to note that in Karpelevich *et al.* [35], a "sufficient interval" for p of $\mathcal{P}' = [1 - \frac{1}{m}, 1]$ was obtained. For $m \in \overline{\mathbb{N}}_1$, we, of course, have that $\mathcal{P}' \subseteq \mathcal{P}$. For $m = 1$, $\mathcal{P}$ is the degenerate interval $[1, 1]$ but $\mathcal{P}'$ is the entire interval $[0, 1]$, which is not valid[2] because we have already seen that the inequality is in the opposite direction for $\mathcal{F}_0(x) = H(x)$. It is also worthwhile to note that $\mathcal{P}$ and $\mathcal{P}'$ display precisely the opposite behaviour when $m \to \infty$, i.e., $\mathcal{P}'$ contracts to the single point 1 and $\mathcal{P} \to (0, 1]$.

The fact that for large m, $\mathcal{P}$ is almost the entire interval $[0, 1]$ is not unreasonable. As $m \to \infty$, the discrepancy between the maximum, $\max\limits_{j \in \mathbb{N}_m} \left[W^{(j)} + u^{(j)}\right]$, and the sum, $\sum\limits_{j \in \mathbb{N}_m} \left[W^{(j)} + u^{(j)}\right]$, becomes greater. Allowing p to assume more values is, therefore, sensible.

Finally we note that $\forall\, m \;\; 1 \in \mathcal{P}$ and $1 \in \mathcal{P}'$. This implies that, independently of m, if $\mathcal{F}_0(x) = \mathcal{F}_0^{(\nu,1)}(x) = H(x)\,[1 - e^{-\nu x}]$, then $\mathcal{F}_n(x) \geq \mathcal{F}_{n-1}(x) \; \forall\, n \in \mathbb{N}$.

To summarise

$$0 \;\leq\; \mathcal{F}_0^{(\nu,p)}(x) \;\leq\; \mathcal{F}_n^{(\nu,p)}(x) \;\leq\; 1 \quad \forall\, n \in \mathbb{N} \quad \left\{ \begin{array}{c} \nu \in [\alpha, \beta] \\ p \in [\frac{1}{m}, 1] \end{array} \right\} \tag{6.3.34}$$

[2] although, admittedly, we are mostly concerned with $m > 1$

in which case $\mathcal{F}^{(\nu,p)}(x) = \lim_{n\to\infty} \mathcal{F}_n^{(\nu,p)}(x)$ exists. Thus

$$
\begin{aligned}
1 \;=\; \lim_{x\to\infty}\lim_{n\to\infty} \mathcal{F}_0^{(\nu,p)}(x) \;&\le\; \lim_{x\to\infty}\lim_{n\to\infty} \mathcal{F}_n^{(\nu,p)}(x) \\
&=\; \lim_{x\to\infty} \mathcal{F}^{(\nu,p)}(x) \qquad \le\; 1
\end{aligned}
\tag{6.3.35}
$$

In addition

$$
\begin{aligned}
& & \mathcal{F}_0^{(\nu,p)}(x) &= H(x)\left[1 - pe^{-\nu x}\right] \le H(x) = \mathcal{F}_0(x) \\
\Rightarrow & & \mathcal{F}_n^{(\nu,p)}(x) &\le \mathcal{F}_n(x) \quad \text{by Lemma 6.3.4} \\
\Rightarrow & & \mathcal{F}_n^{(\nu,p)}(x) &\le \mathcal{F}_0(x) \quad \text{by (6.3.33)} \\
\Rightarrow & & 0 \le \lim_{x\to-\infty} \mathcal{F}^{(\nu,p)}(x) &= \lim_{x\to-\infty}\lim_{n\to\infty} \mathcal{F}_n^{(\nu,p)}(x) \\
& & &\le \lim_{x\to-\infty} \mathcal{F}_0(x) = 0
\end{aligned}
\tag{6.3.36}
$$

(6.3.35) and (6.3.36), together with the monotonicity inherited by the limit, gives us the fact that — subject to the "conditions" in (6.3.34) — $\mathcal{F}^{(\nu,p)}$ is a proper distribution function. More significantly, the second line of (6.3.36) provides us with $\lim_{x\to\infty} \mathcal{F}(x) = 1$.

(b) (i) Use the usual argument involving the Borel-Cantelli Lemma and the extended version of Chebyshev's inequality, as in §6.2.

(ii) (1) Since (6.3.30) is of precisely the same form as (6.3.27), we deduce from Lemma 6.3.6 (equation (6.3.28)) that $\Lambda\mathcal{F}(x) = \mathcal{F}(x) + O\left(e^{-ax}\right)$ holds. This, in turn, coincides with (6.3.24) which, by Corollary 6.3.5(a), implies the existence of a pointwise limit, $G^{(c)}$. If $\mathcal{F}(x)$ and $\mathcal{F}^*(x)$ are two different members of $\boldsymbol{\mathcal{F}}(c)$, then $\mathcal{F}(x) - \mathcal{F}^*(x) = O\left(e^{-ax}\right)$ so that (6.3.22) is satisfied. As a result, Lemma 6.3.5 yields

$$|\Lambda^n \mathcal{F}(x) - \Lambda^n \mathcal{F}^*(x)| \leq he^{-ax}b^n \underset{\infty}{\overset{n}{\rightarrow}} 0$$

from which the uniqueness of the solution is deduced.

(2) Consider the following shifted exponential distribution functions[3]

$$\left.\begin{array}{rcl} \mathcal{F}^{(c)}(x) & = & [1 - ce^{-\alpha x}]^+ \, H(x) \\ \text{i.e.} \quad \mathcal{F}^{(c)}(x) & = & \left\{ \begin{array}{cc} 1 - ce^{-\alpha x} & x \geq \left[\frac{1}{\alpha}\ln c\right]^+ \\ 0 & x \leq \left[\frac{1}{\alpha}\ln c\right]^+ \end{array} \right. \end{array}\right\} c \geq 0 \tag{6.3.37}$$

so (certainly for $x \geq X = \left[\frac{1}{\alpha}\ln c\right]^+$) $\mathcal{F}^{(c)} \in \boldsymbol{\mathcal{F}}(c)$ implying that the sequence $\left\langle \mathcal{F}_n^{(c)}(x) = \Lambda^n \mathcal{F}^{(c)}(x), n \in \mathbb{Z}_\oplus \right\rangle$ tends to $G^{(c)}$.

[3] "$\ln 0 = -\infty$"

Now if $c_1 < c_2$, then

$$
\begin{aligned}
& \mathcal{F}^{(c_1)}(x) \quad - \quad \mathcal{F}^{(c_2)}(x) \\
= & \begin{cases} (c_2 - c_1)e^{-\alpha x} & x \geq \left[\frac{1}{\alpha}\ln c_2\right]^+ \\ 1 - c_1 e^{-\alpha x} & \left[\frac{1}{\alpha}\ln c_1\right]^+ \leq x \leq \left[\frac{1}{\alpha}\ln c_2\right]^+ \\ 0 & x \leq \left[\frac{1}{\alpha}\ln c_1\right]^+ \end{cases}
\end{aligned}
$$

$$
\Rightarrow \quad \mathcal{F}^{(c_1)}(x) \quad > \quad \mathcal{F}^{(c_2)}(x) \quad \forall\, x > \left[\tfrac{1}{\alpha}\ln c_1\right]^+
$$

$$
\begin{aligned}
\Rightarrow \quad & \mathcal{F}_n^{(c_1)}(x) \quad > \quad \mathcal{F}_n^{(c_2)}(x) \\
& {\scriptstyle n}\downarrow{\scriptstyle \infty} \qquad\qquad\quad {\scriptstyle n}\downarrow{\scriptstyle \infty} \\
& G^{(c_1)}(x) \quad > \quad G^{(c_2)}(x) \quad x > \left[\tfrac{1}{\alpha}\ln c_1\right]^+
\end{aligned}
$$

by the (strict) monotonicity properties of operator Λ inherited by the limit

$$
\Longrightarrow \quad \overline{G}^{(0)}(x) \quad < \quad \overline{G}^{(c_1)}(x) \quad < \quad \overline{G}^{(c_2)}(x)
$$

$\Rightarrow$ the minimality of the random variable with distribution function $G^{(0)}(x) = G(x)$.

(c) If necessary, one may define a new variable u' with $\inf_{a>0} \mathrm{E}e^{au'} > \frac{1}{m}$, and satisfying

$$
\begin{array}{rrcll}
 & P[u' = -\infty] & = & p \in [0,1) & \text{where } p \geq P[u = -\infty] \\
 & P[u' = k\delta] & = & p_k & k \in [-l, l']_{\mathbb{Z}};\ \ l, l' \in \mathbb{N} \\
\text{where} & \delta & > & 0 & \\
 & \sum_{k=-l}^{l'} p_k & = & 1 - p & \\
\text{and} & \prod_{k=-l}^{l'} p_k & > & 0 & \\
\text{such that} & P[u \geq u'] & = & 1 &
\end{array}
\tag{6.3.38}
$$

It follows that if a solution does not exist in the case of u', then the same is true for u. From now on it will be convenient not to make any notational distinctions for the two cases.

Assume to the contrary that a solution does exist. From (6.3.11)

$$
\begin{aligned}
\mathcal{F}(x) &= K * [(\mathcal{F}H)(x)]^m \qquad\qquad x \in \mathbb{R} \\
&= \sum_y P[u = y]\,[(\mathcal{F}H)(x - y)]^m \\
&= P[u = -\infty]\,[(\mathcal{F}H)(\infty)]^m + \sum_{k=-l}^{l'} P[u = k\delta]\,[(\mathcal{F}H)(x - k\delta)]^m \\
&= p + \sum_{k=-l}^{l'} p_k\,[(\mathcal{F}H)(x - k\delta)]^m
\end{aligned}
\tag{6.3.39}
$$

Write

$$y_n = 1 - \mathcal{F}(x_0 + n\delta) \quad n \in \mathbb{Z}_{\oplus},\ x_0 \geq l'\delta \tag{6.3.40}$$

Then for $n \in \overline{\mathbb{N}}_{l'-1}$

$$\begin{aligned}
y_n &= 1 - \mathcal{F}(x_0 + n\delta) \\
&= 1 - \left[p + \sum_{k=-l}^{l'} p_k \left[(\mathcal{F}H)(x_0 + (n-k)\delta)\right]^m\right] && \text{from (6.3.39)} \\
&= \sum_{k=-l}^{l'} p_k \left[1 - \left[\mathcal{F}(x_0 + (n-k)\delta)\right]^m\right] && \text{using (6.3.38)} \\
&= \sum_{k=-l}^{l'} p_k \left[1 - (1 - y_{n-k})^m\right] && \text{by (6.3.40)} \\
&= -\sum_{k=-l}^{l'} p_k \sum_{j=1}^{m} \binom{m}{j} (-y_{n-k})^j \\
&= \sum_{j=1}^{m} \binom{m}{j} (-1)^{j+1} \sum_{k=-l}^{l'} p_k y_{n-k}^j
\end{aligned} \tag{6.3.41}$$

After some careful reflection, it emerges that Siegel's Theorem may be applied to (6.3.41) (where u' may, if necessary, be altered without changing the essential assumptions) to deduce that that there exists a positive, bounded sequence $\langle y_n^*, n \in \mathbb{Z}_\oplus \rangle$ such that

$$y_n^* = m \sum_{k=-l}^{l'} p_k y_{n-k}^* \quad n \geq l' \tag{6.3.42}$$

The characteristic equation $\chi(\lambda) = 0$ may be determined informally by writing $y_{n-k}^* = \lambda^{-k}$. Substitution in (6.3.42) then yields

$$\sum_{k=-l}^{l'} p_k \lambda^{-k} = \frac{1}{m} \tag{6.3.43}$$

As explained in Karpelevich *et al.* [35], u' may be manipulated in such a way that $\exists\, \lambda_0 \in (0,1)$ (w.l.g. $\lambda_0 = e^{-a_0 \delta}$) which solves (6.3.43), but

$$\frac{1}{m} = \sum_{k=-l}^{l'} p_k \lambda_0^{-k} = \sum_{k=-l}^{l'} p_k e^{a_0(k\delta)} + \text{“}p e^{a_0(-\infty)}\text{”} = \mathrm{E} e^{a_0 u}$$

in contradiction with (6.3.31).

Hence our assumption on the existence of a solution is invalid.

□

6.4 THE EXPONENTIAL CASE

As in Chapter 5, we investigate the case in which the mutually independent IID sequences of interarrival times and service times are Exp(λ^{-1})- and Exp(μ^{-1})-distributed respectively. The analysis is inevitably more tractable than the general case of the previous section. Remark 6.2.1 has already alluded to this fact.

If the infimum in (6.2.9) (or its non-strict form) is attained, then to find expressions for α and β (cf. (6.3.5a) and (6.3.5b)) i.t.o. λ and μ, consider

$$\begin{aligned} \tfrac{\lambda\mu}{(\lambda+a)(\mu-a)} &\leq \tfrac{1}{m} \qquad \text{(cf. (6.2.10))} \\ \Rightarrow \quad a^2-(\mu-\lambda)a+\lambda\mu(m-1) &< 0 \\ \Rightarrow \quad a \in \left(\tfrac{1}{2}[\mu-\lambda-\sqrt{\Delta}], \tfrac{1}{2}[\mu-\lambda+\sqrt{\Delta}]\right) \end{aligned}$$

where $\Delta = (\mu-\lambda)^2 - 4\lambda\mu(m-1)$, i.e.

$$\alpha = \tfrac{1}{2}[\mu-\lambda-\sqrt{(\mu-\lambda)^2-4\lambda\mu(m-1)}] \tag{6.4.1a}$$

$$\beta = \tfrac{1}{2}[\mu-\lambda+\sqrt{(\mu-\lambda)^2-4\lambda\mu(m-1)}] \tag{6.4.1b}$$

We first note that equations (5.5.5)–(5.5.7) are equally valid and we repeat them here for the sake of clarity:

$$g(x) = \begin{cases} \frac{\lambda\mu}{\lambda+\mu}e^{-\mu x} & x \geq 0 \\ \frac{\lambda\mu}{\lambda+\mu}e^{\lambda x} & x \leq 0 \end{cases} \tag{6.4.2}$$

$$g'(x) = \begin{cases} -\mu g(x) & x > 0 \\ \lambda g(x) & x < 0 \end{cases} \tag{6.4.3}$$

$$g''(x) = \begin{cases} \mu^2 g(x) & x > 0 \\ \lambda^2 g(x) & x < 0 \end{cases} \tag{6.4.4}$$

In a completely analogous fashion, therefore, we obtain the following equations for $\mathcal{F}(x)$ and its first derivative:

$$\mathcal{F}(x) = \int_0^\infty [\mathcal{F}(r)]^m g(x-r)dr \tag{6.4.5}$$

$$\mathcal{F}'(x) = \int_0^\infty [\mathcal{F}(r)]^m g'(x-r)dr \tag{6.4.6}$$

As before, we have the relationships

$$g''(x) = \lambda\mu g(x) + (\lambda-\mu)g'(x) \quad x \neq 0 \tag{6.4.7}$$

$$g'(0+) - g'(0-) = -\lambda\mu \tag{6.4.8}$$

which lead to

$$\begin{aligned}
\mathcal{F}''(x) &= \int_0^\infty [\mathcal{F}(r)]^m g''(x-r)dr + J\left(\mathcal{F}^m H\right)(x) \quad x \neq 0 \\
&= \int_0^\infty [\mathcal{F}(r)]^m [(\lambda-\mu)g'(x-r) + \lambda\mu g(x-r)]dr \\
&\quad -\lambda\mu[\mathcal{F}(x)]^m H(x) \\
&= (\lambda-\mu)\mathcal{F}'(x) + \lambda\mu\left[\mathcal{F}(x) - [\mathcal{F}(x)]^m H(x)\right]
\end{aligned}$$

$$\Rightarrow \quad \mathcal{F}''(x) = \begin{cases} (\lambda-\mu)\mathcal{F}'(x) + \lambda\mu\mathcal{F}(x)\left[1 - [\mathcal{F}(x)]^{m-1}\right] & x > 0 \quad \text{(6.4.9a)} \\ (\lambda-\mu)\mathcal{F}'(x) + \lambda\mu\mathcal{F}(x) & x < 0 \quad \text{(6.4.9b)} \end{cases}$$

Finally, before considering the differential equations (6.4.9a) and (6.4.9b) in greater detail, we note the following properties, similar to (5.5.13)–(5.5.15), of $\mathcal{F}$:

$$\begin{aligned}
\mathcal{F}(0) &= \int_0^\infty [\mathcal{F}(r)]^m g(-r)dr \\
&= \int_0^\infty [\mathcal{F}(r)]^m \tfrac{\lambda\mu}{\lambda+\mu} e^{-\lambda r} dr
\end{aligned} \tag{6.4.10}$$

$$
\begin{aligned}
x < 0 \quad & \Rightarrow \\
\mathcal{F}(x) \quad & = \quad \int_0^\infty [\mathcal{F}(r)]^m \tfrac{\lambda\mu}{\lambda+\mu} e^{\lambda(x-r)} dr \\
& = \quad e^{\lambda x} \int_0^\infty [\mathcal{F}(r)]^m \tfrac{\lambda\mu}{\lambda+\mu} e^{-\lambda r} dr \\
& = \quad e^{\lambda x} \mathcal{F}(0) \qquad\qquad \text{from (6.4.10)}
\end{aligned}
\tag{6.4.11}
$$

$$
\begin{aligned}
\mathcal{F}'(0) \quad & = \quad \int_0^\infty [\mathcal{F}(r)]^m g'(-r) dr \\
& = \quad \lambda \int_0^\infty [\mathcal{F}(r)]^m g(-r) dr \quad \text{from (6.4.3)} \\
& = \quad \lambda \mathcal{F}(0) \qquad\qquad \text{from (6.4.10)}
\end{aligned}
\tag{6.4.12}
$$

The solution to (6.4.9b) is provided by (6.4.11), so that one need only investigate the situation further for (6.4.9a).

When $m = 1$, the differential equation is easily solved, as follows:

$$
\begin{aligned}
\mathcal{F}''(x) \quad & = \quad (\lambda - \mu)\mathcal{F}'(x) \\
\Rightarrow \quad \int \tfrac{\mathcal{F}''(x)}{\mathcal{F}'(x)} dx \quad & = \quad \int (\lambda - \mu) dx \\
\Rightarrow \quad \mathcal{F}'(x) \quad & = \quad c_1 e^{-(\mu-\lambda)x} \\
\Rightarrow \quad \mathcal{F}(x) \quad & = \quad \tfrac{c_1}{\lambda-\mu} e^{-(\mu-\lambda)x} + c_2
\end{aligned}
$$

To evaluate the constants c_1 and c_2 we first observe that $c_2 = \mathcal{F}(\infty) = 1$. Thus $\mathcal{F}(x) = 1 + \frac{c_1}{\lambda-\mu}e^{-(\mu-\lambda)x}$.

Next (6.4.12) implies that $c_1 = \mathcal{F}'(0) = \lambda\mathcal{F}(0) = \frac{\lambda(\lambda-\mu+c_1)}{\lambda-\mu} \Rightarrow c_1 = \rho(\mu-\lambda)$.

Finally, therefore, we have

$$\begin{array}{rrcl} & \mathcal{F}(x) & = & 1-\rho e^{-(\mu-\lambda)x} \\ \text{or} & X & \sim & \mathrm{Exp}\left((\mu-\lambda)^{-1},\rho\right) \quad (\rho<1) \end{array} \tag{6.4.13}$$

which we recognise as the stationary $M/M/1$ waiting-time distribution. This is more than mere coincidence. Since we are considering $X > 0$ (in fact, $X \geq 0$) and $m = 1$, $X = X^+ \sim (W+u)^+ \sim W$.

Remark 6.4.1 (Basic Lindley Equation)

Before we proceed to consider more general values of m, we note the equivalence (except for $x = 0$) of (6.4.9a)/(6.4.9b) and (5.5.22) when $m = 1$, so that the derivation of (6.4.13) is also valid for (5.5.22).

For $m \in \overline{\mathbb{N}}_1$, it is convenient to write (6.4.9a) as

$$\mathcal{F}''(t) = (\lambda-\mu)\mathcal{F}'(t) + \lambda\mu\mathcal{F}(t)\left[1-[\mathcal{F}(t)]^{m-1}\right] \quad t>0 \tag{6.4.14}$$

where t may be interpreted as time.

Initially we let $x(t) = \mathcal{F}(t)$, so that $\ddot{x}(t) = \phi\left[t, x(t), \dot{x}(t)\right]$ and treat the system as *autonomous*, i.e., we obtain a second-order nonlinear differential equation

$$\ddot{x} = (\lambda-\mu)\dot{x} + \lambda\mu x(1-x^{m-1}) \tag{6.4.15}$$

In accordance with a standard technique in dynamical systems (with one degree of freedom), we reduce (6.4.15) to a *first-order plane autonomous system* (i.e., a system of two first-order differential equations) by letting $y = \dot{x}$:

$$\left\{ \begin{array}{lcrcr} \dot{x} & = & & & y \\ \dot{y} & = & \lambda\mu(x - x^m) & - & (\mu - \lambda)y \end{array} \right\} \tag{6.4.16}$$

which, in view of (6.4.12), is subject to $\left.\frac{y}{x}\right]_{t=0} = \lambda$.

For convenience write $\boldsymbol{r} = (x, y)$. To find equilibrium points[4] of (6.4.16), we simply solve $\dot{\boldsymbol{r}} = \mathbf{0}$, which yields $\boldsymbol{r} \in \{\boldsymbol{r}_1 = (0,0), \boldsymbol{r}_2 = (1,0)\}\ \forall\, m$ and $(-1,0)$ if m is odd. Since $x(t) = \mathcal{F}(t)$, a distribution function, the point $(-1,0)$ does not interest us, so we restrict our attention to $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$. More especially, we require $\boldsymbol{r} \in [0,1] \times \mathrm{I\!R}_{\oplus}$ $(t > 0)$ (and $x(\infty) = 1$).

Using the fact that $m \geq 2$, linearisation of the system at $\boldsymbol{r}_1$ takes the form

$$\dot{\boldsymbol{r}} = \boldsymbol{r}\mathbf{A}_1 \quad \text{where} \quad \mathbf{A}_1 = \begin{pmatrix} 0 & \lambda\mu \\ 1 & -(\mu - \lambda) \end{pmatrix}$$

The characteristic roots of $\mathbf{A}_1$ are $\gamma_1 = \lambda > 0$ and $\gamma_2 = -\mu < 0$, which means that $(0,0)$ is a *saddle point* of the original system in (6.4.16).

A little more effort is required to deal with the point $\boldsymbol{r}_2$. Let $x = z + 1$. Then

$$\left\{ \begin{array}{lcrcr} \dot{z} & = & & & y \\ \dot{y} & = & \lambda\mu(z+1)[1 - (z+1)^{m-1}] & - & (\mu - \lambda)y \end{array} \right\} \tag{6.4.17}$$

[4]also known as *critical points* or *singular points* in the literature

with the linearisation

$$\left\{\begin{array}{rcrcr} \dot{z} & = & & & y \\ \dot{y} & = & \lambda\mu(1-m)z & - & (\mu-\lambda)y \end{array}\right\} \tag{6.4.18}$$

and critical point $\boldsymbol{r}'_0 = (0,0)$, where $\boldsymbol{r}' = (z,y)$.

(6.4.18) may then be written more succinctly as

$$\dot{\boldsymbol{r}}' = \boldsymbol{r}'\mathbf{A}_2 \quad \text{where} \quad \mathbf{A}_2 = \begin{pmatrix} 0 & \lambda\mu(1-m) \\ 1 & -(\mu-\lambda) \end{pmatrix}$$

from which

$$\mid \mathbf{A}_2 - \delta\mathbf{I} \mid = \delta^2 + (\mu-\lambda)\delta - \lambda\mu(1-m)$$

Consequently the eigenvalues are determined by

$$\delta \;=\; \tfrac{1}{2}\left[-(\mu-\lambda) \pm \sqrt{\Delta}\right]$$

where

$$\begin{aligned} \Delta &= (\mu-\lambda)^2 - 4\lambda\mu(m-1) \\ &= (\mu+\lambda)^2 - 4\lambda\mu m \end{aligned}$$

For convenience write

$$\delta_j = -\tfrac{1}{2}\left[(\mu-\lambda) + (-1)^j\sqrt{\Delta}\right] \qquad j = 1,2$$

In order to find the nature of the critical point $(1,0)$ of (6.4.16), it is necessary to examine the behaviour of the discriminant Δ.

$$\begin{aligned}\Delta = 0 \;&\Rightarrow\; \lambda^2 - 2\mu(2m-1)\lambda + \mu^2 = 0 \\ &\Rightarrow\; \lambda = \left[(2m-1) - \sqrt{(2m-1)^2 - 1}\right]\mu\end{aligned}$$

where we observe that the other solution of the quadratic equation is not feasible (under the circumstances).

Letting $\rho = \frac{\lambda}{\mu}$ (as usual), we note that

$$\begin{aligned}\Delta \geq 0 \;&\Leftrightarrow\; \rho \leq 2m - 1 - 2\sqrt{m(m-1)} \\ &\Rightarrow\; \delta_1, \delta_2 \text{ both negative} \\ &\Rightarrow\; (1,0) \text{ an (asymptotically) stable node}\end{aligned}$$

$$\begin{aligned}\Delta < 0 \;&\Leftrightarrow\; \rho > 2m - 1 - 2\sqrt{m(m-1)} \\ &\Rightarrow\; \delta_1 \text{ and } \delta_2 \text{ are complex conjugates; } \Re(\delta_j) < 0 \\ &\Rightarrow\; (1,0) \text{ an (asymptotically) stable inward–spiralling focus}\end{aligned}$$

Remark 6.4.2

(a) There is a clear connection between Remark 6.2.1 and the present section, where the expressions for the traffic intensity ρ and the disriminant Δ recur.

(b) The interested reader is referred to Karpelevich *et al.* [35] where the discussion of §6.4 is incorporated into a number of theorems on the exponential case.

CHAPTER 7

THE *FAFS* DISCIPLINE ON A GENERAL NETWORK

7.1 INTRODUCTION

The purpose of this chapter is to investigate the operation of a general circuit-switched network 'governed' by the $FAFS$ discipline, which was introduced in Chapter 4 and applied to the starlike networks of the last two chapters.

The reasonably unrestricted nature of the network model prevents our construction of a Lindley-Loynes equation to describe its waiting times. The inclusion of this chapter in the thesis is justified nonetheless because of its relevance not only to the $FAFS$ protocol but also to complex queueing and communication systems.

There is a strong connection between Chapter 6 (for instance, Remark 6.2.1) and §7.5 where simplifying assumptions are made about the network. The more general situation is treated in §7.4. The next two sections are concerned with the inevitable description of new notation and concepts relevant to the current chapter.

7.2 THE STRUCTURE AND OPERATION OF THE NETWORK

The network we consider may be represented by a directed pseudograph $\mathcal{G} = [V, E]$ (Appendix A) where $\mathcal{G}$ is connected and vertex set V is countable, possibly denumerable. Write $\mathcal{N}(v) = \{v' : (v, v') \in E(\mathcal{G})\}$ for the *neighbourhood* of v (which may include v itself if there is a loop). If $m(v, v')$ is the multiplicity of arc (v, v') (Appendix A), then we require that

$$1 \leq \sum_{v' \in \mathcal{N}(v)} m(v, v') \leq M \quad \forall v \in V(\mathcal{G}) \tag{7.2.1}$$

To avoid confusion with the notation for waiting times, we represent a (finite) diwalk by γ and its length by $|\gamma|$. For completeness, we will allow for the trivial case of a walk of zero length, which is then a walk consisting of a single vertex. If we wish to emphasise that the initial vertex, the *source*, is v, we write γ^v. It is also convenient to define collections of walks as follows:

$$\Gamma^v = \{\gamma^v\} = \{\text{walks with (fixed) source } v\} \tag{7.2.2a}$$
$$\Gamma(l) = \{\gamma(l)\} = \{\text{walks of fixed length } l\} \tag{7.2.2b}$$
$$\Gamma = \{\gamma^v; v \in V\} = \{\text{walks of } \mathcal{G}\} \tag{7.2.2c}$$

$\Gamma^v(l)$ and $\gamma^v(l)$ are intuitively defined. Finally, we use the natural notation $\tilde{v} \in \gamma^v$ to represent the fact that $\tilde{v}$ is a vertex on the walk γ^v.

Definition 7.2.1 (Intersecting Walks)

The walks γ^v and $\gamma^{v'}$ *intersect*, written $\gamma^v \cap \gamma^{v'}$, iff $\exists\, \tilde{v} \in V$ s.t. $\tilde{v} \in \gamma^v$ and $\tilde{v} \in \gamma^{v'}$.

Much as before, the total arrival flow $\boldsymbol{\xi} = \boldsymbol{\xi}^0$ may be partitioned into streams $\xi^v = \xi^{0,v}$ to each vertex $v \in V$. The system is assumed to operate from time 0, while the first arrival to source v occurs at the moment t_1^v. More generally, t_n^v determines the time of the nth such arrival. The corresponding interarrival times are given by $\tau_n^v = t_{n+1}^v - t_n^v$ such that $t_{n+1}^v = t_1^v + \sum_{j=1}^{n} \tau_j^v$.

The variables $\langle t_1^v, v \in V \rangle$ are IID and distributed as G_1, independently of $\langle \tau_n^v; n \in \mathbb{N}, v \in V \rangle$ which again is an IID sequence of random variables with common distribution G and finite mean $\mathrm{E}\tau$. Stationarity is supposed so that

$$G_1(t) = \frac{1}{\mathrm{E}\tau} \int_0^t \overline{G}(u) du \tag{7.2.3}$$

which is also the distribution of the excess lifetime (Appendix B). We will make use of this fact in the sections which follow.

Calls are given by $y = [t; (S, \gamma^v)]$ where the mark γ^v (with v explicitly given) generalises $\boldsymbol{i}$ of Chapter 6, and the sequence of service times $\langle S_n^v; n \in \mathbb{N}, v \in V \rangle$ is IID with distribution F. Additional assumptions about F shall be made in §§7.4 and 7.5. As usual, n is indicative of the relative position of the call in its source (from time 0) and v gives its source. An analogous interpretation may be used for the IID sequence $\langle \gamma_n^v; n \in \mathbb{N}, v \in V \rangle$. For each fixed v and any $n \in \mathbb{N}$, γ_n^v has distribution ϕ^v, where

$$\phi^v(\gamma^v) \quad = \quad P[\gamma_n^v = \gamma^v] \quad \gamma^v \in \Gamma^v \tag{7.2.4}$$

The collection $\{\phi^v, v \in V\}$ satisfies the following constraint:

$$\exists \ \langle p_l, l \in \mathbb{Z}_\oplus \rangle \ \text{s.t.} \ \forall\, v \in V, \ \ \forall\, \gamma^v(l) \in \Gamma^v(l)$$

$$\phi^v\left(\gamma^v(l)\right) \ < \ p_l \tag{7.2.5a}$$

$$\text{where} \quad \mathcal{P} \ = \ \sum_{l=1}^{\infty} l^2 M^{l-1} p_{l-1} \ < \ \infty \tag{7.2.5b}$$

Finally, marks $S^\cdot_\cdot$ and $\gamma^\cdot_\cdot$ are mutually independent, in addition to being distributed independently of the (inter)arrival-time sequence.

We now turn to the issue of waiting times. The vector of initial conditions $W^0[\gamma^v]$ (imposed at time 0) shall be denoted by $\boldsymbol{w}^0 = \boldsymbol{w}^0[\gamma^v \in \Gamma^v; v \in V] \in \mathbb{R}_\oplus^{|\Gamma|}$ where the trivial initial condition $\boldsymbol{w}^0 = \mathbf{0}$ often will apply, and $\boldsymbol{w}^0$ is independent of $\boldsymbol{\xi}$.

At this stage it is useful to review briefly how the circuit-switched $FAFS$ network will operate. First a definition is in order.

Definition 7.2.2 (Generalised Predecessor)
Let $y = [t; (S, \gamma^v)]$ and $y' = [t'; (S', \gamma^{v'})]$. Then y' is said to be a *generalised predecessor* of y, written $y' = p(y)$, iff $t' < t$ and $\gamma^v \cap \gamma^{v'}$. More especially, if these walks have vertex $\tilde{v}$ in common, then y' shall be called a *generalised predecessor* of y *at* $\tilde{v}$, denoted $y' = p_{\tilde{v}}(y)$.

Remark 7.2.1
As before $p^n(y)$ for $n \in \overline{\mathbb{N}}_1$ and *generalised successors* $p^{-n}(y)$ $(n \in \mathbb{N})$ may also be considered.

The $FAFS$ discipline for the more general network of this chapter may be summarised as follows. $y = [t; (S, \gamma^v)]$ completes its wait $W^t(y) = W^t[\gamma^v]$ and is served at that first moment after $W^0[\gamma^v]$ when all its generalised predecessors $\{p_{\tilde{v}}(y); \tilde{v} \in \gamma^v\}$ have left the network. After time S has passed, y itself leaves the network (its sojourn time $T^t(y)$ having elapsed).

Yet some more formalities will be needed before we will be able to state and prove some theorems.

7.3 THE WAITING-TIME PROCESS & GENERALISATIONS OF THE PREDECESSOR PATH

In this section our initial concern is with monotonicity of the waiting-time process and the effect of shifts.

If we generalise the notation of Remark B.2.2, we may write $N^v(s)$ for the number of arrivals to v in $(0, s]$. One may then consider shifted random variables

$$X_n^{s,v} \;=\; T^{N^v(s)} X_n \;=\; X_{N^v(s)+n}$$

where t_n^v, τ_n^v, S_n^v and γ_n^v may be substituted for X_n, in which case the shifted arrival stream shall be denoted by $\boldsymbol{\xi}^s$.

Under the stationarity assumption associated with (7.2.3) (and the nature of the $FAFS$ discipline), we have

$$\boldsymbol{\xi}^s \;=_d\; \boldsymbol{\xi}^{s'} \quad \forall\, s, s' \in \mathbb{R}_\oplus \tag{7.3.1}$$

We wish to investigate the waiting-time process with trivial initial condition $\mathbf{0}$. The process may be viewed as a vector of virtual waiting times $W^t[\gamma^v]$ (based on fictitious calls $y_t = [t; (S, \gamma^v)]$), thus $\boldsymbol{w}^t = \boldsymbol{w}^t[\gamma^v \in \Gamma^v, v \in V]$. More generally $\boldsymbol{w}^t(\boldsymbol{w})$ is the waiting-time process arising from the initial condition $\boldsymbol{w}$. For reasons similar to those which give rise to (7.3.1), we have

$$\boldsymbol{w}^{t+s}(\mathbf{0}) \;=\; \boldsymbol{w}^{t+s} \;=_d\; \boldsymbol{w}^t(\boldsymbol{w}^s) \tag{7.3.2a}$$

but, because the process is monotonic w.r.t. the initial condition, $\boldsymbol{w}^t(\mathbf{0}) \leq \boldsymbol{w}^t(\boldsymbol{w}^s)$ which, on taking (7.3.2a) into account, implies

$$\boldsymbol{w}^{t+s}(\mathbf{0}) \;\geq_{st}\; \boldsymbol{w}^t(\mathbf{0}) \tag{7.3.2b}$$

This, in turn, means that $\mathrm{E}(\boldsymbol{w}^t(\mathbf{0})) \leq \mathrm{E}(\boldsymbol{w}^{t+s}(\mathbf{0}))$. Thus, in order to show convergence of $\boldsymbol{w}^t$ to a stationary process, it suffices to prove that $\mathrm{E}(W^t[\gamma^v])$ is bounded uniformly in $t \in \mathbb{R}_\oplus$ and $v \in V$. Alternatively, it is enough to demonstrate uniform boundedness/stability (cf. §1.2) of $W^t[\gamma^v]$. (Recall the related concept that a monotonic nondecreasing sequence converges iff it is bounded.) We use the first approach in Theorem 7.4.1 and the second in Theorem 7.5.1.

Definition 7.3.1 (Chains of Influence)

Consider some fixed call $y^0 = [t_0; (S_0, \gamma^{v_0})]$. Let $\boldsymbol{v} = (v_1, \ldots, v_n) \in V^n$, $\boldsymbol{k} = (k_1, \ldots, k_n) \in \mathbb{N}^n$ and $\boldsymbol{t} = (t_1, \ldots, t_n) \in \mathbb{R}^n_{\oplus}$. With $\boldsymbol{v}$ associate the sequence of walks $\langle \gamma^{v_i}, i \in \mathbb{N}_n \rangle$.

(a) The walks form a *plane chain*, $\boldsymbol{\gamma}(\boldsymbol{v}) = (\gamma^{v_1}, \ldots, \gamma^{v_n})$, w.r.t. y^0 iff $\gamma^{v_{i-1}} \cap \gamma^{v_i}\ \forall\, i \in \mathbb{N}_n$. The collection $\{\boldsymbol{\gamma}(\boldsymbol{v})\}$ of plane chains of length n shall be written as $\boldsymbol{\Gamma}(n, \boldsymbol{v})$.

(b) The sequence of calls $\langle y^i = [t_i; (S_i, \gamma^{v_i})], i \in \mathbb{N}_n \rangle$ forms a *space chain* (of length n), $\boldsymbol{\gamma}(\boldsymbol{k}, \boldsymbol{v}, \boldsymbol{t})$, w.r.t. y^0 iff

(i) $t_{i-1} > t_i\ \forall\, i \in \mathbb{N}_n$ (7.3.3)

(ii) k_i calls arrive to the vertex v_i during the interval $[t_i, t_{i-1}]$ $(i \in \mathbb{N}_n)$ (7.3.4)

(iii) $\boldsymbol{\gamma}(\boldsymbol{v})$ forms a plane chain (of length n) w.r.t. y^0 (7.3.5)

The collection $\{\boldsymbol{\gamma}(\boldsymbol{k}, \boldsymbol{v}, \boldsymbol{t})\}$ of such space chains (of length n) shall be denoted by $\boldsymbol{\Gamma}(n, \boldsymbol{k}, \boldsymbol{v})$.

(c) $\boldsymbol{\gamma}(\boldsymbol{k}, \boldsymbol{v}, \boldsymbol{t}) \in \boldsymbol{\Gamma}(n, \boldsymbol{k}, \boldsymbol{v})$ shall be called an *essential space chain* iff

(i) $\forall\, i \in \mathbb{N}_n$ call y^{i-1} is served immediately after y^i has left the network

(ii) $W^{t_n}(y^n) = 0$

$\boldsymbol{\Gamma}_E(n, \boldsymbol{k}, \boldsymbol{v}) \subseteq \boldsymbol{\Gamma}(n, \boldsymbol{k}, \boldsymbol{v})$ refers to the collection of essential space chains.

Remark 7.3.1 (Notation for Chains of Influence)

(a) For the plane chain $\boldsymbol{\gamma}(\boldsymbol{v})$ of Definition 7.3.1, we shall let $\boldsymbol{\gamma}_i(\boldsymbol{v}) = \gamma^{v_i}$.

(b) It is important to realise that one of the k_i calls arriving to v_i during the period $[t_i, t_{i-1}]$ is y^i itself, which implies that $k_i \in \mathbb{N}$. For $i \in \mathbb{N}_n$, it is therefore convenient to define

$$\begin{aligned} k'_i &= k_i - 1 \in \mathbb{Z}_{\oplus} \\ \boldsymbol{k}' &= \boldsymbol{k} - \boldsymbol{1} \in \mathbb{Z}^n_{\oplus} \\ K_n &= \sum_{i=1}^{n} k_i \\ K'_n &= \sum_{i=1}^{n} k'_i = K_n - n \end{aligned}$$

Much as in the last chapter, we need to demonstrate that any call y^0 is influenced by finitely many other calls w.p.1. Before we are able to do so, a lemma is required.

We recall the following combinatorial results for $r \in \mathbb{Z}_{\oplus}$:

$$\sum_{j=0}^{n} \binom{r+j}{j} = \binom{n+r+1}{n} \tag{7.3.6a}$$

$$\sum_{j=0}^{r} \binom{r}{j} = 2^r \tag{7.3.6b}$$

Lemma 7.3.1

Suppose $n \in \mathbb{N}$ and $K \in \mathbb{Z}_{\oplus}$. Then

$$\left|\left\{(k_1, \ldots, k_n) \in \mathbb{Z}_{\oplus}^n : \sum_{i=1}^{n} k_i = K\right\}\right| = \binom{K+n-1}{n-1} = \binom{K+n-1}{K} \qquad (7.3.7)$$

<u>PROOF</u>

We proceed by induction on n.

Obviously $\left|\left\{k_1 \in \mathbb{Z}_{\oplus} : \sum_{i=1}^{1} k_i = K\right\}\right| = 1 = \binom{K}{K} = \binom{K+1-1}{K}$, so suppose (7.3.7) is true for some $n \in \mathbb{N}$.

Then

$$\begin{aligned}
&\left|\left\{(k_1, \ldots, k_n, k_{n+1}) \in \mathbb{Z}_{\oplus}^{n+1} : \sum_{i=1}^{n+1} k_i = K\right\}\right| \\
&= \sum_{K'=0}^{K} \left|\left\{(k_1, \ldots, k_n) \in \mathbb{Z}_{\oplus}^n : \sum_{i=1}^{n} k_i = K'\right\}\right| && \text{where } K' = K - k_{n+1} \\
&= \sum_{K'=0}^{K} \binom{K'+n-1}{K'} && \text{by (7.3.7)} \\
&= \binom{K+n}{K} && \text{by (7.3.6a)} \\
&= \binom{K+(n+1)-1}{K}
\end{aligned}$$

which completes the proof.

□

Remark 7.3.2

The combinatorial result, $\binom{r}{j} \leq 2^r$, which will also be used later, may be determined immediately from (7.3.6b). As a special instance, if $n \in \mathbb{N}$ and $K \in \mathbb{Z}_{\oplus}$, then

$$\binom{K+n-1}{K} \leq 2^{K+n-1} \tag{7.3.8}$$

Proposition 7.3.1

If $\mathcal{N}_n$ denotes the number of plane chains of length n (as described in Definition 7.3.1)[1], then

$$\mathrm{E}\mathcal{N}_n \leq (|\gamma^{v_0}| + 1)\,\mathcal{P}^n = C_1\mathcal{P}^n \tag{7.3.9}$$

<u>PROOF</u>

Initially consider a fixed walk $\overline{\gamma}(l_0) : u_0u_1 \ldots u_{l_0}$. If one of the $l_0 + 1$ (not necessarily distinct) such vertices u_j is to intersect with some other walk $\gamma(l)$, then $\gamma(l)$ must be of the form $\gamma(l) : x_0x_1 \ldots (x_k = u_j) \ldots x_l$.

In view of condition (7.2.1), there are at most M^l such walks for each $k \in \{0\} \cup \mathbb{N}_l$. This, in turn, means that there are at most $(l+1)M^l$ walks for each u_j ($j \in \{0\} \cup \mathbb{N}_{l_0}$) which finally implies a bound of

$$(l_0 + 1)(l + 1)M^l \tag{7.3.10}$$

[1]not to be confused with $\mathcal{N}(v)$ of §7.2

on the number of walks of length l intersecting with a fixed walk of length l_0.

Next write $\boldsymbol{l} = (l_1, \ldots, l_n) \in \mathbb{Z}_{\oplus}^n$.

$$\begin{aligned}
\mathrm{E}\mathcal{N}_n &= \sum_{\boldsymbol{v}\in V^n} \mathrm{E} I_{[\exists\, \boldsymbol{\gamma}(\boldsymbol{v})\in\boldsymbol{\Gamma}(n,\boldsymbol{v})]} \\
&= \sum_{\boldsymbol{v}\in V^n} P[\exists\, \boldsymbol{\gamma}(\boldsymbol{v}) \in \boldsymbol{\Gamma}(n,\boldsymbol{v})] \\
&= \sum_{\boldsymbol{v}\in V^n} \sum_{\boldsymbol{\gamma}(\boldsymbol{v})\in\boldsymbol{\Gamma}(n,\boldsymbol{v}):\boldsymbol{\gamma}_i(\boldsymbol{v})=\gamma^{v_i}, i\in\mathbb{N}_n} \prod_{j=1}^{n} \phi^{v_j}(\gamma^{v_j}) \\
&= \sum_{\boldsymbol{v}\in V^n} \sum_{\boldsymbol{l}\in\mathbb{Z}_{\oplus}} \sum_{\boldsymbol{\gamma}(\boldsymbol{v})\in\boldsymbol{\Gamma}(n,\boldsymbol{v}):\boldsymbol{\gamma}_i(\boldsymbol{v})=\gamma^{v_i}(l_i), i\in\mathbb{N}_n} \prod_{j=1}^{n} \phi^{v_j}(\gamma^{v_j}) \\
&\leq \sum_{\boldsymbol{v}\in V^n} \sum_{\boldsymbol{l}\in\mathbb{Z}_{\oplus}} \sum_{\boldsymbol{\gamma}(\boldsymbol{v})\in\boldsymbol{\Gamma}(n,\boldsymbol{v}):\boldsymbol{\gamma}_i(\boldsymbol{v})=\gamma^{v_i}(l_i), i\in\mathbb{N}_n} \prod_{j=1}^{n} p_{l_j} \qquad \text{by (7.2.5a)}
\end{aligned}$$

Thus, using (7.3.10), we have

$$\begin{aligned}
\mathrm{E}\mathcal{N}_n &\leq (|\gamma^{v_0}|+1) \sum_{\boldsymbol{l}\in\mathbb{Z}_{\oplus}^n} (l_1+1) M^{l_1} p_{l_1} \left[\prod_{j=2}^{n} (l_{j-1}+1)(l_j+1) M^{l_j} p_{l_j} \right] \\
&\leq (|\gamma^{v_0}|+1) \prod_{j=1}^{n} \sum_{l_j=0}^{\infty} (l_j+1)^2 M^{l_j} p_{l_j} \quad (=) \\
&\leq (|\gamma^{v_0}|+1)\, \mathcal{P}^n \qquad (=) \quad \text{by (7.2.5b)} \\
&\leq C_1 \mathcal{P}^n
\end{aligned}$$

which is (7.3.9). □

Proposition 7.3.2

Suppose $\langle T_i \rangle$ is an IID sequence with $T_i \sim G_1 \; \forall\, i \in \mathbb{N}$ and $T_{(n)} = \sum_{i=1}^{n} T_i$. Similarly, let $\langle \tau_i \rangle$ be an IID sequence, independent of $\langle T_i \rangle$, with $\tau_i \sim G \; \forall\, i \in \mathbb{N}$ and $\tau_{(m)} = \sum_{i=1}^{m} \tau_i$. For $t_0 > 0$

$$\sum_{\boldsymbol{k} \in \mathbb{N}^n} P\left[T_{(n)} + \tau_{(K_n')} < t_0\right] \;\leq\; 2^n P[T_{(n)} < t_0] \sum_{K=0}^{\infty} 2^K P[\tau_{(K)} < t_0] \;\leq\; C_2 q^n \tag{7.3.11}$$

where $C_2 > 0$, $q \in (0,1)$ and the bound $C_2 q^n$ holds for $n \geq n_0 \in \mathbb{N}$.

<u>Proof</u>

It is clear[2] that $\exists\, n_0$ s.t. $n, K \geq n_0 \Rightarrow$

$$P[\tau_{(K)} < t_0] \;\leq\; c_1^K \;<\; \left(\tfrac{1}{2}\right)^K \tag{7.3.12a}$$

$$P[T_{(n)} < t_0] \;\leq\; c_2^n \;<\; \left(\tfrac{1}{2} q\right)^n \tag{7.3.12b}$$

where $q\mathcal{P} < 1$.

[2] For a formal proof, refer to Berezner & Malyshev [8, *p.* 367].

Next

$$
\begin{aligned}
\sum_{\boldsymbol{k}\in\mathbb{N}^n} P\left[\tau_{(K'_n)} < t_0\right] &= \sum_{K=0}^{\infty} \sum_{\boldsymbol{k}'\in\mathbb{Z}^n_{\oplus}:K'_n=K} P\left[\sum_{i=1}^{K} \tau_i < t_0\right] \\
&= \sum_{K=0}^{\infty} \binom{K+n-1}{K} P\left[\sum_{i=1}^{K} \tau_i < t_0\right] \qquad \text{by (7.3.7)} \\
&\le \sum_{K=0}^{\infty} 2^{K+n} P\left[\sum_{i=1}^{K} \tau_i < t_0\right]
\end{aligned}
\tag{7.3.13}
$$

Recall that both $T_{(n)}$ and $\tau_{(K'_n)}$ are nonnegative (in fact, positive) quantities. Thus

$$
\begin{aligned}
& T_{(n)} + \tau_{(K'_n)} < t_0 &&\Rightarrow\ T_{(n)} < t_0 \wedge \tau_{(K'_n)} < t_0 \\
\Rightarrow\quad & P\left[T_{(n)} + \tau_{(K'_n)} < t_0\right] &&\le\ P\left[\{T_{(n)} < t_0\} \cap \{\tau_{(K'_n)} < t_0\}\right] \qquad (=) \\
& &&\le\ P[T_{(n)} < t_0] \cdot P[\tau_{(K'_n)} < t_0]
\end{aligned}
$$

using independence

$$
\begin{aligned}
\Longrightarrow\quad \sum_{\boldsymbol{k}\in\mathbb{N}^n} P\left[T_{(n)} + \tau_{(K'_n)} < t_0\right] &\le P[T_{(n)} < t_0] \sum_{\boldsymbol{k}\in\mathbb{N}^n} P\left[\tau_{(K'_n)} < t_0\right] \\
&\le 2^n P[T_{(n)} < t_0] \sum_{K=0}^{\infty} 2^K P[\tau_{(K)} < t_0] \\
&\le C_2 q^n
\end{aligned}
$$

where the penultimate inequality is a result of (7.3.13) and the last step is a consequence of (7.3.12a) & (7.3.12b) and, therefore, applies for $n \ge n_0$. $\square$

Theorem 7.3.1

Any call y^0 is influenced by finitely many other calls w.p.1. Hence the waiting times are finite w.p.1, and the functioning of the system is correctly defined.

PROOF

Let

$$\begin{aligned} A_n(\boldsymbol{k},\boldsymbol{v}) &= \{\exists\, \boldsymbol{\gamma}(\boldsymbol{k},\boldsymbol{v},\boldsymbol{t}) \in \boldsymbol{\Gamma}(n,\boldsymbol{k},\boldsymbol{v})\} \\ A_n &= \bigcup_{\boldsymbol{k}\in\mathbb{N}^n} \bigcup_{\boldsymbol{v}\in V^n} A_n(\boldsymbol{k},\boldsymbol{v}) \\ B_n(\boldsymbol{k},\boldsymbol{v}) &= \{\exists\, \langle y^i = [t_i;(S_i,\gamma^{v_i})], i \in \mathbb{N}_n\rangle \text{ s.t. (7.3.3) and (7.3.4) hold}\} \end{aligned} \tag{7.3.14}$$

Now

(a) $P[B_n(\boldsymbol{k},\boldsymbol{v})]$, which does not depend on $\boldsymbol{v} \in V^n$, clearly satisfies

$$P[B_n(\boldsymbol{k},\boldsymbol{v})] \le P\left[T_{(n)} + \tau_{(K'_n)} < t_0\right] \tag{7.3.15}$$

(b) $$P[A_n(\boldsymbol{k},\boldsymbol{v}) \mid B_n(\boldsymbol{k},\boldsymbol{v})] = P[\exists\, \boldsymbol{\gamma}(\boldsymbol{v}) \in \boldsymbol{\Gamma}(n,\boldsymbol{v})] \tag{7.3.16}$$

since (7.3.16) is informally

$P[(7.3.3) \cap (7.3.4) \cap (7.3.5) \mid (7.3.3) \cap (7.3.4)] = P[(7.3.5)]$.

This conditional probability does not depend on $\boldsymbol{k} \in \mathbb{N}^n$.

Hence, by using (7.3.16) and (7.3.15)

$$\begin{aligned}
P(A_n) &= P\left[\bigcup_{\boldsymbol{k}}\bigcup_{\boldsymbol{v}} A_n(\boldsymbol{k},\boldsymbol{v})\right] && \text{by } (7.3.14) \\
&\le \sum_{\boldsymbol{k}}\sum_{\boldsymbol{v}} P[A_n(\boldsymbol{k},\boldsymbol{v})] && (=) \\
&\le \sum_{\boldsymbol{k}}\sum_{\boldsymbol{v}} P[A_n(\boldsymbol{k},\boldsymbol{v}) \mid B_n(\boldsymbol{k},\boldsymbol{v})] P[B_n(\boldsymbol{k},\boldsymbol{v})] \\
&\le \sum_{\boldsymbol{v}} P[\exists\, \boldsymbol{\gamma}(\boldsymbol{v}) \in \boldsymbol{\Gamma}(n,\boldsymbol{v})] \sum_{\boldsymbol{k}} P\left[T_{(n)} + \tau_{(K'_n)} < t_0\right] \\
&\le \mathrm{E}\mathcal{N}_n \sum_{\boldsymbol{k}} P\left[T_{(n)} + \tau_{(K'_n)} < t_0\right] && \text{by Prop. 7.3.1} \\
&\le C_3(\mathcal{P}q)^n && \text{for } n \ge n_0
\end{aligned}$$

where $C_3 = C_1C_2 < \infty$

$$\begin{aligned}
\Longrightarrow \qquad \sum_{n=1}^{\infty} P(A_n) &= C_4 + C_3 \sum_{n=n_0}^{\infty} (\mathcal{P}q)^n \\
&< \infty && \text{since } \mathcal{P}q < 1 \\
\Longrightarrow \quad P\left(\limsup_n A_n\right) &= 0
\end{aligned}$$

by the Borel-Cantelli Lemma.

□

7.4 CONVERGENT STABILITY OF THE GENERAL NETWORK

Theorem 7.4.1 (Existence of a Limiting Process)
Consider the general network — described in §7.2 — under the additional assumption that the service distribution F has an exponentially-decreasing tail, viz.

$$\exists\, s_0, \mu > 0 \text{ s.t. } s \geq s_0 \Rightarrow \overline{F}(s) \leq e^{-\mu s} \tag{7.4.1}$$

Then $\exists \alpha > 0$ such that the network with arrival stream $\boldsymbol{\xi}^0_\alpha = [\alpha t^v_1, \alpha\tau^v_n, S^v_n, \gamma^v_n]$ has equilibrium waiting-time distributions.

PROOF

As indicated in §7.3, we use the "bounded expectation" approach to prove this theorem.

Consider (7.2.5b). Since $\langle p_l \rangle$ is a positive sequence, it is clear that $\mathcal{P} > 1$. Let us choose q (different from that in §7.3) satisfying

$$0 < q < \frac{1}{2\mathcal{P}} < \frac{1}{2} \tag{7.4.2a}$$

Next find

$$\theta_0 > 1 \text{ s.t. } \theta_0 e^{1-\theta_0} < \frac{q}{2} \tag{7.4.2b}$$

which is always possible because $f(x) = xe^{1-x}$ has a maximum of 1 at $x = 1$ and thereafter decreases rapidly towards its asymptote of 0.

For s_0 and μ as in (7.4.1), let $T_0 > 2\left(s_0 + \frac{1}{\mu}\right) > 0$ such that

$$\theta_0 = \mu\left(\frac{T_0}{2} - s_0\right) \tag{7.4.2c}$$

Since G_1 and G are assumed to be continuous at 0, we have that

$$\begin{array}{rrcccl} \exists\, \alpha > 0 \text{ s.t.} & P[\alpha t_1^v < T_0'] & = & G_1(T_0) & < & \left(\frac{q}{2}\right)^2 \\ \text{and} & P[\alpha \tau^v < T_0'] & = & G(T_0) & < & \left(\frac{q}{2}\right)^2 \end{array} \tag{7.4.3}$$

where $T_0' = \alpha T_0$.

In view of (7.4.1), we may write $S_i = S_i' + S_i''$ where $S_i' \leq \zeta_i$, $S_i'' \leq s_0\ \forall i \in \mathbb{N}_n$ and $\langle\zeta_i\rangle$ is an IID Exp(μ^{-1}) sequence, which implies that

$$\sum_{i=1}^{n} S_i \leq \sum_{i=1}^{n} \zeta_i + n s_0 \tag{7.4.4}$$

Obviously $\sum\limits_{i=1}^{n} \zeta_i \sim \text{Gamma}(n, \mu^{-1})$.

We will need to consider $\mathrm{E}\left[\left(\sum\limits_{i=1}^{n} S_i\right) I_{[\sum S_i > mT_0]}\right]$ where $m \in [0, K_n]_{\mathbb{Z}_\oplus}$, as well as the sum $\sum\limits_{m=0}^{K_n} \binom{K_n}{m} \left(\frac{q}{2}\right)^{2(K_n - m)} \mathrm{E}\left[\left(\sum\limits_{i=1}^{n} S_i\right) I_{[\sum S_i > mT_0]}\right]$, the reason for which will become clearer in a short while. We separate $[0, K_n]_{\mathbb{Z}_\oplus}$ into two subintervals that, at first, are dealt with individually.

$\underline{Case\ 1}: \quad \underline{m \le \left\lfloor \frac{K_n}{2} \right\rfloor \le \frac{K_n}{2}} \;\Rightarrow\; -2m \ge -K_n$

$$\Rightarrow \left(\frac{q}{2}\right)^{-2m} \le \left(\frac{q}{2}\right)^{-K_n} \tag{7.4.5}$$

Obviously we are able to write

$$\mathrm{E}\left[\left(\sum_{i=1}^{n} S_i\right) I_{[\sum S_i > mT_0]}\right] \;\le\; \mathrm{E}\left(\sum_{i=1}^{n} S_i\right) \;=\; n\mathrm{E}(S_1) \tag{7.4.6}$$

so that

$$\sum_{m=0}^{\left\lfloor \frac{K_n}{2} \right\rfloor} \binom{K_n}{m} \left(\frac{q}{2}\right)^{2(K_n-m)} \mathrm{E}\left[\left(\sum_{i=1}^{n} S_i\right) I_{[\sum S_i > mT_0]}\right]$$

$$\le \; n\mathrm{E}(S_1) \sum_{m=0}^{\left\lfloor \frac{K_n}{2} \right\rfloor} 2^{K_n} \left(\frac{q}{2}\right)^{2K_n - K_n} \qquad (=) \quad \text{by (7.4.5) \& (7.4.6)}$$

$$\le \; n\left(\left\lfloor \frac{K_n}{2} \right\rfloor + 1\right) \mathrm{E}(S_1) q^{K_n}$$

$$\le \; c_3 n K_n q^{K_n} \tag{7.4.7}$$

$$\underline{Case\ 2}: \quad \underline{m \geq \left\lfloor \frac{K_n}{2} \right\rfloor + 1 \geq \frac{K_n}{2}} \quad \begin{aligned} &\Rightarrow \quad 2m \geq K_n \geq n \\ &\Rightarrow \quad mT_0 - ns_0 \; \geq mT_0 - 2ms_0 \\ &\qquad\qquad\qquad\quad\; = \tfrac{2m\theta_0}{\mu} \end{aligned} \tag{7.4.8}$$

$$\begin{aligned}
& \mathrm{E}\left[\left(\sum_{i=1}^{n} S_i\right) I_{[\sum S_i > mT_0]}\right] \\
\leq\ & \mathrm{E}\left[\left(\sum_{i=1}^{n} \zeta_i + ns_0\right) I_{[\sum \zeta_i > mT_0 - ns_0]}\right] && \text{by (7.4.4)} \\
\leq\ & \mathrm{E}\left[\left(\sum_{i=1}^{n} \zeta_i + ns_0\right) I_{[\sum \zeta_i > \frac{2m\theta_0}{\mu}]}\right] \qquad (=) && \text{by (7.4.8)} \\
\leq\ & \int_{\frac{2m\theta_0}{\mu}}^{\infty} (z + ns_0) \tfrac{\mu^n}{(n-1)!} z^{n-1} e^{-\mu z} dz \qquad (=) \\
\leq\ & n \int_{\frac{2m\theta_0}{\mu}}^{\infty} \tfrac{(\mu z)^n}{n!} e^{-\mu z} dz + n\mu s_0 \int_{\frac{2m\theta_0}{\mu}}^{\infty} \tfrac{(\mu z)^{n-1}}{(n-1)!} e^{-\mu z} dz \quad (=) \\
\leq\ & \tfrac{n}{\mu} \left[\int_{2m\theta_0}^{\infty} \tfrac{y^n}{n!} e^{-y} dy + \mu s_0 \int_{2m\theta_0}^{\infty} \tfrac{y^{n-1}}{(n-1)!} e^{-y} dy\right] \qquad (=) \\
\leq\ & \tfrac{n}{\mu} \left[\sum_{k=0}^{n} \tfrac{(2m\theta_0)^k}{k!} + \mu s_0 \sum_{k=0}^{n-1} \tfrac{(2m\theta_0)^k}{k!}\right] e^{-2m\theta_0} \\
\leq\ & \tfrac{n}{\mu} (1 + \mu s_0) e^{-2m\theta_0} \sum_{k=0}^{n} \tfrac{(2m\theta_0)^k}{k!} \\
\leq\ & \tfrac{n}{\mu} (1 + \mu s_0) \theta_0^n e^{-2m\theta_0} \sum_{k=0}^{n} \tfrac{(2m)^k}{k!} && \text{by (7.4.2b)} \\
\leq\ & \tfrac{n}{\mu} (1 + \mu s_0) \theta_0^{2m} e^{-2m\theta_0} e^{2m} \qquad (=) \\
\leq\ & \tfrac{n}{\mu} (1 + \mu s_0) \left(\theta_0 e^{1-\theta_0}\right)^{2m} \\
\leq\ & c_4 n \left(\tfrac{q}{2}\right)^{2m}
\end{aligned}$$

(7.4.9)

By using the fact that $q < 1$ several times, we have

$$\begin{aligned}
&\sum_{m=\lfloor \frac{K_n}{2} \rfloor+1}^{K_n} \binom{K_n}{m} \left(\tfrac{q}{2}\right)^{2(K_n-m)} \mathrm{E}\left[\left(\sum_{i=1}^{n} S_i\right) I_{[\sum S_i > mT_0]}\right] \\
\le\ & c_4 n \sum_{m=\lfloor \frac{K_n}{2} \rfloor+1}^{K_n} 2^{K_n} \left(\tfrac{q}{2}\right)^{2K_n} \left(\tfrac{q}{2}\right)^{-2m} \left(\tfrac{q}{2}\right)^{2m} \qquad \text{by (7.4.9)} \\
\le\ & c_4 n \sum_{m=\lfloor \frac{K_n}{2} \rfloor+1}^{K_n} \left(\tfrac{q^2}{2}\right)^{K_n} \\
\le\ & c_4 n \sum_{m=\lfloor \frac{K_n}{2} \rfloor+1}^{K_n} q^{K_n} \\
\le\ & c_5 n K_n q^{K_n}
\end{aligned} \tag{7.4.10}$$

On combining (7.4.7) and (7.4.10), we deduce that

$$\sum_{m=0}^{K_n} \binom{K_n}{m} \left(\tfrac{q}{2}\right)^{2(K_n-m)} \mathrm{E}\left[\left(\sum_{i=1}^{n} S_i\right) I_{[\sum S_i > mT_0]}\right] \ \le\ c n K_n q^{K_n} \tag{7.4.11}$$

Recall Proposition 7.3.2 and (7.3.15) involving the sum of n G_1-distributed r.v.s $\langle T_i \rangle$ and K_n' independent G-distributed random variables $\langle \tau_i \rangle$. Denote the joint sequence of these K_n variables, based on $\boldsymbol{t}$ in $\boldsymbol{\gamma}(\boldsymbol{k}, \boldsymbol{v}, \boldsymbol{t})$, by $\langle \sigma_i \rangle$, and let

$$\begin{aligned} I(\boldsymbol{k},\boldsymbol{v},m) &= I_{[|\{\sigma_i > T_0\}|=m]} \\ \Rightarrow \quad \mathrm{E}I(\boldsymbol{k},\boldsymbol{v},m) &= \binom{K_n}{m} P[|\{\sigma_i > T_0\}| = m] P[|\{\sigma_i \le T_0\}| = K_n - m] \\ &= \binom{K_n}{m} \left[1 - \left(\tfrac{q}{2}\right)^2\right]^m \left(\tfrac{q}{2}\right)^{2(K_n - m)} \quad \text{by (7.4.3)} \\ &\le \binom{K_n}{m} \left(\tfrac{q}{2}\right)^{2(K_n - m)} \end{aligned} \tag{7.4.12}$$

Next let

$$\begin{aligned} I_n^E(\boldsymbol{k},\boldsymbol{v}) &= I_{[\boldsymbol{\gamma}(\boldsymbol{k},\boldsymbol{v},\boldsymbol{t}) \in \boldsymbol{\Gamma}_E(n,\boldsymbol{k},\boldsymbol{v})]} \\ &\le I_{[\boldsymbol{\gamma}(\boldsymbol{k},\boldsymbol{v},\boldsymbol{t}) \in \boldsymbol{\Gamma}(n,\boldsymbol{k},\boldsymbol{v})]} \\ &\stackrel{\text{def}}{=} I_n(\boldsymbol{k},\boldsymbol{v}) \\ \Longrightarrow \quad \mathrm{E}I_n^E(\boldsymbol{k},\boldsymbol{v}) &\le \mathrm{E}I_n(\boldsymbol{k},\boldsymbol{v}) \\ &\le \mathrm{E}\mathcal{N}_n \qquad (=) \\ &\le \left[\sum_{\boldsymbol{v}} P[\exists\, \boldsymbol{\gamma}(\boldsymbol{v}) \in \boldsymbol{\Gamma}(n,\boldsymbol{v})]\right] \quad \text{by Prop. 7.3.1} \end{aligned} \tag{7.4.13}$$

We note that the random variable

(a) $\left(\sum_{i=1}^{n} S_i\right) I_{[\sum S_i > mT_0]}$ depends on $\langle S_i \rangle$, but not on the nature of the walks we consider

(b) $I(\boldsymbol{k}, \boldsymbol{v}, m)$ is associated with $\langle \sigma_i \rangle$

(c) $I_n(\boldsymbol{k}, \boldsymbol{v})$, on the other hand, *does* depend on the configuration of the various walks

Thus $\left(\sum_{i=1}^{n} S_i\right) I_{[\sum S_i > mT_0]}$, $I(\boldsymbol{k}, \boldsymbol{v}, m)$ and $I_n(\boldsymbol{k}, \boldsymbol{v})$ (unlike $I_n^E(\boldsymbol{k}, \boldsymbol{v})$) are independent.

We shall also use the fact that if the waiting time is obtained on an essential space chain of length n, then it is of the form $\left[\sum_{i=1}^{n} S_i - \sum_j \sigma_j\right]^+ \leq \sum_{i=1}^{n} S_i$. This leads to $I_{[\sum S_i > mT_0]} \geq I_{[\sum S_i > \sum \sigma_j]}$, a result also used in the sequel.

Hence, by applying (7.4.12), (7.4.13) and the independence of the three recently-mentioned random variables, we have the sequence of equalities and inequalities on the following page.

$$
\begin{aligned}
& \mathrm{E}W^{t_0}[\gamma^{v_0}] \\
= \;& \mathrm{E}\left[\sum_{n=1}^{\infty}\sum_{\boldsymbol{k}\in\mathbb{N}^n}\sum_{\boldsymbol{v}\in V^n} W^{t_0}[\gamma^{v_0}] I_n^E(\boldsymbol{k},\boldsymbol{v})\sum_{m=0}^{K_n} I(\boldsymbol{k},\boldsymbol{v},m)\right] \\
\leq \;& \mathrm{E}\left[\sum_{n=1}^{\infty}\sum_{\boldsymbol{k}}\sum_{\boldsymbol{v}} I_n(\boldsymbol{k},\boldsymbol{v})\sum_{m=0}^{K_n} I(\boldsymbol{k},\boldsymbol{v},m)\left(\sum_{i=1}^{n} S_i\right) I_{[\sum S_i > mT_0]}\right] \quad (=) \\
\leq \;& \sum_{n=1}^{\infty}\sum_{\boldsymbol{v}}\sum_{\boldsymbol{k}} E[I_n(\boldsymbol{k},\boldsymbol{v})]\sum_{m=0}^{K_n} E[I(\boldsymbol{k},\boldsymbol{v},m)] E\left[\left(\sum_{i=1}^{n} S_i\right) I_{[\sum S_i > mT_0]}\right] \\
\leq \;& \sum_{n=1}^{\infty}\left[\sum_{\boldsymbol{v}} P[\exists\, \boldsymbol{\gamma}(\boldsymbol{v}) \in \boldsymbol{\Gamma}(n,\boldsymbol{v})]\right]\sum_{\boldsymbol{k}}\sum_{m=0}^{K_n}\binom{K_n}{m}\left(\frac{q}{2}\right)^{2(K_n-m)} \mathrm{E}\left[\left(\sum_{i=1}^{n} S_i\right) I_{[\sum S_i > mT_0]}\right] \\
\leq \;& \sum_{n=1}^{\infty} C_1\mathcal{P}^n \sum_{\boldsymbol{k}} cnK_n q^{K_n} \\
\leq \;& C_1\sum_{n=1}^{\infty}\mathcal{P}^n\sum_{K=0}^{\infty}\sum_{\boldsymbol{k}\in\mathbb{N}^n : K_n = K+n} cnK_n q^{K_n} \\
\leq \;& C'\sum_{n=1}^{\infty} n\mathcal{P}^n\sum_{K=0}^{\infty}(K+n)\binom{K+n-1}{K} q^{K+n} \qquad C' = cC_1 \\
\leq \;& C'\sum_{n=1}^{\infty} n^2\mathcal{P}^n\sum_{K=0}^{\infty}\binom{K+n}{K} q^{K+n} \\
\leq \;& C'\sum_{n=1}^{\infty} n^2(\mathcal{P}q)^n\sum_{K=0}^{\infty} 2^{K+n} q^{K} \\
\leq \;& C\sum_{n=1}^{\infty} n^2(2\mathcal{P}q)^n \qquad C < \infty \text{ since } 2q < 1 \text{ by (7.4.2a)} \\
< \;& \infty \qquad \text{by (7.4.2a)}
\end{aligned}
$$

independently of t_0 and v_0, as required.

□

7.5 A SPECIAL CASE

We now consider a specific instance of the more general network described in §7.2. Our simplifying assumptions are as follows:

(a) (i) $V \simeq \mathbb{Z}$

(ii) $E = \{(v, v+1); v \in \mathbb{Z}\}$

(iii) $P[\gamma^v = \hat{\gamma}^v] = 1$ where $\hat{\gamma}^v = v,\ v+1,\ v+2$

As a result, for the associated essential space chain $\boldsymbol{\gamma}(\boldsymbol{k}, \boldsymbol{v}, \boldsymbol{t})$, $k_i = 1$ and $v_i - v_{i-1} \in \mathcal{M} = \{-1, 0, 1\}\ \forall\, i$.

(b) (i) The interarrival times $\langle \tau_n^v \rangle \sim \mathrm{Exp}(\lambda^{-1})$. Thus (Appendix B) $G_1(t) = G(t) = 1 - e^{-\lambda t}$.

(ii) The service times $\langle S_n^v \rangle \sim \mathrm{Exp}(\mu^{-1})$ so that $F(s) = 1 - e^{-\mu s}$.

The essential space chain of length n will consist of the vertices $v_0 + \sum_{i=1}^{k} \theta_i$, where $k \in \{0\} \cup \mathbb{N}_n$ and $\theta_i \in \mathcal{M}$. Much as in Chapter 6, a correspondence between the essential space chain and the vector $\boldsymbol{\theta} \in \mathcal{M}^n$ may be established. We shall write $\{\boldsymbol{\gamma} = \boldsymbol{\theta}\}$ for the event that the actual wait is obtained on the essential space chain corresponding to $\boldsymbol{\theta}$.

In proving the following theorem, we will also use other notation and ideas similar to those in Chapter 6.

Theorem 7.5.1

Suppose

$$\left.\begin{array}{lcccc} \inf_{a>0} \mathrm{E}e^{a(S-\tau)} & = & \inf_{a>0} \frac{\lambda\mu}{(\lambda+a)(\mu-a)} & = \; q' & < \; \frac{1}{3} \\ \text{or} \qquad \rho & < & 5-2\sqrt{6} & & \end{array}\right\} \tag{7.5.1}$$

Then there exists a limiting process of (stationary) waiting times.

<u>PROOF</u>

As we remarked earlier, it suffices to prove stability. Now for $T \in \mathbb{R}_{\oplus}$

$$\begin{aligned}
& \sum_{n=n_0+1}^{\infty} \sum_{\boldsymbol{\theta}\in\mathcal{M}^n} P\left[\{W^{t_0}[\gamma^{v_0}] > T\} \cap \boldsymbol{\gamma} = \boldsymbol{\theta}\right] \quad (=) \\
\leq \; & \sum_{n=n_0+1}^{\infty} \sum_{\boldsymbol{\theta}\in\mathcal{M}^n} P\left[W^{t_0}[\gamma^{v_0}] > T \mid \boldsymbol{\gamma} = \boldsymbol{\theta}\right] P[\boldsymbol{\gamma} = \boldsymbol{\theta}] \\
\leq \; & \sum_{n=n_0+1}^{\infty} \sum_{\boldsymbol{\theta}\in\mathcal{M}^n} P\left[\sum_{i=1}^{n} \left(S_i^{\boldsymbol{\theta}} - \tau_i^{\boldsymbol{\theta}}\right) > T\right] P[\boldsymbol{\gamma} = \boldsymbol{\theta}] \\
\leq \; & \sum_{n=n_0+1}^{\infty} \sum_{\boldsymbol{\theta}\in\mathcal{M}^n} P\left[\sum_{i=1}^{n} (S_i - \tau_i) > T\right] \\
\leq \; & \sum_{n=n_0+1}^{\infty} |\mathcal{M}|^n P\left[\sum_{i=1}^{n} (S_i - \tau_i) > 0\right] \\
\leq \; & \sum_{n=n_0+1}^{\infty} (3q')^n \qquad\qquad (=) \quad \text{by Cor. 6.2.1}b \\
\leq \; & \frac{(3q')^{n_0}}{1-3q'} \\
< \; & \frac{\epsilon}{2}
\end{aligned}$$

for n_0 sufficiently large.

Next, letting E_n denote the event that the actual wait is obtained on an essential space chain of length n, we have

$$\forall\, \epsilon > 0,\ \forall\, n \in \mathbb{N}_{n_0}\ \exists\, T_0 \text{ s.t. } T > T_0 \ \Rightarrow\ P\left[\{W^{t_0}[\gamma^{v_0}] > T\} \cap E_n\right] \ < \ \frac{\epsilon}{2n_0}$$

Thus $\forall\, \epsilon > 0,\ \exists\, T_0$ s.t. $T > T_0 \Rightarrow$

$$\begin{aligned} P[W^{t_0}[\gamma^{v_0}] > T] &= \sum_{n=1}^{\infty} P\left[\{W^{t_0}[\gamma^{v_0}] > T\} \cap E_n\right] \\ &= \sum_{n=1}^{n_0} \tfrac{\epsilon}{2n_0} + \sum_{n=n_0+1}^{\infty} \sum_{\boldsymbol{\theta} \in \mathcal{M}^n} P\left[\{W^{t_0}[\gamma^{v_0}] > T\} \cap \boldsymbol{\gamma} = \boldsymbol{\theta}\right] \\ &< \epsilon \end{aligned}$$

□

Remark 7.5.1

There is an obvious connection between (7.5.1) and Remark 6.2.1, where the derivation (for the exponential case) of ρ from $\inf_{a>0} \mathrm{E}e^{a(S-\tau)}$ is given.

CHAPTER 8

THE *FAAFS* DISCIPLINE ON A STARLIKE NETWORK

8.1 INTRODUCTION

As its title suggests, this last, and very short, chapter is concerned with the $FAAFS$ discipline introduced in Chapter 4, where details on the starlike network were also provided.

In §8.2 we briefly review results — from Berezner *et al.* [9] — pertaining to the nonpreemptive and preemptive (resume) models.

It is indeed appropriate that the final chapter of the thesis be devoted to new and exciting research.

8.2 THE *FAAFS* PROTOCOL

While the $FAAFS$-type discipline may be the most effective in practical terms, it is the most difficult of the three protocols of Chapter 4 to analyse mathematically. In the first place, waiting-times are system-dependent: they are determined not only by the times of arrival $\{t\}$ to the network, but also by the shifted Poisson process $\{t + w^{[1]}\}$.

In the $FAAFS_N$ starlike network, the waiting time, $W(y)$, of some call y may be derived partially or completely from future arrivals. This dependence on the future required us to estimate the workload on the quasidelayer, rather than the true delayer, path (Definition 4.3.5). As a result, we were able to prove the stability/uniform boundedness of the network under condition (6.2.3), but not convergence to a stationary regime.

In our analysis of the $FAAFS_P$ network we were faced with the problem of "indirect conditional future dependence". In its second-stage of processing, a call y will experience an initial wait (possibly 0), followed by a period of service which is likely to be interrupted before the service is complete. Consequently, the call y will again wait before completing the next segment of its service. This wait-service cycle for any call y is mathematically problematic because it results in there being a random number of calls which make a direct contribution to the workload of y. Again, therefore, we were not in a position to prove convergence, and we had to be satisfied with uniform stability.

We may summarise the main result of Berezner *et al.* [9] as follows:

Theorem 8.2.1 (The Circuit-Switched Network)
Suppose that for some $\lambda > 0$

$$\inf_{a>0} \mathrm{E}e^{aS}\frac{\lambda}{\lambda + a} < \frac{1}{2} \tag{8.2.1}$$

Then the (nonpreemptive or preemptive) $FAAFS$ starlike network is stable *uniformly* in the size (N) of the network.

Although the proof of Theorem 8.2.1 could proceed in a similar way for both the nonpreemptive and preemptive forms of the $FAAFS$ discipline, an approach based on the direct comparison $W_{FAAFS} \leq W_{FAFS}$ (between the waiting times of the $FAAFS_P$ and $FAFS$ models) was adopted in Berezner *et al.* [9].

The analysis of the message-switched network is more tractable. In particular, several results — of especial interest when one considers the simulation study data given in Chapter 4 — were proved about this network. The most useful of these are included in the theorem below.

Theorem 8.2.2 (The Message-Switched Network)

(a) For each of the three disciplines, $FCFS$, $FAFS$ and $FAAFS$, the first stage of the message-switching network operates precisely as the first stage of the network under the other two disciplines. In particular, each queue in the first stage is simply of the form $GI/GI/1/\infty/FCFS$.

(b) Suppose now that $\rho^{[1]} < 1$ and $\rho^{[2]} < 1$. Then for both the $FCFS$ and $FAAFS$ networks, $\left(w_n^{[1]}, w_n^{[2]}\right) \overset{n}{\underset{\infty}{\to}} \left(w^{[1]}, w^{[2]}\right)$ and $\mathrm{E}w_{FCFS}^{[i]} = \mathrm{E}w_{FAAFS}^{[i]}$ $(i = 1, 2)$.

The research continues ...

REFERENCES

References with author(s) listed in **boldface** are those researched in relatively extensive detail.

[1] Apostol, T.M. *Mathematical Analysis,* 2nd ed. Addison-Wesley, Reading, Massachusetts, 1974.

[2] Arnold, V.I. *Geometrical Methods in the Theory of Ordinary Differential Equations,* 2nd ed. Springer-Verlag, New York, 1988.

[3] Baccelli, F., Gelenbe, E. and Plateau, B. An end-to-end approach to the resequencing problem. *Journal of the Association for Computing Machinery* **31**(3), 474–485, 1984.

[4] **Baccelli, F., Karpelevich, F.I., Kelbert, M.Ya., Puhalskii, A.A., Rybko, A.N.** and **Suhov, Yu.M.** A mean-field limit for a class of queueing networks (Originally: Asymptotical study of centralized message-switched networks under synchronization constraints). *Journal of Statistical Physics* **66**(3–4), 803–825, 1992.

[5] **Baccelli, F.** and **Makowski, A.M.** Queueing models for systems with synchronization constraints. *Proceedings of the Institute of Electrical and Electronics Engineers* **77**(1), 138–161, 1989.

[6] Baccelli, F., Makowski, A.M. and Shwartz, A. The fork-join queue and related systems with synchronization constraints: stochastic ordering and computable bounds. *Advances in Applied Probability* **21**(3), 629–660, 1989.

[7] Baccelli, F., Massey, W.A. and Towsley, D. Acyclic fork-join queuing networks. *Journal of the Association for Computing Machinery* **36**(3), 615–642, 1989.

[8] **Berezner, S.A.** and **Malyshev, V.A.** The stability of infinite-server networks with random routing. *Journal of Applied Probability* **26**(2), 363–371, 1989.

[9] Berezner, S.A., Rose D.M. and Suhov, Yu.M. Starlike networks with synchronization constraints. Submitted to the *IMA Series in Probability* , 30pp., 1994.

[10] Bernussou, J. and Garcia, J.M. Switching networks: real-time management. In *Systems & Control Encyclopedia*, ed. Singh, M.G., Vol. 7, 4733–4741. Pergamon Press, Oxford, 1987.

[11] Billingsley, P. *Convergence of Probability Measures.* Wiley, New York, 1968.

[12] Borovkov, A.A. *Stochastic Processes in Queueing Theory.* Springer-Verlag, New York, 1976.

[13] Borovkov, A.A. Ergodicity and stability theorems for a class of stochastic equations and their applications. *Theory of Probability and its Applications* **23**(2), 227–247, 1978.

[14] Borovkov, A.A. *Asymptotic Methods in Queuing Theory.* Wiley, Chichester, 1984.

[15] **Brandt, A.** On stationary waiting times and limiting behaviour of queues with many servers I: The general $G/G/m/\infty$ case. *Elektronische Informationsverarbeitung und Kybernetik* **21**(1–2), 47–64, 1985.

[16] Brandt, A. On stationary waiting times and limiting behaviour of queues with many servers II: The general $G/GI/m/\infty$ case. *Elektronische Informationsverarbeitung und Kybernetik* **21**(3), 151–162, 1985.

[17] Brandt, A. and Lisek, B. On the continuity of $G/GI/m$ queues. *Mathematische Operationsforschung und Statistik Series Statistics* **14**(4), 577–587, 1983.

[18] Bronshtein, I.N. and Semendyayev, K.A. *Handbook of Mathematics,* 3rd ed. Verlag, Thun and Frankfurt, 1985.

[19] Brown, T.C. and Pollett, P.K. Some distributional approximations in Markovian queueing networks. *Advances in Applied Probability* **14**(3), 654–671, 1982.

[20] Butkov, E. *Mathematical Physics.* Addison-Wesley, Reading, Massachusetts, 1968.

[21] Castel, C. and Le Gall, F. Switching networks: modelling. In *Systems & Control Encyclopedia*, ed. Singh, M.G., Vol. 7, 4726–4732. Pergamon Press, Oxford, 1987.

[22] Chemouil, P. and Gauthier, P. Switching networks: traffic routing and real-time control. In *Systems & Control Encyclopedia*, ed. Singh, M.G., Vol. 7, 4749–4751. Pergamon Press, Oxford, 1987.

[23] Cohen, J.W. *The Single Server Queue.* North-Holland, Amsterdam, 1982.

[24] Cooper, R.B. *Introduction to Queueing Theory,* 2nd ed. Edward Arnold, London, 1981.

[25] de Souza e Silva, E. and Muntz, R.R. Queueing networks: solutions and applications. In *Stochastic Analysis of Computer and Communication Systems*, ed. Takagi, H., 319–399. North-Holland, Amsterdam, 1990.

[26] Despres, R. Packet-switching networks. In *Systems & Control Encyclopedia*, ed. Singh, M.G., Vol. 6, 3581–3590. Pergamon Press, Oxford, 1987.

[27] Dettman, J.W. *Mathematical Methods in Physics and Engineering,* 2nd ed. Dover, New York, 1988.

[28] Devaney, R.L. *An Introduction to Chaotic Dynamical Systems.* Addison-Wesley, Redwood City, California, 1987.

[29] Dobrushin, R.L. and Sukhov, Yu.M. An asymptotical study of star-like message switching networks with a large number of radial rays [Russian]. *Problemy Peredachi Informatsii* **12**(1), 70–94, 1976.

[30] Doob, J.L. *Stochastic Processes.* Wiley, New York, 1953.

[31] Franken, P., König, D., Arndt, U. and Schmidt, V. *Queues and Point Processes.* Wiley, Chichester, 1982.

[32] Gross, D. and Harris, C.M. *Fundamentals of Queueing Theory,* 2nd ed. Wiley, New York, 1985.

[33] Hardy, G.E., Littlewood, J.E. and Pólya, G. *Inequalities.* Cambridge University Press, London, 1934.

[34] Jarry-Lacombe, B. and Legros, J.F. Switching networks: short-term planning. In *Systems & Control Encyclopedia*, ed. Singh, M.G., Vol. 7, 4741–4749. Pergamon Press, Oxford, 1987.

[35] **Karpelevich, F.I., Kelbert, M.Ya.** and **Suhov, Yu.M.** Higher-order Lindley equations (Originally: A queueing network model leading to a second-order Lindley equation *and* Higher-order Lindley equations: existence and non-uniqueness of a solution). Submitted to *Stochastic Processes and their Applications* , 30pp., 1993.

[36] Kelbert, M.Ya., Kontsevich, M.E. and Rybko, A.N. On Jackson networks on denumerable graphs. *Theory of Probability and its Applications* **33**(2), 358–361, 1988.

[37] Kelbert, M.Ya., Kopeika, R.P., Shamsiev, R.N. and Sukhov, Yu.M. Perturbation theory approach for a class of hybrid switching networks with small transit flows. *Advances in Applied Probability* **22**(1), 211–229, 1990.

[38] Kelbert, M.Ya. and Sukhov, Yu.M. Mathematical theory of queuing networks. *Journal of Soviet Mathematics* **50**(3), 1527–1600, 1990.

[39] Kelly, F.P. One-dimensional circuit-switched networks. *The Annals of Probability* **15**(3), 1166–1179, 1987.

[40] Kendall, D.G. Some problems in the theory of queues. *Journal of the Royal Statistical Society, Series* B **13**(2), 153–185, 1951.

[41] Kiefer, J. and Wolfowitz, J. On the theory of queues with many servers. *Transactions of the American Mathematical Society* **78**, 1–18, 1955.

[42] Kleinrock, L. *Queueing Systems,* Vol. 1: *Theory.* Wiley, New York, 1975.

[43] Kleinrock, L. *Queueing Systems,* Vol. 2: *Computer Applications.* Wiley, New York, 1976.

[44] Kolmogorov, A.N. and Fomin, S.V. *Introductory Real Analysis.* Dover, New York, 1975.

[45] Lindley, D.V. The theory of queues with a single server. *Proceedings of the Cambridge Philosophical Society* **48**(2), 277–289, 1952.

[46] Loève, M. *Probability Theory II,* 4th ed. Springer-Verlag, New York, 1978.

[47] **Loynes, R.M.** The stability of a queue with non-independent inter-arrival and service times. *Proceedings of the Cambridge Philosophical Society* **58**(3), 497–520, 1962.

[48] Narendra, K.S. Switching networks: learning algorithms for control policy. In *Systems & Control Encyclopedia*, ed. Singh, M.G., Vol. 7, 4716–4721. Pergamon Press, Oxford, 1987.

[49] Nelson, R. and Tantawi, A.N. Approximate analysis of fork/join synchronization in parallel queues. *IEEE Transactions on Computers* **37**(6), 739–743, 1988.

[50] Peyrade, M. Switching networks: general structures. In *Systems & Control Encyclopedia*, ed. Singh, M.G., Vol. 7, 4710–4716. Pergamon Press, Oxford, 1987.

[51] Rosenblatt, M. Dependence and asymptotic independence for random processes. In *Studies in Probability Theory*, ed. Rosenblatt, M., 24–45. The Mathematical Association of America, U.S.A., 1978.

[52] Rubin, I. and Baker, J.E. Media access control for high-speed local area and metropolitan area communications networks. *Proceedings of the Institute of Electrical and Electronics Engineers* **78**(1), 168–203, 1990.

[53] Rybko, A.N. and Mikhailov, V.A. Range of carrying capacity for communication networks with channel switching [Russian]. *Problemy Peredachi Informatsii* **22**(1), 66-84, 1986.

[54] Saaty, T.L. *Elements of Queueing Theory With Applications.* Dover, New York, 1983.

[55] Saito, H. and Kawashima, K. Performance modeling of integrated voice and data communications networks. In *Stochastic Analysis of Computer and Communication Systems*, ed. Takagi, H., 807–849. North-Holland, Amsterdam, 1990.

[56] Shiryayev, A.N. *Probability.* Springer-Verlag, New York, 1984.

[57] Tobagi, F.A. Fast packet switch architectures for broadband integrated services digital networks. *Proceedings of the Institute of Electrical and Electronics Engineers* **78**(1), 133–167, 1990.

[58] Venkatasubramanyan, N. and Sanders, J.L. Order restoration in queues. *Queueing Systems* **12**, 409–429, 1992.

[59] Wolff, R.W. *Stochastic Modeling and the Theory of Queues.* Prentice-Hall, Englewood Cliffs, New Jersey, 1989.

APPENDIX A
GRAPHS

A.1 INTRODUCTORY CONCEPTS

The first two definitions given below use the term *graph* in a fairly general sense.

Definition A.1.1 (Graph)
A *graph* $\mathcal{G} = [V, E]$ comprises a nonempty set of *vertices* or *nodes* (or *points*), $V = V(\mathcal{G})$, called the *vertex set* of $\mathcal{G}$, together with a possibly empty *edge set*, $E = E(\mathcal{G})$, of pairs of vertices of $\mathcal{G}$, called *edges* (or *lines*).

Definition A.1.2 (Classes of Graphs)
If a graph $\mathcal{G}$ may have *loops*, i.e., edges $e = vv$ joining a vertex to itself, then $\mathcal{G}$ is called a *loop-graph.*
A graph $\mathcal{G}$ in which *multiple* or *parallel edges* (more than one edge joining the same pair of vertices) are permissible, but loops are not, is termed a *multigraph.* The number of parallel edges joining a distinct pair of vertices is called the *multiplicity* of the "edge".
Should both loops and multiple edges be allowed in graph $\mathcal{G}$, then $\mathcal{G}$ is a *pseudograph.*
If $\mathcal{G}$ has edge set $E(\mathcal{G})$ of *ordered* pairs of vertices (so $e = (v_1, v_2) = \overrightarrow{v_1 v_2}$), then $\mathcal{G}$ is a *directed graph* or *digraph* with *directed edges* or *arcs*, and *arc set* $E(\mathcal{G})$.

That graph $\mathcal{G}$ which has an edge set of *unordered* pairs of vertices ($e = v_1v_2$ or, equivalently, $e = v_2v_1$) and has neither loops nor multiple edges is referred to as a *simple graph* or merely a *graph.*

Definition A.1.3 (Symmetric and Asymmetric Digraphs)
Let $\mathcal{D}$ be a digraph with arc set $E(\mathcal{D})$.
Then $\mathcal{D}$ is a *symmetric digraph* iff

$$(u,v) \in E(\mathcal{D}) \Leftrightarrow (v,u) \in E(\mathcal{D}).$$

$\mathcal{D}$ is an *asymmetric digraph* or *oriented graph* iff

$$(u,v) \in E(\mathcal{D}) \Rightarrow (v,u) \notin E(\mathcal{D}).$$

Definition A.1.4 (Underlying Graphs)
The *underlying graph of a digraph* $\mathcal{D}$ is that graph $\mathcal{G}$ obtained from $\mathcal{D}$ by replacing each arc $a = (u,v)$ of $\mathcal{D}$ by the edge $e = uv$ and deleting any occurrence of a multiple edge.
The *underlying graph of a pseudograph* $\mathcal{G}$ is that graph $\mathcal{G}^*$ obtained from $\mathcal{G}$ by deleting all loops and replacing any set of multiple edges of $\mathcal{G}$ with a single edge.

Definition A.1.5 (Edges and Vertices)
If $|V(\mathcal{G})| = p$ and $|E(\mathcal{G})| = q$, then p and q are called the *order* and *size* respectively of $\mathcal{G}$ (a graph or digraph). Then $\mathcal{G}$ is a (p,q)-(di)graph.
If each of p and q is finite, then $\mathcal{G}$ is termed a *finite graph*; otherwise it is an *infinite graph.*

Suppose $\mathcal{G}_1$ is a simple graph with edges $e = uv$ and $f = vw$. Then e is said to *join* u and v, which are called *adjacent vertices.* Each of u and v is *incident* with e. Furthermore, e and f are *adjacent edges* (because they have a vertex in common).

Next consider a digraph $\mathcal{G}_2$ with arcs $a_1 = (u, v)$ and $a_2 = (v, w)$. The terminology used is similar to that above except that the prepositions *to* and *from* are used to convey the direction of the arcs. Vertex u is called the *initial endpoint* or *tail* of arc a_1; v is its *terminal endpoint* or *head.*

Remark A.1.1 (Diagram of a Graph)

It is customary to represent a graph by means of a diagram, which is often also referred to as a graph. Line segments or curves are used for (undirected) edges; arcs are represented by directed (or "arrowed") line segments. Vertices are indicated by small circles or dots.

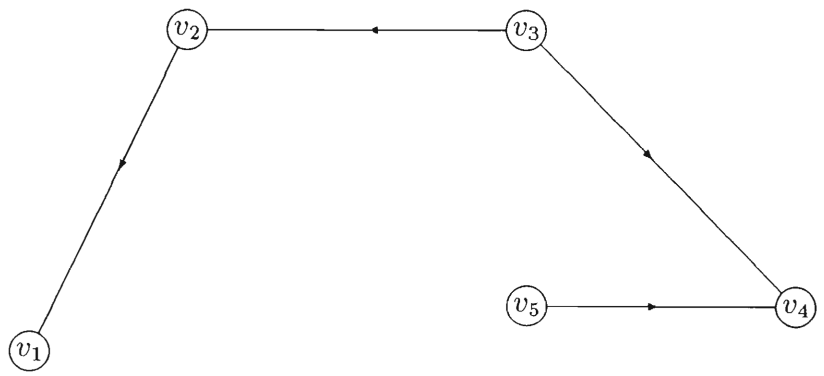


Fig. A.1

(5, 4)-Simple Asymmetric (Finite) Digraph

Vertex set $V(\mathcal{D}) = \{v_1, v_2, v_3, v_4, v_5\}$

Arc set $E(\mathcal{D}) = \{(v_2, v_1); (v_3, v_2); (v_3, v_4); (v_5, v_4)\}$

A.2 CONNECTEDNESS IN GRAPHS

Definition A.2.1 (Walks, Paths and Cycles)

Let $\mathcal{G}$ be a simple graph.

A *walk* $W : v_0 e_1 v_1 \ldots v_{n-1} e_n v_n$ of $\mathcal{G}$ is a (finite) alternating sequence of edges and vertices, beginning and ending with vertices, in which $e_i = v_{i-1} v_i \ \forall\, i \in \mathbb{N}_n$. The *length* of W is n (the number of occurrences of edges).

Without ambiguity, we may also write

$$W : v_0\, v_1\, \ldots\, v_{n-1}\, v_n.$$

v_0 and v_n are the *initial* and *terminal vertices* respectively. The remaining vertices are called *inner* or *intermediate vertices.*

W is sometimes referred to as a v_0-v_n *walk*; it is *closed* iff $v_0 = v_n$ and *open* otherwise.

On the other hand, a *path* of $\mathcal{G}$ is a walk in which vertices (and thus, necessarily, all edges) are distinct.

W is a *cycle* of $\mathcal{G}$ iff it is a closed walk of length $n \geq 3$ with no repetitions of vertices allowed (except, of course, that we have $v_0 = v_n$).

Remark A.2.1 (Walks in Digraphs)

The definitions of a walk, path and cycle in a digraph are analogous to those for a graph, except that in a digraph *we always proceed in the direction of the arcs.* To emphasise this fact, we may use the prefix *di-*, for instance, *diwalk.*

Definition A.2.2 (Connectedness in Graphs & Digraphs)

Let u and v be vertices in a graph $\mathcal{G}$.

Then u is *connected to* v iff $\mathcal{G}$ contains a u-v path; $\mathcal{G}$ itself is *connected* iff u is connected to v $\forall\ (u, v) \in [V(\mathcal{G})]^2$.

A digraph $\mathcal{D}$ is *connected* iff its underlying graph is connected.

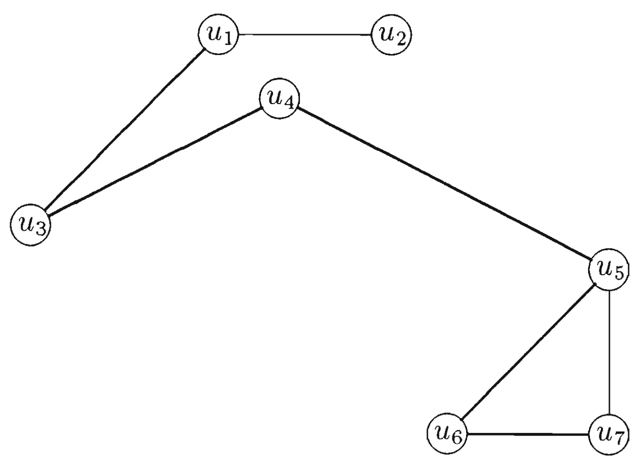


Fig. A.2

$\mathcal{P} : u_1\, u_3\, u_4\, u_5\, u_6\, u_7$

A path of length 5 in a connected graph

A.3 TREES

This section and the next are concerned with two special types of graphs which are of particular interest in this thesis.

Definition A.3.1 (Tree)

Let $\mathcal{G}$ be a graph of order p.

Then $\mathcal{G}$ is a *tree* iff one of the following equivalent conditions holds.

(i) $\mathcal{G}$ is connected and has size $p-1$.

(ii) $\mathcal{G}$ is acyclic and has size $p-1$.

(iii) $\mathcal{G}$ is a minimal connected graph $(p \geq 2)$.

(iv) $\mathcal{G}$ is a maximal acyclic graph $(p \geq 2)$.

Definition A.3.2 (Directed Tree)

A *directed tree* is an asymmetric digraph, the underlying graph of which is a tree.

Definition A.3.3 (Rooted Tree)

A *rooted tree* is a directed tree $\mathcal{T}$ with a distinguished vertex r, called the *root*, with the property that $\mathcal{T}$ contains an r-v path $\forall\, v \in V(\mathcal{T})$.

Definition A.3.4 (n-Trees)

Let $\mathcal{T}$ be a rooted tree with root r.

For a fixed vertex v of $\mathcal{T}$, let P_v be the shortest r-v path, say of length l.

Consider now another vertex u which lies on P_v, such that the shortest r-u path has length $l-1$. Vertex v is then known as the *successor* of u.
A simple example of $\mathcal{T}$ is one in which each vertex has exactly n successors; such a rooted tree is called an *n-tree*.

Remark A.3.1 (Branching Number)
The *branching number* of an n-tree is n. For a more general rooted tree, the idea of a branching number is more complicated. Intuitively-speaking, if we suitably define the "average" number of branches issuing from a typical vertex of our tree, we obtain its branching number. The interested reader is referred to the article of Lyons [A.1] for more details.

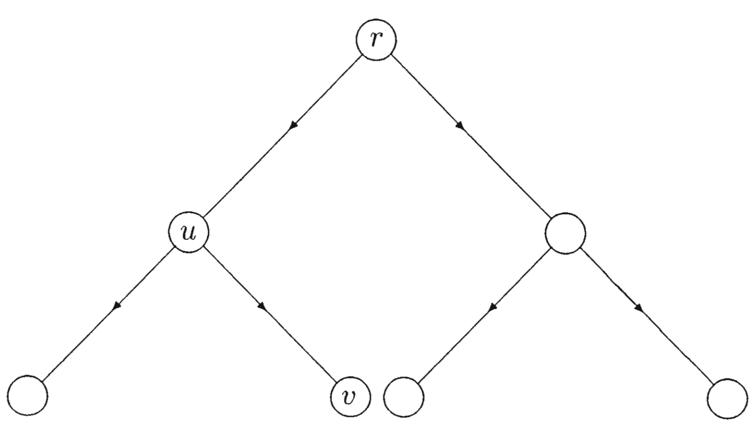


Fig. A.3

A rooted tree with root r (and branching 2)

v is a successor of u

A.4 n-PARTITE GRAPHS

Definition A.4.1 (n-Partite Graphs)

Let $n \in \mathbb{N}$. A graph $\mathcal{G}$ is an *n-partite graph* iff there exists a partition of its vertex set: $V(\mathcal{G}) = \bigcup_{i=1}^{n} V_i \;\wedge\; V_j \cap V_k \neq \emptyset \Rightarrow j = k$ such that

$$uv \in E(\mathcal{G}) \Longrightarrow u \in V_j \;\wedge\; v \in V_k \;\wedge\; j \neq k.$$

If $n = 2$, then $\mathcal{G}$ is a called a *bipartite graph.*

In the following definition we assume that $|V_i| = m_i \;\forall\, i \in \mathbb{N}_n$ (so that $|V(\mathcal{G})| = \sum_{i=1}^{n} m_i$).

Definition A.4.2 (Complete n-Partite Graphs)

Let $n \in \mathbb{N}$. A graph $\mathcal{G}$ is a *complete n-partite graph*, denoted by $K(m_1, \ldots, m_n)$ or $K_{m_1,\ldots,m_n}$, iff it is an n-partite graph and $u \in V_j \;\wedge\; v \in V_k \;\wedge\; j \neq k \Longrightarrow uv \in E(\mathcal{G})$.

Definition A.4.3 (Star Graph)

Let $N \in \mathbb{N}$. The complete bipartite graph $K_{1,N}$ is called a *star* (*graph*).

Remark A.4.1 (Star Network)

In communications applications (such as computing), a *star network* is one in which there is a single central coordinating node, and distant devices are connected along communications lines which radiate like spokes from the central node.

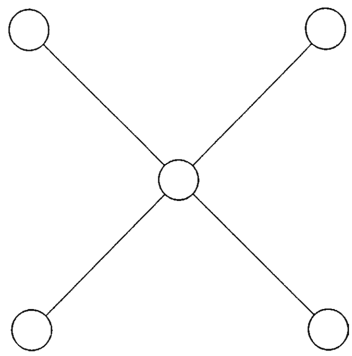


Fig. A.4

Star graph $K_{1,4}$

REFERENCE

[A.1] Lyons, R. Random Walks and Percolation on Trees. *The Annals of Probability* **18**(3), 931–958, 1990.

APPENDIX B
POINT PROCESSES

B.1 INTRODUCTORY CONCEPTS

Consider a Polish space E (Franken *et al.* [B.1, *p.* 62], Kallenberg [B.3]) with Borel σ-algebra $\mathcal{E}$. Write $\mathcal{B}$ for the ring of bounded (relatively compact) sets in $\mathcal{E}$.

Definition B.1.1 (Measures on E)
Let ϕ be a measure on E.

(a) The set of *locally finite (Radon) measures* on E is given by

$$M = \{\phi \in E : \phi(B) < \infty \; \forall \, B \in \mathcal{B}\}$$

(b) The collection of *point* or *counting measures* is defined by

$$M_p = \{\phi \in M : \phi(B) \in \mathbb{Z}_{\oplus} \; \forall \, B \in \mathcal{B}\}$$

(c) The set of *simple point measures* is given by

$$M_s = \{\phi \in M_p : \phi(\{e\}) \leq 1 \; \forall \, e \in E\}$$

Write $\mathcal{M}, \mathcal{M}_p$ and $\mathcal{M}_s$ for the σ-algebras generated by the sets given in (a), (b) and (c) respectively.

Remark B.1.1

For convenience we write $\phi(e)$ for $\phi(\{e\})$.

Definition B.1.2 (Dirac Measure)

Let $e \in E$ be fixed, and $B \in \mathcal{B}$. Then

$$\delta_e(B) = I_B(e) = \begin{cases} 1 & e \in B \\ 0 & e \notin B \end{cases}$$

The fundamental point measure $\delta_e(\cdot)$ is called the *Dirac measure.*

B.2 RANDOM POINT PROCESSES

Definition B.2.1 (Point Process)

A *(random) point process* Φ is a random element of the measurable space $(M_p, \mathcal{M}_p)$. Consequently Φ may be given as a probability space $(M_p, \mathcal{M}_p, P)$ which implies that Φ is a *random counting measure* on $\mathcal{B}$.

Remark B.2.1 (Alternative Definition of a Point Process)

An alternative definition, which is most appropriate when some other random process generates point process Φ, is often encountered in the literature. In this case, given a probability space $(\Omega, \mathcal{F}, P)$, Φ is defined as the measurable mapping $(\Omega, \mathcal{F}) \mapsto (M_p, \mathcal{M}_p)$.

Definition B.2.2 (Simple Point Process)

A point process Φ with associated probability space $(M_p, \mathcal{M}_p, P)$ is said to be *simple* iff $P[\Phi \in M_s] = 1$.

$\forall\, B \in \mathcal{B}$ define the $\mathbb{Z}_\oplus$-valued random variable $N(B)$ by

$$P[N(B) = n] = P[\phi \in M_p : \phi(B) = n] \quad (n \in \mathbb{Z}_\oplus)$$

Thus $N(B)$ gives the random number of points of Φ in B.

Remark B.2.2 (Point Processes on $\mathbb{R}$)
When Polish space $E = \mathbb{R}$ (or $\mathbb{R}_\oplus$), then the random points of the process Φ are most commonly times of events $\langle t_n, n \in \mathbb{Z}\rangle$ (or $\langle t_n, n \in \mathbb{N}\rangle$) (for instance, moments of arrivals to a network) and $\tau_n = t_{n+1} - t_n$ are interevent (interarrival) times. The sequence $\langle \tau_n \rangle$ constitutes the *incremental process* of $\langle t_n \rangle$. In the time-context, a simple point process (informally) is one in which $\tau_n > 0\ \forall\, n$, so that there are a.s. no points where two or more events occur simultaneously. (To understand why, consider Definition B.1.1(c). Since ϕ is a counting measure, and for every t $\phi(t) \leq 1$, $\phi(t) = 0$ or 1 (and no greater).)

When the (time) points are nonnegative, one conventionally writes $N(t) = N((0,t])$ for the number of events (arrivals) in the bounded interval $(0,t]$. It is obvious that $N(t) \geq n \Leftrightarrow t_n \leq t$.

We now briefly consider a well-known class of point processes.

Example B.2.1 (Renewal and Poisson Processes)
A *renewal process* Φ is a point process characterised by an IID incremental sequence $\langle \tau_n, n \in \mathbb{N}\rangle \sim G$ which is independent of the time $t_1 \sim G_1$ of the first event. When G_1 coincides with G, then Φ is said to be an *ordinary*

renewal process; otherwise it is called a *modified* or *delayed renewal process.*

Φ is an *equilibrium* or *stationary renewal process* iff

$$G_1(s) = \frac{1}{\mathrm{E}(\tau)} \int_0^s \overline{G}(x)dx \qquad (\mathrm{E}\tau < \infty) \tag{B.2.1}$$

iff the process has stationary increments ($L[N(s+h) - N(s)] = L[N(h)]$ $\forall\, s \geq 0, \forall\, h > 0$). The equivalence of these two conditions is demonstrated in Grimmett & Stirzaker [B.2, *pp.* 295–296].

When one considers a renewal process, two quantities of interest are the *age* (or *current lifetime*) at t and the *excess lifetime* at t, given by $\delta(t) = t - t_{N(t)}$ and $\gamma(t) = t_{N(t)+1} - t$ respectively. In other words, $\delta(t)$ is the time since the last event before t and $\gamma(t)$ is the time to the next event after t. It is interesting to note that, if t_1 is nonlattice, then

$$\lim_{t\to\infty} P[\gamma(t) \leq s] = \frac{1}{\mathrm{E}(\tau)} \int_0^s \overline{G}(x)dx \tag{B.2.2}$$

which is the same distribution as in (B.2.1).

A renewal process of particular importance (certainly in queueing theory) is the *Poisson process*, which is usually introduced by means of postulates or axioms (e.g., Parzen [B.5, *p.* 118]). For the purposes of this discussion, it suffices to note that $G(s) = 1 - e^{-\lambda s}$ characterises the Poisson process with rate λ. (See, for example, Medhi [B.4, Chapter 4].)

According to its axiomatic definition, the Poisson process has stationary increments so that it is valid to consider (B.2.1) which reduces to

$$G_1(s) = \lambda \int_0^s e^{-\lambda x} dx = 1 - e^{-\lambda s} = G(s)$$

which means, of course, that we are concerned with an ordinary renewal process, and, using (B.2.2), that $\lim_{t\to\infty} P[\gamma(t) \leq s] = G(s)$. Parzen [B.5, *p.* 173] proves a more significant result (without the need for a limit), namely that $P[\gamma(t) \leq s] = 1 - e^{-\lambda s}$.

Before we proceed to investigate the concept of a *marked* point process, we outline what is meant by a Palm distribution $\mathcal{P}_\cdot$. Historically this distribution was first introduced for (stationary) simple point processes Φ (on the real line), in which case $\mathcal{P}_t$ may be regarded as the conditional distribution of Φ given that an event (arrival) occurs at t. For instance, the man after whom this distribution was named investigated the conditional probability of the absence of telephone calls in the interval $(0, t)$, given the occurrence of a call at time 0.

Definition B.2.3 (Palm Distributions and Probabilities)
The *Palm distributions* $\mathcal{P}_e$ of point process Φ are given by

$$\mathcal{P}_e(L) = \frac{\mathrm{E}\left(d\Phi(e) I_{\{\Phi \in L\}}\right)}{\mathrm{E}\left(d\Phi(e)\right)} \qquad L \in \mathcal{M},\ e \in E \text{ (a.s.} - \mathrm{E}\Phi)$$

More generally, *Palm probabilities* are defined by

$$\mathcal{P}_e(A) = \frac{\mathrm{E}\left(d\Phi(e) I_{\{\Phi \in A\}}\right)}{\mathrm{E}\left(d\Phi(e)\right)} \qquad A \in \mathcal{F},\ e \in E \text{ (a.s.} - \mathrm{E}\Phi)$$

Remark B.2.3 (Campbell Measure)

Definition B.2.3 is based on the idea of the *Campbell measure* of Φ, which is the measure

$$\mathcal{C}(B \times L) = \mathrm{E}\left[\Phi(B) I_{\{\Phi \in L\}}\right]$$

defined on $\mathcal{B} \times \mathcal{M}$.

B.3 RANDOM MARKED POINT PROCESSES: RMPPs

Random point processes were originally derived to model the occurrence of events *of the same kind* at random points of $\mathbb{R}$. It often becomes necessary, however, to distinguish between different types of events in a process. In the context of queues or networks, there are service times, priorities, customer types, route-determining vertices and so on, which cause one customer to differ from another. As explained in the following definition, RMPPs allow for such an eventuality.

Definition B.3.1 ((Random) Marked Point Process)

Consider $E, \mathcal{E}$ and $\mathcal{B}$ as before. In addition, let $\mathcal{X}$ be a Polish space, called the *mark space*, and $M_p^{\mathcal{X}}$ be the set of all counting measures on $\mathcal{B} \times \sigma(\mathcal{X})$. An RMPP $\Phi^{\mathcal{X}}$ (with mark space $\mathcal{X}$) is a random element of the measurable space $(M_p^{\mathcal{X}}, \mathcal{M}_p^{\mathcal{X}})$, and thus can be determined as the probability space $(M_p^{\mathcal{X}}, \mathcal{M}_p^{\mathcal{X}}, P^{\mathcal{X}})$.

Remark B.3.1

The RMPP $\Phi^{\mathcal{X}}$ is nothing more than a point process on the space $E \times \mathcal{X}$, so that the discussion in §B.2 is applicable. Consequently theoretical details will not be reproduced (in their appropriately modified forms) here. We do, however, note that random attribute $x \in \mathcal{X}$ is called a *mark*, which may obviously be a vector $\boldsymbol{x} = (x_1, \ldots, x_k)$ where the x_i may, for example, specify service times at the various nodes of a network, or the nodes themselves.

Example B.3.1

A *compound Poisson process* (Parzen [B.5, *pp.* 128–131]) is an example of an RMPP in which marks are IID and independent of the points of the Poisson process they label. The most common marked point process in queuing theory is some version of a compound Poisson process. As an example, consider the $M^{[X]}$-queues, i.e., queues with batch arrivals/service.

Remark B.3.2 (Independence of Marks)

In contrast with the compound Poisson process of Example B.3.1, one may always consider RMPPs for which marks are not necessarily IID, or even those that are not independent of the point process they label.

REFERENCES

[B.1] Franken, P., König, D., Arndt, U. and Schmidt, V. *Queues and Point Processes.* Wiley, Chichester, 1982.

[B.2] Grimmett, G.R. and Stirzaker, D.R. *Probability and random processes.* Oxford University Press, Oxford, 1983.

[B.3] Kallenberg, O. *Random Measures,* 3rd ed. Akademie-Verlag, Berlin, 1983.

[B.4] Medhi, J. *Stochastic Processes.* Wiley Eastern, New Delhi, 1982.

[B.5] Parzen, E. *Stochastic Processes.* Holden-Day, San Francisco, 1962.

APPENDIX C
STOCHASTIC ORDERING

Definition C.1 (Stochastic Ordering of Random Variables and Distributions)

Suppose X and Y are random variables with respective distribution functions F and G.

Then X is *stochastically smaller* (or *smaller in distribution*) than Y, written $X \leq_{st} Y$ (or $X \leq_d Y$), iff

$$\overline{F}(x) \leq \overline{G}(x) \;\; \forall\, x \in \mathbb{R}$$

in which case F is said to be *stochastically smaller* than G, written $F \leq_d G$.

X is *stochastically larger* than Y (F is *stochastically larger* than G), $X \geq_{st} Y$ ($F \geq_d G$), iff Y is stochastically smaller than X.

X is *stochastically equivalent* (*equal*) to Y, $X =_{st} Y$, iff

$$F(x) = G(x) \;\; \forall\, x \in \mathbb{R}$$

in which case it may also be said that F and G are *stochastically equivalent* (*equal*), written as $F =_d G$.

Remark C.1

The symbols $_{st}$ and $_d$ often are used interchangeably. Also, the concept of stochastic ordering may be intuitively extended to sequences of random variables and their joint distributions.

Proposition C.1

Stochastic ordering is a partial ordering on the class of real-valued random variables (or their distributions).

PROOF

Reflexivity, transitivity and antisymmetry are easily shown to hold.

Remark C.2

The ordering is, however, not complete since not all pairs of distributions (random variables) can be stochastically ordered.

Remark C.3

If X, Y are as in Definition C.1, then $F(x) = G(x)\ \forall\, x \in \mathbb{R} \Longrightarrow X =_{st} Y$ but it is not necessarily true that $X = Y$. An obvious and most likely instance would be that in which X and Y are identically distributed.

Theorem C.1

$X \leq_{st} Y$ iff $\mathrm{E}[g(X)] \leq \mathrm{E}[g(Y)]\ \forall$ nondecreasing functions g, whenever the expectations are well-defined.

PROOF

Ross [C.1, *p*. 252] or Wolff [C.2, *p*. 486].

REFERENCES

[C.1] Ross, S.M. *Stochastic Processes.* Wiley, New York, 1983.

[C.2] Wolff, R.W. *Stochastic Modeling and the Theory of Queues.* Prentice-Hall, Englewood Cliffs, New Jersey, 1989.

APPENDIX D
STARLIKE NETWORK COMPUTER SIMULATION PROGRAM

In this appendix, a short explanation is given of some of the aspects of the simulation program mentioned in Chapter 4, where a summary of its output was also given. After a few theoretical details have been provided, the program itself — written by the author in Turbo Pascal Version 6.0 — is reproduced with brief comments.

The simulation program is written for the case of a Poisson arrival stream and exponential service times. The properties of the exponential distribution (especially its Markovian/memoryless character) allow us to model the network using the principles of a Markov chain. This is not to say, however, that more general distribution assumptions could not have been made (using the approximation $F(x) \approx \sum_i p_i F_i(x)$ — where $\sum_i p_i = 1$ and F_i is a gamma distribution function with parameters β_i and n_i — for the distribution, F, under consideration).

One of the essential ideas of the program is certainly fairly general and may be applied to a broad class of simulation studies. We refer to the concept of *discrete event simulation* which is a useful mechanism for avoiding the complexities of dealing with the dynamics of continuous time. Two different approaches exist for discretizing the time.

The first approach, *interval-oriented simulation*, uses *fixed time increments.* Time is advanced by a constant amount Δt, and the situation is reviewed at $\ldots, t-\Delta t, t, t+\Delta t, \ldots$. The method applied in our program is the second type, *event-oriented simulation*, in which time increments vary: the system time is advanced to that of the next event, and the system evaluated at that moment. This latter approach has several advantages over the first which may update the time from t to $t+\Delta t$ without any events having occurred during the interval $(t, t+\Delta t]$. On the other hand, a considerable number of events can occur, but these may be detected only at $t+\Delta t$. Both of these situations are wasteful.

Our preferred variable-time increment method, in accordance with our assumptions on exponential times, yields a useful result on how to establish by random means the time and nature of the events (arrivals/departures/service). In the context of the program, α_n in the theoretical discussion below should be interpreted as the rate of the nth event happening (given the current state of the system). For instance, in the case of circuit-switching, program variable *arrivalrate* gives the rate of arrival to the network which is a constant $N\lambda$, but *departurerate* depends on how many destination channels are nonempty (and on time-ordering if the $FAFS$ protocol is in use). α, on the other hand, represents the (program variable) *totalrate* for any of the protocols.

Let $\{T_n\}$ be a collection of independent random variables (representing interevent times) where T_n has the exponential distribution $F_{T_n}(t) = 1 - e^{-\alpha_n t}$ with mean α_n^{-1}. Write $T_{(1)} = \min_n T_n$ and $\alpha = \sum_n \alpha_n$. Then applying the independence, we have

$$
\begin{aligned}
P\left[T_{(1)} > t\right] &= P\left[\bigcap_n \{T_n > t\}\right] \\
&= \prod_n P\left[T_n > t\right] \\
&= e^{-\sum_n \alpha_n t} \\
&= e^{-\alpha t}
\end{aligned}
$$

$$
\begin{aligned}
P\left[T_k = T_{(1)}\right] &= \int_{s=0}^{\infty} P\left[T_k = s\right] P\left[\bigcap_{n \neq k} \{T_n \geq s\}\right] ds \\
&= \int_{s=0}^{\infty} \alpha_k e^{-\alpha_k s} \prod_{n \neq k} e^{-\alpha_n s} ds \\
&= \alpha_k \int_{s=0}^{\infty} e^{-\alpha s} ds \\
&= \frac{\alpha_k}{\alpha}
\end{aligned}
$$

Also, by again using the independence of the random variables:

$$\begin{aligned}
P\left[T_{(1)} > t, T_k = T_{(1)}\right] &= \int_{u=t}^{\infty} P\left[T_k = u\right] P\left[\bigcap_{n \neq k} \{T_n \geq u\}\right] du \\
&= \int_{u=t}^{\infty} \alpha_k e^{-\alpha_k u} \prod_{n \neq k} e^{-\alpha_n u} du \\
&= \alpha_k \int_{u=t}^{\infty} e^{-\alpha u} du \\
&= \frac{\alpha_k}{\alpha} e^{-\alpha t} \\
&= P\left[T_k = T_{(1)}\right] P\left[T_{(1)} > t\right]
\end{aligned}$$

Consequently the *type* of the next event and the *time interval* until the next event (which is $T_{(1)}$) may be independently determined (by *separately* using the (pseudo)random number generator).

To understand how the random number generator is actually used in this program is not difficult as we apply some well-known properties of distributions.

Let U be a uniform random variate with distribution function G_U:

$$G_U(u) = P[U \leq u] = u \tag{D.1}$$

If T is exponentially distributed with mean α^{-1} and distribution F, then

$$\begin{aligned}
P[T \leq t] &= F(t) \\
&= P[U \leq F(t)] \qquad \text{by (D.1)} \\
&= P[F^{-1}(U) \leq t]
\end{aligned}$$

so solving $T = F^{-1}(U)$ or

$$\begin{aligned} U &= F(T) \\ \Rightarrow \quad U &= 1 - e^{-\alpha T} \\ \Rightarrow \quad e^{-\alpha T} &= 1 - U \\ \Rightarrow \quad T &= -\frac{\ln(1-U)}{\alpha} \end{aligned}$$

and if $U \in [0,1)$ (as is the case with Turbo Pascal's uniform random number generator), then $-\frac{\ln(1-U)}{\alpha}$ yields an (honest) exponential random variable with mean α^{-1}.

The program-listing is given in the pages that follow.

```
PROGRAM Network;  {Interactive version}
{Two-stage starlike network: disrete event-oriented simulation
for different disciplines & switching rules: channel in stage i
(i=1,2) is symmetric w.r.t. others in that stage: exponential
interarrival and service times}

USES CRT, DOS;

CONST
    bigspace                    =    '        ';
    space                       =    '    ';
    underline                   =    '------------';

TYPE
    boolarray                   =    ARRAY[1..250] OF BOOLEAN;
    countarray                  =    ARRAY[1..250] OF LONGINT;
    protocol                    =    (faafs, fafs, fcfs);
    q1pointer                   =    ^q1node;
    q1node                      =    RECORD
           destindex            :  BYTE;
           destrank1            :  LONGINT;
           arrivetime1          :  REAL;
           link1                :  q1pointer
    END;
    q2pointer                   =    ^q2node;
    q2node                      =    RECORD
           sourceindex          :  BYTE;
           destrank2            :  LONGINT;
           arrivetime2          :  REAL;
           link2                :  q2pointer
    END;
    queue1                      =    RECORD
           front1, rear1        :  q1pointer
    END;
    queue1array                 =    ARRAY[1..250] OF queue1;
    queue2                      =    RECORD
           main2, subsid2       :  q2pointer
    END;
    queue2array                 =    ARRAY[1..250] OF queue2;
    servicetype                 =    (nonpre, pre);
                                     {nonpreemptive/preemptive}
    switchtype                  =    (cir, mes);
                                     {circuit/message}
```

```
VAR
    numberofdests, numberofsources,
    numberoftrials, trial          :    BYTE;
    factor, maxmessages            :    LONGINT;
    discipline                     :    protocol;
    arrivalrate,
    intensity, invnumdests,
    lambda, mu1, mu2, musystem,
    overallav,
    overallav1, overallav2,
    overallwait1, overallwait2,
    rho1, rho2, rhosystem,
    wait1, wait2, waitsystem       :    REAL;
    servicerule                    :    servicetype;
    switchrule                     :    switchtype;
    out                            :    TEXT;

{--------------------------------------------------------------}
```

```
PROCEDURE ReadData   {Interactive keyboard response on number of
                      source/destination channels, calls &
                      trials, and rates of arrival & service,
                      and so forth}
(VAR lambda, mu1, mu2: REAL; VAR numberofdests, numberofsources,
numberoftrials: BYTE; VAR maxmessages: LONGINT; VAR discipline:
protocol; VAR servicerule: servicetype; VAR switchrule:
switchtype);
VAR
    selection                    :    SHORTINT;
BEGIN
 CLRSCR;
 WRITELN;
 WRITELN(' ':3,
  '*** WELCOME TO THE STARLIKE NETWORK SIMULATION PROGRAM ***');
 WRITELN; WRITELN; WRITELN;
 WRITELN(' ':5,'DISCIPLINE');
 WRITELN(bigspace, '1. FAAFS');
 WRITELN(bigspace, '2. FAFS');
 WRITELN(bigspace, '3. FCFS');
 WRITE(' ':5,'Enter selection:  ');
 READLN(selection);
 CASE selection OF
  1: discipline:=faafs;
  2: discipline:=fafs;
  3: discipline:=fcfs
 END;
 IF discipline=faafs
  THEN
   BEGIN
    WRITELN; WRITELN;
    WRITELN(' ':5,'SERVICE RULE');
    WRITELN(bigspace, '1. Nonpreemptive');
    WRITELN(bigspace, '2. Preemptive');
    WRITE(' ':5,'Enter selection:  ');
    READLN(selection);
    CASE selection OF
     1: servicerule:=nonpre;
     2: servicerule:=pre
    END
   END
  ELSE servicerule:=nonpre;
 WRITELN; WRITELN;
 WRITELN(' ':5,'SWITCHING RULE');
```

```
WRITELN(bigspace, '1. Circuit');
WRITELN(bigspace, '2. Message');
WRITE(' ':5,'Enter selection:  ');
READLN(selection);
CASE selection OF
 1: switchrule:=cir;
 2: switchrule:=mes;
END;
WRITELN; WRITELN;
WRITE('Please enter the mean arrival rate LAMBDA:',' ':23);
READLN(lambda);
CASE switchrule OF
 cir: mu1:=0;
 mes: BEGIN
       WRITE('Please enter the stage 1 mean service rate MU1:',
             ' ':18);
       READLN(mu1)
      END
END;
WRITE('Please enter the stage 2 mean service rate MU2:',' ':18);
READLN(mu2);
WRITE('Please enter the number of source queues<251:',' ':20);
READLN(numberofsources);
WRITE('Please enter the number of destination queues<251:',
      ' ':15);
READLN(numberofdests);
WRITE('Please enter the number of calls/messages to be',
      ' processed<2E9:   ');
READLN(maxmessages);
WRITE('Please enter the number of trials<251:',' ':27);
READLN(numberoftrials);
WRITELN('PROCESSING BEGINS');
END;   {ReadData}
```

```
PROCEDURE InitialiseAllTrials  {Set to zero wait statistics that
                                 will accumulate over all trials;
                                 assign values to quantities that
                                 remain constant over all trials}
(numberofdests, numberofsources: BYTE; maxmessages: LONGINT; VAR
factor: LONGINT; lambda, mu1, mu2: REAL; VAR arrivalrate,
intensity, invnumdests, musystem, overallwait1, overallwait2,
rho1, rho2, rhosystem: REAL; switchrule: switchtype);
BEGIN
 factor:=TRUNC(maxmessages/10);
 WHILE factor>5000 DO
  factor:=TRUNC(factor/2);
 arrivalrate:=lambda*numberofsources;
 invnumdests:=1/numberofdests;
 intensity:=arrivalrate*invnumdests;
 overallwait1:=0;
 overallwait2:=0;
 CASE switchrule OF
  cir: BEGIN
        rho2:=intensity/mu2;
        musystem:=mu2-intensity;
        rhosystem:=lambda/musystem
       END;
  mes: BEGIN
        rho1:=lambda/mu1;
        rho2:=intensity/mu2
       END
 END
END;    {InitialiseAllTrials}

FUNCTION TrueMean  {Calculate theoretical values of mean waiting
                    times for FCFS discipline}
(mu, rho: REAL):REAL;
BEGIN
    IF (rho<1)
     THEN TrueMean:=1/(mu*(1-rho))
     ELSE TrueMean:=9999
END;   {TrueMean}
```

```
PROCEDURE PrintData  {Data, based on input & subsequent
                      calculations, pertinent to all trials is
                      written to a text file, to which
                      information will be appended}
(numberofdests, numberofsources: BYTE; maxmessages: LONGINT;
discipline: protocol; intensity, lambda, mu1, mu2, musystem,
rho1, rho2, rhosystem: REAL; VAR wait1, wait2: REAL; servicerule:
servicetype; switchrule: switchtype; VAR out: TEXT);
BEGIN
 WRITELN(out); WRITELN(out);
 WRITE(out,'           ****      ');
 CASE discipline OF
  faafs: CASE servicerule OF
          nonpre: WRITE(out, 'NONPREEMPTIVE FAAFS');
          pre:    WRITE(out, '  PREEMPTIVE FAAFS ')
         END;
  fafs:  WRITE(out, '       FAFS       ');
  fcfs:  WRITE(out, '       FCFS       ')
 END;
 CASE switchrule OF
  cir: WRITE(out, ' CIRCUIT-SWITCHING');
  mes: WRITE(out, ' MESSAGE-SWITCHING')
 END;
 WRITELN(out,'      ****');
 WRITELN(out); WRITELN(out);
 WRITELN(out,'Number of arrivals to network:', bigspace,
           maxmessages:7, bigspace);
 WRITELN(out); WRITELN(out); WRITELN(out);
 WRITELN(out,' ':43, 'Stage',' ':16,'Total', ' ':11);
 WRITELN(out,' ':38,'1',' ':13,'2');
 WRITELN(out); WRITELN(out);
 WRITELN(out, 'Number of channels',' ':19, numberofsources:3,
        ' ':11, numberofdests:3,' ':7);
 WRITELN(out);
 CASE switchrule OF
  cir: BEGIN
        WRITELN(out, 'Mean arrival rate (per channel)',
                lambda:13:6, ' ', intensity:13:6,' ':7);
        WRITELN(out);
        WRITE(out, 'Mean service rate MU',' ':25,mu2:13:6);
        IF NOT(discipline=fafs) THEN
         WRITELN(out, musystem:14:6)
         ELSE WRITELN(out);
        WRITELN(out);
```

```
        WRITE(out,'Traffic intensity (per channel)',' ':14,
              rho2:13:6);
        IF NOT(discipline=fafs) THEN
         WRITELN(out, rhosystem:14:6)
         ELSE WRITELN(out);
        WRITELN(out);
        wait2:=TrueMean(mu2, rho2);
        waitsystem:=TrueMean(musystem, rhosystem);
        wait1:=waitsystem-wait2;
        WRITELN(out,'Theoretical waiting time', bigspace,
                wait1:13:6,' ', wait2:13:6, waitsystem:14:6);
        WRITELN(out,'(FCFS circuit-switching)')
       END;
 mes: BEGIN
        WRITELN(out,'Mean arrival rate (per channel)',
                lambda:13:6, ' ', intensity:13:6, ' ':7);
        WRITELN(out);
        WRITELN(out,'Mean service rate (per channel)', mu1:13:6,
                ' ', mu2:13:6, ' ':7);
        WRITELN(out);
        WRITELN(out,'Traffic intensity (per channel)', rho1:13:6,
                ' ', rho2:13:6, ' ':7);
        WRITELN(out);
        wait1:=TrueMean(mu1, rho1);
        wait2:=TrueMean(mu2, rho2);
        WRITELN(out,'Theoretical waiting time', bigspace,
                wait1:13:6, ' ', wait2:13:6, (wait1+wait2):14:6);
        WRITELN(out,'(FCFS message-switching)')
       END
 END;
 WRITELN(out); WRITELN(out); WRITELN(out);
 WRITELN(out,' ':5,'  ','No. leaving',' ','No. leaving','  ',
        'Stage 1 mean ',' ','Stage 2 mean ',' ','Network mean',
        ' ':8);
 WRITELN(out,'Trial','  ','-----------',' ','-----------','  ',
        underline,'  ',underline, '  ',underline);
 WRITELN(out,'-----','  ','  Stage 1  ',' ','  Stage 2  ','  ',
        'waiting time ',' ','waiting time ',' ','waiting time',
        ' ':8);
 WRITELN(out,' ':5,'  ','  -------  ',' ','  -------  ','  ',
        underline,'  ',underline, '  ',underline);
 WRITELN(out);
 CLOSE(out)
END;  {PrintData}
```

```
PROCEDURE InitialiseATrial  {Assign initial values (for a single
                             trial) to wait, count, boolean and
                             other quantities}
(VAR canleave, windowempty: boolarray; numberofdests,
numberofsources: BYTE; VAR nonemptydests, nonemptysources: BYTE;
VAR nexttoleave1, nexttoleave2, numarray1, numarray2,
orderarray1, orderarray2: countarray; VAR arrivalcount,
departurecount, interstagecount: LONGINT; discipline: protocol;
VAR stage1array: queue1array; VAR stage2array: queue2array; VAR
currenttime, totalwait1, totalwait2: REAL; servicerule:
servicetype; switchrule: switchtype);
VAR
    counter1                      :    BYTE;
BEGIN
 FOR counter1:=1 TO numberofdests DO
  canleave[counter1]:=FALSE;
 IF (discipline=faafs) AND (servicerule=nonpre)
  THEN
   BEGIN
    FOR counter1:=1 TO numberofdests DO
     windowempty[counter1]:=TRUE
   END;
 nonemptydests:=0;
 nonemptysources:=0;
 FOR counter1:=1 TO numberofsources DO
  BEGIN
   IF switchrule=cir
    THEN
     BEGIN
      nexttoleave1[counter1]:=1;
      orderarray1[counter1]:=1
     END;
   numarray1[counter1]:=0;
  END;
 FOR counter1:=1 TO numberofdests DO
  BEGIN
   nexttoleave2[counter1]:=1;
   numarray2[counter1]:=0;
   orderarray2[counter1]:=1
  END;
 arrivalcount:=0;
 departurecount:=0;
 interstagecount:=0;
```

```
 FOR counter1:=1 TO numberofsources DO
  BEGIN
   stage1array[counter1].front1:=NIL;
   stage1array[counter1].rear1:=NIL
  END;
 FOR counter1:=1 TO numberofdests DO
  BEGIN
   stage2array[counter1].main2:=NIL;
   stage2array[counter1].subsid2:=NIL
  END;
 currenttime:=0;
 totalwait1:=0;
 totalwait2:=0;
END;   {InitialiseATrial}

{---------------------------------------------------------------}
PROCEDURE InterEvent  {Use the uniform random number generator to
                       produce an exponential random variable
                       which gives the time between successive
                       events; the current time is updated to
                       that of the new event}
(nonemptydests, nonemptysources: BYTE; arrivalrate, mu1, mu2:
REAL; VAR currenttime, departurerate, servicerate, totalrate:
REAL; switchrule: switchtype);
VAR
    exponvariate,
    uniformvariate              :    REAL;
BEGIN
 servicerate:=nonemptysources*mu1;
 departurerate:=nonemptydests*mu2;
 CASE switchrule OF
  cir: totalrate:=arrivalrate+departurerate;
  mes: totalrate:=arrivalrate+servicerate+departurerate
 END;
 uniformvariate:=RANDOM;
 exponvariate:=-LN(1-uniformvariate)/totalrate;
 currenttime:=currenttime+exponvariate
END;   {InterEvent}
```

```
PROCEDURE EventType  {Use the random number generator again, this
                      time to determine the nature of the next
                      event (arrival to, or departure from, the
                      system and, in the case of
                      message-switching, possibly completion of
                      first-stage service)}
(VAR arrival, departure, service: BOOLEAN; arrivalrate,
servicerate, totalrate: REAL; switchrule: switchtype);
VAR
    arrivalprob, randomprob,
    serviceprob                   :    REAL;
BEGIN
  randomprob:=RANDOM;
  arrival:=FALSE;
  departure:=FALSE;
  service:=FALSE;
  arrivalprob:=arrivalrate/totalrate;
  CASE switchrule OF
   cir: IF randomprob<arrivalprob
         THEN arrival:=TRUE
         ELSE departure:=TRUE;
   mes: BEGIN
         serviceprob:=servicerate/totalrate;
         IF randomprob<arrivalprob
          THEN arrival:=TRUE
          ELSE IF (arrivalprob<=randomprob) AND
                  (randomprob<(arrivalprob+serviceprob))
           THEN service:=TRUE
           ELSE departure:=TRUE
        END
  END
END;   {EventType}
```

```
PROCEDURE EventIndex  {Marks are assigned to, or extracted from,
                       the call (or message)}
(canleave: boolarray; arrival, departure, service: BOOLEAN;
nonemptydests, nonemptysources, numberofdests, numberofsources:
BYTE; VAR destination, source: BYTE; numarray1, numarray2:
countarray; discipline: protocol; switchrule: switchtype);
VAR
    counter, destposition,
    sourceposition                :    BYTE;
BEGIN
 IF arrival
  THEN
   BEGIN                                      {source & destin.}
    source:=SUCC(RANDOM(numberofsources));   {determined randomly}
    destination:=SUCC(RANDOM(numberofdests)) {for an arriving}
   END                                        {call}
  ELSE IF departure
   THEN
    BEGIN                                       {nonempty queue}
     destposition:=SUCC(RANDOM(nonemptydests)); {from which call}
     counter:=1;                                {will leave the}
     destination:=0;                            {network is}
     WHILE counter<=destposition DO             {randomly establ.}
      BEGIN
       destination:=SUCC(destination);
       CASE discipline OF
        faafs, fcfs: IF numarray2[destination]>0
                      THEN counter:=SUCC(counter);
        fafs       : IF (numarray2[destination]>0)
                        AND canleave[destination]
                      THEN counter:=SUCC(counter)
       END;
      END;
    END;
  IF (switchrule=mes) AND service
   THEN
    BEGIN                                           {in case of}
     sourceposition:=SUCC(RANDOM(nonemptysources)); {mes-switch:}
     counter:=1;                                    {randomly}
     source:=0;                                     {find in}
```

```
    WHILE counter<=sourceposition DO                    {which non-}
     BEGIN                                              {empty first-}
      source:=SUCC(source);                             {stage queue}
      IF numarray1[source]>0                            {service}
       THEN counter:=SUCC(counter);                     {completion}
     END                                                {occurs}
   END;
END;   {EventIndex}

{---------------------------------------------------------------}

PROCEDURE FindPosition  {locate destination-queue position, which
                         is based on arrival times to the
                         network}
(destination: BYTE; temppointer2: q2pointer; VAR next, previous:
q2pointer; stage2array: queue2array);
BEGIN
 previous:=NIL;
 next:=stage2array[destination].main2;
 WHILE (temppointer2^.destrank2>next^.destrank2)
       AND (next<>NIL) DO
  BEGIN
   previous:=next;
   next:=next^.link2
  END
END;   {FindPosition}

PROCEDURE PutInPosition  {join queue at the appropriate (time-
                          dictated) position}
(destination: BYTE; next: q2pointer; VAR previous, temppointer2:
q2pointer; VAR stage2array: queue2array);
BEGIN
 IF previous=NIL
  THEN
   BEGIN
    temppointer2^.link2:=stage2array[destination].main2;
    stage2array[destination].main2:=temppointer2
   END
  ELSE
   BEGIN
    previous^.link2:=temppointer2;
    temppointer2^.link2:=next
   END
END;   {PutInPosition}
```

```
PROCEDURE PutInLine  {combines previous two procedures}
(destination: BYTE; VAR next, previous, temppointer2: q2pointer;
VAR stage2array: queue2array);
BEGIN
 FindPosition(destination, temppointer2, next, previous,
              stage2array);
 PutInPosition(destination, next, previous, temppointer2,
               stage2array);
END;   {PutInLine}

PROCEDURE PutInWindow  {for nonpreemptive FAAFS: call enters
                        service 'window' since queue is empty;
                        statistics are adjusted}
(VAR windowempty: boolarray; destination: BYTE; VAR
nonemptydests: BYTE; VAR temppointer2: q2pointer; VAR
stage2array: queue2array);
BEGIN
 windowempty[destination]:=FALSE;                {subsid2 is}
 nonemptydests:=SUCC(nonemptydests);             {service 'window'}
 temppointer2^.link2:=NIL;
 stage2array[destination].subsid2:=temppointer2
END;   {PutInWindow}
```

```
PROCEDURE RemoveFromLine  {if not the nonpreemptive FAAFS
                           discipline, leave from front of queue
                           'proper'}
(destination: BYTE; VAR source: BYTE; numarray2: countarray;
discipline: protocol; VAR stage2array: queue2array; VAR
timeofarrival2: REAL);
VAR
    temppointer3                    :    q2pointer;
BEGIN
 temppointer3:=stage2array[destination].main2;
 timeofarrival2:=temppointer3^.arrivetime2;
 source:=temppointer3^.sourceindex;
 IF discipline=fcfs
  THEN
   BEGIN
    CASE numarray2[destination] OF
     1:   BEGIN
           stage2array[destination].main2:=NIL;
           stage2array[destination].subsid2:=NIL
          END;
     2:   stage2array[destination].main2:=
          stage2array[destination].subsid2;
     ELSE stage2array[destination].main2:=
          stage2array[destination].main2^.link2
    END
   END
  ELSE stage2array[destination].main2:=
       stage2array[destination].main2^.link2;
 DISPOSE(temppointer3)
END;   {RemoveFromLine}
```

```
PROCEDURE RemoveFromWindow  {otherwise, for the nonpreemptive
                             FAAFS discipline, call leaves
                             service window & head of queue
                             proceeds to service window}
(VAR windowempty: boolarray; destination: BYTE; VAR source: BYTE;
numarray2: countarray; VAR stage2array: queue2array; VAR
timeofarrival2: REAL);
VAR
    temppointer3                 :    q2pointer;
BEGIN
 temppointer3:=stage2array[destination].subsid2; {subsid2 is}
 timeofarrival2:=temppointer3^.arrivetime2;      {service window}
 source:=temppointer3^.sourceindex;
 CASE numarray2[destination] OF
  1:   BEGIN
        stage2array[destination].subsid2:=NIL;
        windowempty[destination]:=TRUE
       END
  ELSE BEGIN
        stage2array[destination].subsid2:=
        stage2array[destination].main2;
        stage2array[destination].main2:=
        stage2array[destination].main2^.link2
       END
 END;
 DISPOSE(temppointer3)
END;   {RemoveFromWindow}
```

```
PROCEDURE CheckWindow  {If, in the nonpreemptive FAAFS situation,
                        the destinations of the departing call
                        & head of the associated source-queue
                        (also) coincide, then the latter customer
                        may have arrived prior to the head of the
                        destination queue. Thus that 'front call'
                        must have its position in the service
                        window replaced before advancing the time}
(destination: BYTE; numarray2: countarray; VAR stage2array:
queue2array);
VAR
    temppointer4            :    q2pointer;
BEGIN
 IF numarray2[destination]>1
  THEN
   BEGIN
    IF stage2array[destination].subsid2^.destrank2>
       stage2array[destination].main2^.destrank2
     THEN
      BEGIN
       temppointer4:=stage2array[destination].subsid2;
       stage2array[destination].subsid2:=
       stage2array[destination].main2;
       temppointer4^.link2:=stage2array[destination].main2^.link2;
       stage2array[destination].main2:=temppointer4;
       stage2array[destination].subsid2^.link2:=NIL
      END
   END
END;   {CheckWindow}

{---------------------------------------------------------------}

PROCEDURE NPFAAFSArr2  {arrival to stage 2: nonpreemptive FAAFS}
(VAR windowempty: boolarray; destination: BYTE; VAR
nonemptydests: BYTE; VAR next, previous, temppointer2: q2pointer;
VAR stage2array: queue2array);
BEGIN
 IF windowempty[destination]
  THEN PutInWindow(windowempty, destination, nonemptydests,
                   temppointer2, stage2array)
  ELSE PutInLine(destination, next, previous, temppointer2,
                 stage2array)
END;   {NPFAAFSArr2}
```

```
PROCEDURE PFAAFSArr2  {arrival to stage 2: preemptive FAAFS}
(destination: BYTE; VAR nonemptydests: BYTE; numarray2:
countarray; VAR next, previous, temppointer2: q2pointer; VAR
stage2array: queue2array);
BEGIN
 PutInLine(destination, next, previous, temppointer2,
           stage2array);
                                         {incr. from 0 to 1 in}
 IF numarray2[destination]=0             {destination queue}
                                         {means an increase in}
  THEN nonemptydests:=SUCC(nonemptydests) {the number of nonempty}
                                         {destination queues}
END;   {PFAAFSArr2}

PROCEDURE FAFSArr2  {arrival to stage 2: FAFS}
(VAR canleave: boolarray; destination: BYTE; VAR nonemptydests:
BYTE; nexttoleave2: countarray; VAR next, previous, temppointer2:
q2pointer; VAR stage2array: queue2array);
BEGIN
 PutInLine(destination, next, previous, temppointer2, stage2array);
 IF previous=NIL   {If customer joins front of FAFS queue, then}
  THEN             {either the queue is empty or customer has the}
   BEGIN           {earliest arrival of those in the queue}
    IF stage2array[destination].main2^.destrank2=
       nexttoleave2[destination]          {Possibly then call may}
     THEN                                 {leave next. If so,}
      BEGIN                               {number of 'nonempty'}
       nonemptydests:=SUCC(nonemptydests); {destination queues}
       canleave[destination]:=TRUE        {increases by 1}
      END
   END
END;   {FAFSArr2}
```

```
PROCEDURE FCFSArr2  {arrival to stage 2: FCFS - simply join back
                     of queue}
(destination: BYTE; VAR nonemptydests: BYTE; numarray2:
countarray; VAR temppointer2: q2pointer; VAR stage2array:
queue2array);
BEGIN
 temppointer2^.link2:=NIL;
 IF numarray2[destination]=0
  THEN stage2array[destination].main2:=temppointer2
  ELSE stage2array[destination].subsid2^.link2:=
       temppointer2;
 stage2array[destination].subsid2:=temppointer2;
                                         {incr. from 0 to 1 in}
 IF numarray2[destination]=0             {destination queue}
                                         {means an increase in}
  THEN nonemptydests:=SUCC(nonemptydests) {the number of nonempty}
                                         {destination queues}
END;   {FCFSArr2}

{-----------------------------------------------------------------}

PROCEDURE Arrival1  {Call joins back of appropriate source queue}
(destination, source: BYTE; VAR nonemptysources: BYTE;
orderarray2: countarray; VAR numarray13: LONGINT;
VAR stage1array3: queue1; currenttime: REAL);
VAR
    temppointer1                  :    q1pointer;
BEGIN
 NEW(temppointer1);
 WITH temppointer1^ DO
  BEGIN
   destindex:=destination;
   destrank1:=orderarray2[destination];
   arrivetime1:=currenttime;
   link1:=NIL
  END;
 IF numarray13=0
  THEN stage1array3.front1:=temppointer1
  ELSE stage1array3.rear1^.link1:=temppointer1;
  stage1array3.rear1:=temppointer1;
  numarray13:=SUCC(numarray13);
  IF numarray13=1
   THEN nonemptysources:=SUCC(nonemptysources);
END;   {Arrival1}
```

```
PROCEDURE Removal1  {Customer leaves front of appropriate source
                     queue}
(source: BYTE; VAR destination, nonemptysources: BYTE;
VAR numarray15: LONGINT; VAR temppointer2: q2pointer;
VAR stage1array5: queue1; currenttime: REAL;
VAR timeofarrival1: REAL);
VAR
    temppointer                  :    q1pointer;
BEGIN
 NEW(temppointer2);
 temppointer:=stage1array5.front1;
 timeofarrival1:=temppointer^.arrivetime1;
 destination:=temppointer^.destindex;
 temppointer2^.sourceindex:=source;
 WITH temppointer^ DO
  BEGIN
  temppointer2^.destrank2:=destrank1;    {obtain update on marks}
  temppointer2^.arrivetime2:=currenttime {& statistics of call}
  END;
 DISPOSE(temppointer);
 WITH stage1array5 DO
  BEGIN
   CASE numarray15 OF
    1:   BEGIN
          front1:=NIL;
          rear1:=NIL
         END;
    2:   front1:=rear1
    ELSE front1:=front1^.link1
   END
  END;
 numarray15:=PRED(numarray15);
                                               {decr. from 1 to 0}
 IF numarray15=0                               {in source queue}
                                               {means a decr. in}
  THEN nonemptysources:=PRED(nonemptysources); {the number of}
                                               {nonempty sources}

END;   {Removal1}
```

```
PROCEDURE Arrival2  {customer joins appropriate destination queue
                     in a position determined by the protocol in
                     use}
(VAR canleave, windowempty: boolarray; destination: BYTE; VAR
nonemptydests: BYTE; nexttoleave2: countarray; VAR numarray2:
countarray; VAR interstagecount: LONGINT; discipline: protocol;
VAR temppointer2: q2pointer; VAR stage2array: queue2array;
servicerule: servicetype);
VAR
    previous, next                :    q2pointer;
BEGIN
 CASE discipline OF
  faafs: CASE servicerule OF
          nonpre: NPFAAFSArr2(windowempty, destination,
                              nonemptydests, next, previous,
                              temppointer2, stage2array);
          pre:    PFAAFSArr2(destination, nonemptydests,
                             numarray2, next, previous,
                             temppointer2, stage2array)
         END;

  fafs:  FAFSArr2(canleave, destination, nonemptydests,
                  nexttoleave2, next, previous, temppointer2,
                  stage2array);
  fcfs:  FCFSArr2(destination, nonemptydests, numarray2,
                  temppointer2, stage2array)
 END;
 numarray2[destination]:=SUCC(numarray2[destination]); {update}
 interstagecount:=SUCC(interstagecount);               {stats}
END;   {Arrival2}

{----------------------------------------------------------------}
```

```
PROCEDURE CircuitArrival1  {Handles arrivals to circuit-switched
                            network: since there is no 1st-stage
                            service, call may be able to proceed
                            to stage 2 immediately}
(VAR canleave, windowempty: boolarray; destination, source: BYTE;
VAR nonemptydests, nonemptysources: BYTE; nexttoleave12,
orderarray12: LONGINT; nexttoleave2, orderarray2: countarray; VAR
numarray12: LONGINT; VAR numarray2: countarray; VAR
interstagecount: LONGINT; discipline: protocol; VAR temppointer2:
q2pointer; VAR stage1array2: queue1; VAR stage2array:
queue2array; currenttime: REAL; servicerule: servicetype);
BEGIN
 IF orderarray12=nexttoleave12 {new arrival may proceed directly}
  THEN                         {to second stage}
   BEGIN
    NEW(temppointer2);
    WITH temppointer2^ DO
     BEGIN
      sourceindex:=source;
      destrank2:=orderarray2[destination];
      arrivetime2:=currenttime
     END;
    Arrival2(canleave, windowempty, destination, nonemptydests,
             nexttoleave2, numarray2, interstagecount, discipline,
             temppointer2, stage2array, servicerule)
   END
  ELSE Arrival1(destination, source, nonemptysources,
                orderarray2, numarray12, stage1array2, currenttime)
    {otherwise join stage 1 queue}
END;   {CircuitArrival1}

{-----------------------------------------------------------------}
```

```
PROCEDURE EnterSystem  {Controls arrivals to the first stage
                        (hence the network)}
(VAR canleave, windowempty: boolarray; destination, source: BYTE;
VAR nonemptydests, nonemptysources: BYTE; nexttoleave11: LONGINT;
nexttoleave2: countarray; VAR numarray11: LONGINT; VAR
orderarray11: LONGINT; VAR numarray2, orderarray2: countarray;
factor: LONGINT; VAR arrivalcount, interstagecount: LONGINT;
discipline: protocol; VAR temppointer2: q2pointer; VAR
stage1array1: queue1; VAR stage2array: queue2array; currenttime:
REAL; servicerule: servicetype; switchrule: switchtype);
BEGIN
 CASE switchrule OF  {enter network}
  cir: CircuitArrival1(canleave, windowempty, destination,
                       source, nonemptydests, nonemptysources,
                       nexttoleave11, orderarray11, nexttoleave2,
                       orderarray2, numarray11, numarray2,
                       interstagecount, discipline, temppointer2,
                       stage1array1, stage2array, currenttime,
                       servicerule);
  mes: Arrival1(destination, source, nonemptysources,
                orderarray2, numarray11, stage1array1,
                currenttime)
 END;
 arrivalcount:=SUCC(arrivalcount);  {update statistics on}
 IF (arrivalcount MOD factor)=0     {counting ...}
  THEN WRITELN(arrivalcount:8, ' calls/messages arrived to the',
               ' network.');
 orderarray11:=SUCC(orderarray11);  {... and 'order'}
 orderarray2[destination]:=SUCC(orderarray2[destination]);
END;   {EnterSystem}
```

```
PROCEDURE Stage1ToStage2  {Controls transfer of customers from
                           first to second stage in the network
                           for either of the switching rules}
(VAR canleave, windowempty: boolarray; source: BYTE; VAR
destination: BYTE; VAR nonemptydests, nonemptysources: BYTE;
nexttoleave2: countarray; VAR numarray14: LONGINT; VAR numarray2:
countarray; VAR interstagecount: LONGINT; discipline: protocol;
VAR temppointer2: q2pointer; VAR stage1array4: queue1; VAR
stage2array: queue2array; currenttime: REAL; VAR totalwait1:
REAL; servicerule: servicetype);
VAR
    stage1wait, timeofarrival1    :    REAL;
BEGIN
 Removal1(source, destination, nonemptysources, numarray14,
          temppointer2, stage1array4, currenttime,
          timeofarrival1); {leave stage 1}
 Arrival2(canleave, windowempty, destination, nonemptydests,
          nexttoleave2, numarray2, interstagecount, discipline,
          temppointer2, stage2array, servicerule); {enter sta. 2}
 stage1wait:=currenttime-timeofarrival1;
 totalwait1:=totalwait1+stage1wait
END;   {Stage1ToStage2}

PROCEDURE LeaveSystem  {Controls departure from the second stage
                        (hence the network) for all protocols}
(VAR canleave, windowempty: boolarray; destination: BYTE; VAR
nonemptydests, nonemptysources, olddestination, source: BYTE; VAR
nexttoleave1, nexttoleave2, numarray1, numarray2: countarray; VAR
departurecount, interstagecount: LONGINT; discipline: protocol;
VAR temppointer2: q2pointer; VAR stage1array: queue1array; VAR
stage2array: queue2array; currenttime: REAL; VAR totalwait1,
totalwait2: REAL; servicerule: servicetype; switchrule:
switchtype);
VAR
    stage2wait, timeofarrival2    :    REAL;
BEGIN
 IF (discipline=faafs) AND (servicerule=nonpre)
   {distinction between nonpreemptive & preemptive service, which
    matters in FAAFS}
  THEN RemoveFromWindow(windowempty, destination, source,
                        numarray2, stage2array, timeofarrival2)
  ELSE RemoveFromLine(destination, source, numarray2, discipline,
                      stage2array, timeofarrival2);
```

```
stage2wait:=currenttime-timeofarrival2; {update stats: wait/}
totalwait2:=totalwait2+stage2wait;      {numbers in queues/}
IF switchrule=cir                       {orders & priorities}
 THEN nexttoleave1[source]:=SUCC(nexttoleave1[source]);
nexttoleave2[destination]:=SUCC(nexttoleave2[destination]);
numarray2[destination]:=PRED(numarray2[destination]);
departurecount:=SUCC(departurecount);
CASE discipline OF
    {decrease in number of 'nonempty' queues is dictated, inter
     alia, by the discipline}
 faafs,
 fcfs  : IF numarray2[destination]=0
          THEN nonemptydests:=PRED(nonemptydests);
 fafs:   BEGIN
          IF (numarray2[destination]=0) OR
             (canleave[destination] AND
             (stage2array[destination].main2^.destrank2<>
              nexttoleave2[destination]))
           THEN
            BEGIN
             canleave[destination]:=FALSE;
             nonemptydests:=PRED(nonemptydests)
            END
         END
 END;
 IF switchrule=cir
  THEN
   BEGIN
    IF numarray1[source]>=1
     THEN
      BEGIN {next customer with same source may leave stage 1}
       olddestination:=destination;
       Stage1ToStage2(canleave, windowempty, source, destination,
                      nonemptydests, nonemptysources,
                      nexttoleave2, numarray1[source], numarray2,
                      interstagecount, discipline, temppointer2,
                      stage1array[source], stage2array,
                      currenttime, totalwait1, servicerule);
       IF (destination=olddestination) AND (discipline=faafs) AND
          (servicerule=nonpre) {coincidence of destinations}
        THEN CheckWindow(destination, numarray2, stage2array)
      END
    END;
END;   {LeaveSystem}
```

```
PROCEDURE PrintTrialData  {Output analysis for a specific trial}
(trial: BYTE; departurecount, interstagecount : LONGINT;
averagewait1, averagewait2: REAL; VAR out:TEXT);
BEGIN
 APPEND(out);
 WRITELN(out, trial:3,'. ',' ',interstagecount:11,' ',
        departurecount:11, averagewait1:14:6, averagewait2:14:6,
        (averagewait1+averagewait2):14:6);
 CLOSE(out)
END;   {PrintTrialData}

PROCEDURE TrialStatistics  {Data analysis: calculations & output
                            for a single trial}
(trial: BYTE; departurecount, interstagecount: LONGINT;
totalwait1, totalwait2: REAL; VAR overallwait1, overallwait2:
REAL; VAR out: TEXT);
VAR
    averagewait1, averagewait2    :    REAL;
BEGIN
 IF interstagecount>0
  THEN averagewait1:=totalwait1/interstagecount
  ELSE averagewait1:=0;
 IF departurecount>0
  THEN averagewait2:=totalwait2/departurecount
  ELSE averagewait2:=0;
 PrintTrialData(trial, departurecount, interstagecount,
               averagewait1, averagewait2, out);
 overallwait1:=overallwait1+averagewait1;
 overallwait2:=overallwait2+averagewait2
END;   {TrialStatistics}

{----------------------------------------------------------------}
```

```
PROCEDURE EachTrial  {Controls the operation of a complete trial}
(numberofdests, numberofsources, trial: BYTE; factor,
maxmessages: LONGINT; discipline: protocol; arrivalrate, mu1,
mu2: REAL; VAR overallwait1, overallwait2: REAL; servicerule:
servicetype; switchrule: switchtype; VAR out: TEXT);
VAR
    canleave, windowempty          :    boolarray;
    arrival, departure, service    :    BOOLEAN;
    destination, nonemptydests,
    nonemptysources,
    olddestination, source         :    BYTE;
    nexttoleave1, nexttoleave2,
    numarray1, numarray2,
    orderarray1, orderarray2       :    countarray;
    arrivalcount, departurecount,
    interstagecount                :    LONGINT;
    temppointer2                   :    q2pointer;
    stage1array                    :    queue1array;
    stage2array                    :    queue2array;
    currenttime, departurerate,
    finish, servicerate,
    start, totalrate,
    totalwait1, totalwait2         :    REAL;
    hr, hundth, min, sec           :    WORD;

BEGIN
 WRITELN;
 WRITELN('Trial number',trial:4, ' initiated.');
 GetTime(hr, min, sec, hundth);
 start:=3600*hr+60*min+sec+1/100*hundth;
 InitialiseATrial(canleave, windowempty, numberofdests,
                  numberofsources, nonemptydests,
                  nonemptysources, nexttoleave1, nexttoleave2,
                  numarray1, numarray2, orderarray1, orderarray2,
                  arrivalcount, departurecount, interstagecount,
                  discipline, stage1array, stage2array,
                  currenttime, totalwait1, totalwait2,
                  servicerule, switchrule);
 WHILE arrivalcount<maxmessages DO {stopping condition within}
                                   {a trial}
  BEGIN
   InterEvent(nonemptydests, nonemptysources, arrivalrate, mu1,
              mu2, currenttime, departurerate, servicerate,
              totalrate, switchrule);
```

```
  EventType(arrival, departure, service, arrivalrate,
            servicerate, totalrate, switchrule);
  EventIndex(canleave, arrival, departure, service,
             nonemptydests, nonemptysources, numberofdests,
             numberofsources, destination, source, numarray1,
             numarray2, discipline, switchrule);
  IF arrival
   THEN EnterSystem(canleave, windowempty, destination, source,
                    nonemptydests, nonemptysources,
                    nexttoleave1[source], nexttoleave2,
                    numarray1[source], orderarray1[source],
                    numarray2, orderarray2, factor,
                    arrivalcount, interstagecount, discipline,
                    temppointer2, stage1array[source],
                    stage2array, currenttime, servicerule,
                    switchrule)
   ELSE IF departure
    THEN LeaveSystem(canleave, windowempty, destination,
                     nonemptydests, nonemptysources,
                     olddestination, source, nexttoleave1,
                     nexttoleave2, numarray1, numarray2,
                     departurecount, interstagecount,
                     discipline, temppointer2, stage1array,
                     stage2array, currenttime, totalwait1,
                     totalwait2, servicerule, switchrule)
   ELSE IF service AND (switchrule=mes)
    THEN Stage1ToStage2(canleave, windowempty, source,
                        destination, nonemptydests,
                        nonemptysources, nexttoleave2,
                        numarray1[source], numarray2,
                        interstagecount, discipline,
                        temppointer2, stage1array[source],
                        stage2array, currenttime, totalwait1,
                        servicerule)
  END;
 TrialStatistics(trial, departurecount, interstagecount,
                 totalwait1, totalwait2, overallwait1,
                 overallwait2, out);
 GetTime(hr, min, sec, hundth);
 finish:=3600*hr+60*min+sec+1/100*hundth;
 WRITE('Trial number',trial:4, ' completed.');
 WRITELN('   Duration of trial: ',(finish-start):15:8,
        ' seconds.');
END;   {EachTrial}
```

```
BEGIN {program}  {Input of data -- generate and analyse system
                  model -- output analysis}
 ASSIGN(out,'A:\Network.out');
 APPEND(out);
 RANDOMIZE;
 WRITELN(out,'-----------------------------------------------',
         '------------------------------');
 WRITELN(out);WRITELN(out);
 ReadData(lambda, mu1, mu2, numberofdests, numberofsources,
          numberoftrials, maxmessages, discipline, servicerule,
          switchrule);
 InitialiseAllTrials(numberofdests, numberofsources, maxmessages,
                     factor, lambda, mu1, mu2, arrivalrate,
                     intensity, invnumdests, musystem,
                     overallwait1, overallwait2, rho1, rho2,
                     rhosystem, switchrule);
 PrintData(numberofdests, numberofsources, maxmessages,
           discipline, intensity, lambda, mu1, mu2, musystem,
           rho1, rho2, rhosystem, wait1, wait2, servicerule,
           switchrule, out);
 FOR trial:= 1 TO numberoftrials DO {overall stopping condition}
  EachTrial(numberofdests, numberofsources, trial, factor,
            maxmessages, discipline, arrivalrate, mu1, mu2,
            overallwait1, overallwait2, servicerule, switchrule,
            out);
 APPEND(out); WRITELN(out); WRITELN(out);
 overallav1:=overallwait1/numberoftrials;
 overallav2:=overallwait2/numberoftrials;
 overallav:=overallav1+overallav2;
 WRITELN(out, 'Mean over all trials',' ':11,overallav1:13:6,' ',
         overallav2:13:6,' ', overallav:13:6,' ':7);
 WRITELN(out);
 WRITE(out,'Ratio: Grand mean/FCFS Theory', '  ',
       (overallav1/wait1):13:6, ' ', (overallav2/wait2):13:6,' ');
 CASE switchrule OF
  cir: WRITELN(out, overallav/waitsystem:13:6);
  mes: WRITELN(out, overallav/(wait1+wait2):13:6)
 END;
 WRITELN(out);WRITELN(out);
 WRITELN(out,'-----------------------------------------------',
         '------------------------------');
 WRITELN(out);WRITELN(out); CLOSE(out);
 WRITELN; WRITELN('PROCESSING COMPLETE')
END.  {program}
```