SOME PROBLEMS CONCERNING
CLUSTER PROCESSES
AND OTHER POINT PROCESSES

by

Geoffrey Mark Laslett

A thesis submitted to the
Australian National University
for the degree of
Doctor of Philosophy
in the
Department of Statistics
Research School of Social Sciences
May, 1975

Unless otherwise stated, the material presented in this thesis is the product of my own original research.

G.M. Laslett

(G.M. Laslett)

CONTENTS

ACKNOWLEDGEMENTS

I am especially indebted to my supervisor, Dr D.J. Daley, for suggesting a substantial proportion of the problems investigated in this thesis, for his lively encouragement towards their solution and readiness to listen to any difficulties along the way. To Professor P.A.P. Moran, who supervised me for four months while Dr Daley was overseas, I am grateful for suggesting the material of Section 4.3 and for wise advice whenever it was required. Thanks are also due to Professor E.J. Hannan, who supervised me for six weeks in the absence of Dr Daley and Professor Moran, for allowing me to tap his extensive knowledge of ergodicity and related matters.

The friendly and stimulating environment of the whole Department of Statistics was a significant factor in the evolution of this thesis. In particular, I would like to mention Mr P.G. Hall for informative exchanges and for indirect assistance in demonstrating that one of the best ways of learning a new topic is attempting a problem on it, Dr R.L. Tweedie for providing sensible interpretations of some mathematical mysteries, and Mr D.B. Pollard for numerous enlightening conversations and genuine and constructive criticism, in particular concerning the desirability of a suggestive and consistent notation. Also, it is a pleasure to acknowledge the cooperation of Mrs B.C. Cranston in the solution of many problems of a non-mathematical nature.

Finally, I am grateful to Mrs B.M. Geary for her excellent typing of this thesis, and to S. Pollard and Dr I.C. Maclean for meticulous attention to detail in the proof-reading stage.

I also wish to acknowledge the hospitality of the Australian National University for the past three years, and the support of a Commonwealth Post Graduate Award.

obtain finer results. We believe that we have already partly countered the first objection, in that generality leads to clarity, which is a necessary first step to building more realistic generalized models, and, in addition, we feel that some idea of robustness of already available models is desirable. We fully concur with the second objection, but frequently the general method obtains optimal results (often more simply). Even if it does not, it is worth knowing what it can achieve, in order to elucidate the advantages of tools which exploit the distinctive features of particular models.

The transformation referred to in the opening paragraph will usually take the form of a mapping between measure spaces. Sometimes (rarely) measurability of this mapping is all the 'smoothness' we will require of it. Other times we will demand that it be more coherent, in the sense of preserving the component structure, and as this is usually not provided, we must impose conditions guaranteeing it. Typically we will find an approximating 'coherent' mapping, and the conditions will arise in demonstrating that the differences are irrelevant.

The question arises as to whether our approach can be generalized still further: if there are two or more component processes, can their independence be removed? The answer to this, in some cases, is in the affirmative. For example, it will be clear from the proof of Theorem 4.4.1 that mixing of cluster processes will follow from joint mixing of the centre and subsidiary processes. However, in those situations in which 'coherency' conditions on our mapping must be imposed, it becomes correspondingly more difficult to interpret their meaning, and probably also to verify. For this reason too (ease of writing down conditions), we have required throughout the thesis that our processes be stationary, although often, but not always, this assumption can, with care, be removed.

Chapter One, as well as providing a brief historical introduction and a definition of point processes, aims at setting down those concepts and classes of point processes which are frequently employed in the body of the dissertation. This still does not render the thesis self-contained, but minimizes to some extent the amount of external referencing required later.

Chapter Two investigates weak forms of asymptotic independence and (ordinary) limit laws for the number of busy servers process associated with the $G/G/\infty$ queue. In particular we demonstrate that stationarity, mixing, strong laws, the central limit theorem and rates of functional convergence derive from those of the arrival and server processes under reasonable conditions. Ergodicity does not, and rates of ordinary convergence seem to depend on a factor associated with the transformation operation.

Functional limit laws for cluster processes are examined in Chapter Three. A sufficient condition for existence of our generalized cluster process is given, and the measurability of the mapping defining the cluster process is investigated. Then, as in Chapter Two, the functional central limit theorem for the cluster process is shown, under suitable conditions predictably more severe than those of Daley (1972) for ordinary convergence, to flow from those of its components; the same is true for functional laws of the iterated logarithm for processes with right-hand clusters, but an unsolved problem is the extension to processes with left-hand clusters. Related topics (functional strong laws, the law of the iterated logarithm for the $G/G/\infty$ queue, and limit laws for doubly stochastic Poisson processes) are then mentioned.

Chapter Four deals with the preservation of strong forms of asymptotic independence under the clustering operation, initially attempted with a view to weakening the theorems in Chapter Three. This has not eventuated, however, but the theorems are of interest in their own right. Strong, ϕ- and complete mixing possess an increasing degree of uniformity of their asymptotic independence, and this turns out to be a significant factor in their preservation. In particular, it is indicated that ϕ-mixing may be maintained only under very stringent conditions (bounded clusters), whereas strong mixing is maintained under considerably milder conditions. It is also argued that the probability generating functional is not an applicable tool in these circumstances, even if the subsidiary processes are independent.

Chapter Five consists of characterization problems associated with renewal point processes, of which the main contribution is contained in a forthcoming paper (Laslett (1975)), in which we conclude that the output process of a finite capacity $GI/M/1$ queue (i.e., with renewal input) is never renewal. This is not unexpected, since the property of being a renewal process is very much a local one. This subsection is written in the language of queueing theory, unlike the rest of the thesis. We also look briefly at the problem of n (≥ 2) independent, identically distributed point processes being superposed to produce a renewal process. It is conjectured that all processes must then be Poisson, and proved in the case of the superposed processes being alternating renewal; counter-examples in the non-identical case have come from this area.

Chapter Six provides a list of unsolved and partially solved problems and generalizations. In addition to those associated with the bulk of the thesis, it also exhibits identifiability of the cluster structure of a stationary Poisson cluster process from a complete centre process-cluster process record. The problem is included in this chapter, because it is not solved in the generality required of the rest of this thesis, although it is conjectured that it can be.

For record and referral purposes, we have included a relatively extensive bibliography. It will be apparent that some of these references are only of an incidental nature, whereas others are more extensively applicable, in that they are required in the original sections of the thesis.

Once the Introduction and the definition of a cluster process (Sections 3.2 and 4.4) have been assimilated, the chapters may, apart from the unsolved problems, be read independently. The section 'Symbols and Abbreviations' should be perused before commencing on the thesis itself.

SYMBOLS AND ABBREVIATIONS

Some of the details which follow are almost conventional and are given simply to remove any possible ambiguity.

At times we write ab/cd rather than the more cumbersome $(ab)/(cd)$.

$\mathcal{B}(R)$	the usual Borel σ-field of the real line R
$\mathcal{B}(X)$	the Borel σ-field generated by the open sets of a specified topology on the space X
card$\{\cdot\}$	cardinality of the set $\{\cdot\}$
ch.fl.	characteristic functional
cov$\{\cdot, \cdot\}$	the covariance with respect to the appropriate probability measure
$E\{\cdot\}$	expectation with respect to the appropriate probability measure
i.i.d.	independent and identically distributed
P	(the probability measure of) a point process or other specified process
Pr	probability measure on the appropriate measure space
p.g.fl.	probability generating functional
$\Pi_i z_i$	product of real or complex numbers z_i
R	the real line
R_+	$[0, \infty)$
$\mathcal{R}(X)$	a specified ring of subsets of a set X
$\sigma(X)$	a specified σ-field of subsets of a set X
$\sigma(X)^T$	the restricted product σ-field on $X^{T'}$, where $T \subset T' \subset R$, i.e., the smallest σ-field containing sets of the form $\left\{x(\cdot) \in X^{T'} : x(t_1) \in C_1, \ldots, x(t_n) \in C_n\right\}$ for $n \in Z_+$, $t_1, \ldots, t_n \in T$ and $C_i \in \sigma(X)$. Unless

	otherwise specified $T' = T$, and then this corresponds to the usual product σ-field.
$\sigma_{\text{gen.}}(C)$	the minimal σ-field generated by sets in the class C
$\sigma(X_1) \times \sigma(X_2)$	the product σ-field of $\sigma(X_1)$ and $\sigma(X_2)$, i.e., $\sigma_{\text{gen.}}\{C_1 \times C_2 : C_1 \in \sigma(X_1), C_2 \in \sigma(X_2)\}$.
supp.f	the support of the function f
Var{•}	the variance with respect to the appropriate probability measure
X^T	the set of mappings $T \to X$. We will occasionally abuse this notation e.g. if card$\{T\} = n$, we may write X^n and the members of X^n may be indexed other than by $1, \ldots, n$.
Z	the set of all integers $\{0, \pm 1, \ldots\}$
Z_+	the set of non-negative integers $\{0, 1, \ldots\}$
$\bar{Z}_+$	$Z_+ \cup \{\infty\}$
$A + t$	$\{x+t : x \in A\}$ where A is a subset of R , $t \in R$.
A^c	complement of the set A
$\bar{A}$	closure of the set A
$\Sigma_i A_i$	disjoint union of sets A_i
$1_A(\cdot)$	indicator function of the set A
$\|A\|$	Lebesgue measure of the set A , where $A \in \mathcal{B}(R)$
$\|z\|$	absolute value of the real or complex number z
$\mathbf{a}.\mathbf{b}$	scalar product of vectors $\mathbf{a}$ and $\mathbf{b}$
$f \circ g$	composition of mappings f and g

$f(t) \sim g(t)$	$f(t)/g(t)$ tends to 1	as t tends to some limit (which may be infinite)
$f(t) = O\{g(t)\}$	$f(t)/g(t)$ remains bounded	
$f(t) = o\{g(t)\}$	$f(t)/g(t)$ tends to 0	

$\xrightarrow{\text{a.s.}}$ almost sure convergence

$\xrightarrow{p}$ convergence in probability

$\xrightarrow{\mathcal{D}}$ convergence in distribution

$\Rightarrow$ weak convergence (but also used as an implication sign)

$*$ convolution

$\square$ end of proof, or of statement of theorem etc., if no proof given

Section x.y is the yth section of Chapter x. Section x.y.z. is the zth subsection of Section x.y.

Statements of conjectures, definitions, lemmas, propositions and theorems are numbered consecutively (irrespective of type) within each section: thus Statement x.y.z. is the zth such statement in Section x.y.

Equations and similar entities are numbered consecutively within sections and subsections, i.e., equation (y.z) is the zth equation of Section x.y (for any x). Within chapters, we will refer to equation (y.z), but between chapters to equation (y.z) of Chapter x (similarly with equations from subsections).

CHAPTER 1

INTRODUCTION

1.1. Historical background

It is by no means an accident that many of the major works on point processes do not contain an historical account of the subject - the task of compiling such a survey would be formidable indeed. Hence we will attempt only a brief outline of the development of the subject; further references specific to the topics discussed in this thesis will be cited in the appropriate places. In this introduction we will not hesitate to draw on D. Vere-Jones' (1973) excellent account of the early history of point processes.

Fluctuations in the counts of objects in various situations seem to have been recognized as a stochastic problem since the latter half of the nineteenth century, although distributions other than the simple Poisson do not occur until the early 1920's, where they appear in problems on accident proneness (Greenwood and Yule (1920)) and contagion (Polya (1931)) Neyman (1939) introduced the important idea of clustering.

Population processes are intricately part of counting processes, and probably originate from the Bienaymé-Galton-Watson process. Development of the theory (for finite populations) begins with Feller (1950) and Bartlett (1954), and is extended by Moyal (1962) and Harris (1963). The study of particle showers (Bhabha (1950), Ramakrishnan (1950)) should also be mentioned as an early stimulus.

Considerations of distributions of interval lengths can be traced back as far as life-tables themselves, so that, surprisingly, renewal theory seems to have a longer history than counting problems. Attention seems to have been paid to the problem of relating the two properties, counts and intervals, of renewal processes on the line well before the 1940's (see

Lotka (1939)). The first treatment of processes with correlated intervals is by Wold (1949), to whom the term "point process" is due. He also made the first progress with processes with an infinite number of points.

Queueing theory has provided great impetus to the theory, particularly since Khinchin's monograph (1955). Palm (1943) attempted (without complete success) the first limit theorem for point processes (Poisson processes as the limit of n superposed processes, $n \to \infty$), as well as probing the question of Palm measures and functions. An account of the theory of superpositions may be read in Çinlar (1972); Palm functions and measures have been studied by Khinchin (1955), Slivnyak (1962), Ryll-Nardzewski (1961) and others.

During the 1960's and 1970's the theory has evolved rapidly, with expositions on stationarity and general properties (Matthes (1963a), Beutler and Leneman (1966)), statistical analysis (Cox and Lewis (1966)), spectral theory (Bartlett (1963), Daley (1971), Vere-Jones (1974)), Palm-Khinchin theory (Neveu (1968), Papangelou (1974), Leadbetter (1972a), Jagers (1973)), infinitely divisible point processes (Kerstan and Matthes (1964), Lee (1967), (1968)), cluster models (Neyman and Scott (1958), Lewis (1964a,b), (1969)) and probability generating functionals (Vere-Jones (1968), Westcott (1972)). Generalizations to multivariate (Milne (1971)), multidimensional (Fisher (1972) for a survey) and more abstract point processes (Mecke (1967), Kallenberg (1973), Jagers (1974) and Krickeberg (1974)) have been effected.

Of course in the above list we have omitted many topics and references (which are not intended to indicate priority), and more complete coverage can be obtained from Lewis (1972) and Kerstan, Matthes and Mecke (1974) and references therein. Daley and Milne (1973) have compiled a bibliography.

The rest of this chapter sets out the basic definitions and properties of point processes and weak convergence in a more coherent manner than their piecemeal introduction in the text of the thesis allows. After the formal definition in Section 1.2, we catalogue in Section 1.3 some known properties for easy reference later. Many of these have become part of the folklore, in which case reference to their source is omitted. The definitions of a few special processes needed in this thesis are given in Section 1.4, while weak convergence is summarised in Section 1.5.

1.2. The definition of a point process

Let N denote the set of non-negative integer-valued measures on the real line R which are finite on compact sets, and $\sigma(\mathsf{N})$ the σ-field generated by $\{\{N(A) \leq k\},\ k \in Z_+,\ \text{bounded } A \in B(R)\}$. Then we define a point process to be a probability measure on $(\mathsf{N}, \sigma(\mathsf{N}))$. Throughout this thesis we will reserve N (possibly subscripted) for members of N; P (possibly subscripted) will mean a point process, or other (specified) random process.

Denote by (Ω, F, Pr) an (arbitrary) probability space, and η (a random "N") a measurable mapping from (Ω, F, Pr) to $(\mathsf{N}, \sigma(\mathsf{N}))$. Then $P = Pr\eta^{-1}$ specifies a point process, but on some occasions, in particular those in which we wish to emphasize that a fixed underlying probability space is necessary for our arguments to be meaningful, we will refer to η as the point process.

The question of existence of measures P arises. As for ordinary stochastic processes, the problem is solved by extending the finite dimensional distributions of the process.

THEOREM 1.2.1 (Moyal (1962), Harris (1963), p. 53; Nawrotzki (1962)).

Given a set of functions

$\{p(A_1, \ldots, A_s; k_1, \ldots, k_s) : s \geq 1$ *and integral,* $k_i \in Z_+$, A_i *bounded* $\in B(R),\ 1 \leq i \leq s\}$,

satisfying the consistency conditions

(i) $p(A_1, \ldots, A_s; k_1, \ldots, k_s) = p(A_{i_1}, \ldots, A_{i_s}; k_{i_1}, \ldots, k_{i_s})$

for every permutation $(i_1, \ldots, i_s)$ *of* $(1, \ldots, s)$ *;*

(ii) $p(A_1, \ldots, A_s; k_1, \ldots, k_s) \geq 0$ *and*

$$\sum_{k_s=0}^{\infty} p(A_1, \ldots, A_{s-1}, A_s; k_1, \ldots, k_{s-1}, k_s)$$

$$= p(A_1, \ldots, A_{s-1}; k_1, \ldots, k_{s-1}) \ ;$$

(iii) $\sum_{k_1=0}^{\infty} p(A_1, k_1) = 1$ *;*

(iv) whenever $A_1, \ldots, A_s$ *are disjoint,*

$$p\left(\bigcup_{i=1}^{s} A_i; k\right) = \sum_{k_1+\ldots+k_s=k} p(A_1, \ldots, A_s; k_1, \ldots, k_s)$$

and

$$p\left(\bigcup_{i=1}^{s} A_i, A_1, \ldots, A_s; k, k_1, \ldots, k_s\right) = 0$$

unless $\sum_{i=1}^{s} k_i = k$ *, when it equals*

$p(A_1, \ldots, A_s; k_1, \ldots, k_s)$ *;*

(v) $p(A_k; 0) \uparrow 1$ *whenever* $A_k \downarrow \emptyset$ *;*

then there exists a unique probability measure P *on* $(N, \sigma(N))$ *for which*

$$P\{N : N(A_1) = k_1, \ldots, N(A_s) = k_s\} = p(A_1, \ldots, A_s; k_1, \ldots, k_s) \ . \qquad \square$$

It will be required to characterise P more finely (Chapter 6). Since the class of bounded half-open intervals with rational endpoints generates $B(R)$, we may restrict the A_i's to this class. Further, suppose we have functions $p_0(A_1, \ldots, A_s; k_1, \ldots, k_s)$ defined whenever the A's are disjoint; then the functions p_0 may be regarded as defining a joint distribution for random variables $\xi(A_1), \ldots, \xi(A_s)$ defined on a space

(Ω, F, Pr) . Suppose *(iv)* is modified as follows:

(iv)' *Let* $A_1, \ldots, A_s$ *be any mutually disjoint sets in* $B(R)$ *, and suppose* $A_i = \sum_{k=1}^{n_i} A_{ik}$ *,* $A_{ik} \in B(R)$ *, for some* $n_i = 1, 2, \ldots$ *; then the joint distribution of* $\xi(A_1), \ldots, \xi(A_s)$ *is the same as the joint distribution of*

$$\sum_{k=1}^{n_1} \xi(A_{1k}), \ldots, \sum_{k=1}^{n_s} \xi(A_{sk}) .$$

Then

THEOREM 1.2.2 (Harris (1963), p. 54). *A set of functions*

$$\{p_0(A_1, \ldots, A_s; k_1, \ldots, k_s) : s \geq 1 \text{ and integral, } k_i \in Z_+ , \text{ bounded disjoint } A_i \in B(R), 1 \leq i \leq s\}$$

satisfying (i) and (ii) of Theorem 1.2.1 *whenever the* A_i*'s are disjoint, and (iii), (iv)' and (v) can be uniquely extended to functions* $p(A_1, \ldots, A_s; k_1, \ldots, k_s)$ *satisfying (i)-(v) of Theorem* 1.2.1*, and agreeing with the* $p_0(\cdot; \cdot)$ *whenever the* A_i *are disjoint.* □

In particular, if our functions p_0 are generated by a point process, then the requirements of Theorem 1.2.2 are met. Also, we only need to know p_0 for A_i's being bounded half-open intervals with rational endpoints.

Although our processes will only be on the real line, there is no difficulty in extending the preceding concepts to R^n (e.g. Fisher (1972)) and Polish spaces (Jagers (1974)).

We may associate with each $N \in N$ the distance $t_j(N)$ of the jth point of N from the origin (commonly referred to as the jth epoch) defined as follows:

$$t_j \equiv t_j(N) = \inf\{y > 0 : N(0, y] \geq j\} , \quad j = 1, 2, \ldots ,$$

$$= \sup\{y \leq 0 : N[y, 0] > -j\} , \quad j = 0, -1, \ldots . \tag{2.1}$$

Note that $t_j : N \to R$ is measurable, since

$$\{N : t_j \leq x\} = \{N : N(0, x] \geq j\} \in \sigma(N) , \quad j = 1, 2, \ldots ,$$

$$\{N : t_j < x\} = \{N : N[x, 0] \leq -j\} \in \sigma(N), \quad j = 0, -1, -2, \ldots . \tag{2.2}$$

For any $N \in N$, the set $\{t_j(N)\}$ is a finite or countably infinite subset of R , multiple points included but with no finite limit points, and satisfying the inequalities

$$\ldots \leq t_{-1} \leq t_0 \leq 0 < t_1 \leq t_2 \ldots . \tag{2.3}$$

(This is not the conventional indexing (see Daley and Vere-Jones (1972), p. 308), but is employed to simplify the notation in Chapters 2 and 3.) If we denote the class of all such subsets of R by R_t^Z , then we can define $\eta : R_t^Z \to N$ by

$$\eta(A)(\mathbf{t}) = \mathrm{card}\{j : t_j \in \mathbf{t} \cap A\} , \quad A \in \mathcal{B}(R) , \quad \mathbf{t} \in R_t^Z ; \tag{2.4}$$

if also we introduce the σ-algebra $\sigma\left(R_t^Z\right)$ generated by the sets $\{\mathbf{t} : \eta(A)(\mathbf{t}) \leq k\}$, $A \in \mathcal{B}(R)$, $k \in Z_+$, i.e., such that the mappings η are measurable, then we may define a point process as a probability measure on $\left(R_t^Z, \sigma\left(R_t^Z\right)\right)$. It is intuitively clear that we may specify a point process also via the intervals $\{\Pi_j\} \equiv \{t_j - t_{j-1}\}$ and t_0 and t_1 , but, throughout this thesis, unless explicitly stated otherwise, we will consider our event epochs as generated in (2.1). We do this for two reasons: not only is $(N, \sigma(N))$ more amenable to generalization, but at times it will be the convenient to introduce the space N_∞ , the set of all non-negative integer-valued measures on R which may also be infinite on compact sets.

Finally we mention that Matthes and others have extended the scheme to *marked point processes*, where each point t_i is associated with a mark k_i

from a fixed measure space $[K, \sigma(K)]$. As this idea will only be employed circumspectly in this thesis, we will not expand on it. An account may be found in Kerstan, Matthes and Mecke (1974).

1.3. Basic properties of point processes

Define the translation operator $T_y : N \to N$, $y \in R$ by $T_y N(\cdot) = N(\cdot + y)$. A point process P is (strictly) stationary if

$$P(T_y C) = P(C) \text{ , all } C \in \sigma(N) . \tag{3.1}$$

If P is stationary, $P\{N(R) = 0 \text{ or } \infty\} = 1$, and P has no atoms in the sense that $P\{N(\{x\}) > 0\} = 0$ for all singleton point sets $\{x\}$ (Ryll-Nardzewski (1961)).

A stationary point process P is *ergodic* if any member G of the invariant σ-field $T_1 = \left\{G \in \sigma(N) : T_y^{-1} G = G\right\}$ satisfies $P(G) = 0$ or 1. Rosenblatt (1962) demonstrates that ergodicity may be charactized by

$$\lim_{\tau \to \infty} \tau^{-1} \int_0^\tau P(C \cap T_y D) dy = P(C)P(D) \text{ , all } C, D \in \sigma(N) . \tag{3.2}$$

P is *weakly mixing*[1] if

$$\lim_{\tau \to \infty} \tau^{-1} \int_0^\tau |P(C \cap T_y D) - P(C)P(D)| dy = 0 \text{ , all } C, D \in \sigma(N) , \tag{3.3}$$

and *mixing* if

$$\lim_{\tau \to \infty} P(C \cap T_\tau D) = P(C)P(D) \text{ , all } C, D \in \sigma(N) . \tag{3.4}$$

Note that mixing $\Rightarrow$ weak mixing $\Rightarrow$ ergodicity $\Rightarrow$ stationarity.

We will also use the following stronger forms of asymptotic independence. Let

$$\sigma(N(B)) \equiv \sigma_{\text{gen.}}\{\{N(A) = k\}, A \in \mathcal{B}(R) \cap B, k \in Z_+\} \text{ , } B \in \mathcal{B}(R) . \tag{3.5}$$

A stationary point process P is *strong mixing* with rate $\alpha(\tau)$ if

[1] The noun is 'weak mixing'; the adjective is 'weak mixing' or 'weakly mixing'; similarly with other types of mixing.

$$|P(C \cap D)-P(C)P(D)| \leq \alpha(\tau) \tag{3.6}$$

for all $C \in \sigma(N(-\infty, t])$, $D \in \sigma(N(t+\tau, \infty))$, $t \in R$, $\tau \geq 0$. Here $\alpha : [0, \infty) \to [0, 1]$ is a monotone decreasing function satisfying $\lim_{\tau\to\infty} \alpha(\tau) = 0$. Let C and D be as in (3.6). Then P is *ϕ-mixing* if

$$|P(C \cap D)-P(C)P(D)| \leq \phi(\tau)P(C) \tag{3.7}$$

and *completely mixing* (with rate $\gamma(\tau)$) if

$$|P(C \cap D)-P(C)P(D)| \leq \gamma(\tau)P(C)P(D) \tag{3.8}$$

where ϕ and γ have the same properties as α in (3.6). Clearly

complete mixing $\Rightarrow$ ϕ-mixing $\Rightarrow$ strong mixing $\Rightarrow$ mixing.

For a real valued function $g : N \to R$, we will define the expectation E of g by

$$E\{g\} \equiv E_P\{g\} = \int_N g(N)dP(N) . \tag{3.9}$$

Hence we can define the first moment measure M of a point process by $M(\cdot) = E\{N(\cdot)\}$, where existence of M is taken to mean $M(A) < \infty$ for all bounded Borel sets A . Clearly M is indeed a measure. If P is stationary, then easily $M(A) = m|A|$, where $m = EN(0, 1]$ is called the *intensity* of the point process. Higher moments are defined as

$$M_r(A_1 \times \dots \times A_r) = E\{N(A_1) \dots N(A_r)\} , \quad A_i \in \mathcal{B}(R) , \quad 1 \leq i \leq r , \tag{3.10}$$

and are easily shown to be measures in R^r . The cumulant measure C_2 exists if M_2 does, and is defined by

$$C_2(A_1 \times A_2) = M_2(A_1 \times A_2) - M(A_1)M(A_2) \equiv \mathrm{cov}(N(A_1), N(A_2)) . \tag{3.11}$$

C_2 may be a signed measure. If P is stationary, and $M_r(\cdot)$ exists, then $M_r(\cdot)$ is stationary in the sense that $M_r\{(A_1+x) \times \dots \times (A_r+x)\}$ is independent of $x \in R$.

A point process is defined to be *weakly stationary* if M_1 and M_2 exist and are stationary in the above sense (Daley (1971)). Important for us is the fact that the cumulant measure C_2 of a weakly stationary point process may be decomposed into Lebesgue measure and the *reduced covariance measure* $C(\cdot)$ in the following way (using differential notation)

$$C_2\big((t+dt) \times (t+u+du)\big) = C(du)dt \ , \tag{3.12}$$

and in particular (Daley and Vere-Jones (1972), p. 323)

$$\operatorname{Var}\big(N(0,\ y]\big) = C_2\big((0,\ y] \times (0,\ y]\big) = \int_{-y}^{+y} (y-|u|)C(du) \ . \tag{3.13}$$

The *superposition* of n independent point processes $P_1, \ldots, P_n$ is intuitively the overlaying of all the points on one line, but may be specified rigorously via the mapping $\eta : N^n \to N$ defined by

$$\eta(\{N_i\}) = \sum_{i=1}^{n} N_i \ . \tag{3.14}$$

Then $P = \left(\bigtimes_{i=1}^{n} P_i\right)\eta^{-1}$ is the process of superpositions. Note that η may be regarded as the sum of n coordinate mappings η_i each defined on N^n. It will be convenient on occasions to conform to the usual practice of supposing the summands to be random measures $\eta_i : (\Omega,\ F,\ Pr) \to \big(N,\ \sigma(N)\big)$ on a fixed space Ω, and defining their superposition η by $\eta(\omega) = \sum_{i=1}^{n} \eta_i(\omega)$. This scheme allows for dependent η_i.

We recall that for functions $f : R \to R$ of compact support,

$$\int_{-\infty}^{+\infty} f(t)dN(t) = \Sigma_i f\big(t_i\big) \tag{3.15}$$

is well defined. By Fubini's theorem

$$E\left\{\int_{-\infty}^{+\infty} f(t)dN(t)\right\} = \int_{-\infty}^{+\infty} f(t)dM(t) \tag{3.16}$$

if $\int_{-\infty}^{+\infty} |f(t)|dM(t) < \infty$.

Supposing $\log \xi(t)$ to have compact support, we may define the *probability generating functional* (p.g.fl.) of a point process P by

$$G[\xi] = E\left\{\exp \int_{-\infty}^{+\infty} \log \xi(t) dN(t)\right\} . \tag{3.17}$$

We will denote by V the class of measurable functions $\xi : R \to [0, 1]$ such that $1 - \xi$ has compact support. G exists if ξ is in this class, or if $P\{N(R) < \infty\} = 1$. For an extensive discussion of the p.g.fl. see Westcott (1972). Note that if $P_1, \ldots, P_n$ are n independent point processes with p.g.fls. $G_1, \ldots, G_n$, then the superposition P has p.g.fl. $G[\xi] = \prod_{i=1}^{n} G_i[\xi]$.

A point process P is said to be *a.s. orderly* (or without multiple points) if

$$P\{N(\{x\}) = 0 \text{ or } 1 \text{ (all } x)\} = 1 , \tag{3.18}$$

and a *stationary* process P is *orderly* (or analytically orderly) if

$$P\{N(0, h] \geq 2\} = o(h) \quad (h \downarrow 0) . \tag{3.19}$$

If P is stationary with finite intensity m, then analytically orderly $\Longleftrightarrow$ a.s. orderly ($\Leftarrow$ is Dobrushin's Lemma). We shall also require Khinchin's existence theorem, namely that for a stationary point process P the *rate*

$$\lambda = \lim_{h \downarrow 0} P\{N(0, h] > 0\}/h \tag{3.20}$$

exists, although it may be infinite. An interesting relation exists between m and λ: for a stationary point process, the 'batch-size' distribution for the number of events in a multiple occurrence exists, and has mean m/λ. Hence $\lambda \leq m$, and a necessary and sufficient condition for $m = \lambda$ is that the process be orderly.

An important recent discovery is the Rényi-Mönch-Kallenberg Theorem: if P is a.s. orderly, then it is determined by knowing

$$\phi(I) = P\{N(I) = 0\} \tag{3.21}$$

for all I in the semi-ring generated by half-open intervals. Kurtz (1974) has characterized the functions ϕ in terms of a property he calls complete monotonicity.

Further properties may be found in the references cited in Section 1.1.

1.4. Some examples of point processes

The *Poisson process* η with parameter $\lambda(\cdot)$ (a non-atomic Borel measure) is defined by

(i) For $A \in \mathcal{B}(R)$ such that $\lambda(A) < \infty$, $\eta(A)$ has a Poisson distribution with mean $\lambda(A)$. Otherwise $\eta(A) = \infty$ with probability one.

(ii) η is completely random, i.e., $\eta(A_1), \ldots, \eta(A_n)$ are mutually independent for all finite collections of disjoint sets $A_1, \ldots, A_n \in \mathcal{B}(R)$.

The Poisson process has been characterized in many ways: Prékopa (1957a,b) showed that a point process is Poisson if and only if it is atomless, completely random and has no multiple points. On the other hand, if

$$Pr\{\eta(I) = 0\} = e^{-\lambda(I)} \quad \text{and} \quad Pr\{\eta(I) \geq 2\} = o(\lambda(I)) \;, \quad \lambda(I) \downarrow 0 \;, \qquad (4.1)$$

for some non-atomic $\lambda(\cdot)$ and all I consisting of a finite union of intervals, then η is Poisson (Rényi (1967)). If $\lambda(A) = \lambda|A|$, some $\lambda < \infty$, we recover the stationary Poisson process.

If the parameter measure of a Poisson process η is taken to be the realization of a random measure $\Lambda(\cdot)$, we obtain the *doubly stochastic Poisson process*. Kingman (1964) characterized doubly stochastic Poisson processes as stationary Poisson processes with unit parameter subjected to a random change of time independent of the original process, an important and too often neglected result.

The best example of a point process which is generated by its interval properties is the *renewal process*, which starts at time 0 and has i.i.d. inter-event times. The *stationary* renewal process has the time t_1 to the first event delayed according to the distribution

$$Pr\{t_1 \le u\} = \lambda \int_0^u \{1-F(x)\}dx \ , \tag{4.2}$$

where F is the distribution function of inter-event times, and λ^{-1} its (finite) mean.

Alternating renewal processes have independent inter-event times, but their distribution switches successively from one lifetime distribution function F_1 to another F_2. The concept may be generalised. Processes with consecutive pairs of points equidistant we will call *deterministic*, and, as a general rule, are very useful for providing counter-examples.

A point process P is *infinitely divisible* if for every $n = 1, 2, \ldots$ it may be represented as the superposition of n i.i.d. point processes $P_{n,i}$ $(i = 1, \ldots, n)$. Other definitions are possible. Since such processes are the subject of a book by Kerstan, Matthes and Mecke (1974), they will not be discussed in this work, except when they occur incidentally as Poisson cluster processes (see the cited reference).

Point processes may be generated in many other ways: e.g. as the transition times of Markov processes (Rudemo (1973)), or by level crossings (Leadbetter (1972b)). We shall not attempt to expand on these, as this thesis is not concerned with specific point processes, except for illustrative purposes.

1.5. A summary of some weak convergence concepts

If probability measures P, P_n on a separable metric space S with metric d (referred to as (S, d)) and Borel sets $\mathcal{B}(S)$ satisfy

$$\int_S f dP_n \to \int_S f dP \quad (n \to \infty) \tag{5.1}$$

for every bounded, continuous real function f on S, then we say P_n

converges weakly to P and write $P_n \Rightarrow P$. The theory of weak convergence on metric spaces is outlined in Billingsley (1968), and for most theorems we need we will refer to this text. However, we mention the *continuous mapping theorem* (Theorem 5.1 of Billingsley (1968)), which we often use. Let $h : (S, d) \to (S', d')$ be a measurable mapping with a set of discontinuities D_h.

CONTINUOUS MAPPING THEOREM. *If* $P_n \Rightarrow P$ *and* $P(D_h = 0)$, *then* $P_n h^{-1} \Rightarrow Ph^{-1}$. □

Various metrics are associated with weak convergence: since we shall use almost all of them, we give a brief summary (adapted from Whitt (1974b)).

Let $RV \equiv RV(S, d)$ denote the set of all random variables $(\Omega, F, Pr) \to (S, d)$ defined on a fixed space Ω, and $P(S, d)$ the space of all probability measures on $(S, B(S))$.

Firstly, for $X_1, X_2 \in RV$, define

$$\alpha(X_1, X_2) = \inf\{\varepsilon \geq 0 : Pr[d(X_1, X_2) \geq \varepsilon] \leq \varepsilon\} . \tag{5.2}$$

α corresponds to convergence in probability.

Now, for any $\varepsilon > 0$, define

$$A^\varepsilon = \{y : d(x, y) < \varepsilon \text{ for some } x \in A\} . \tag{5.3}$$

The Prohorov metric ρ which induces the topology of weak convergence on P is given by

$$\begin{aligned} \rho(P_1, P_2) &= \max\{\gamma(P_1, P_2), \gamma(P_2, P_1)\} \\ \gamma(P_1, P_2) &= \inf\left\{\varepsilon \geq 0 : P_1(F) \leq \varepsilon + P_2(F^\varepsilon), F \text{ closed}\right\} . \end{aligned} \tag{5.4}$$

Note immediately that ρ may be regarded as acting on RV, by setting $\rho(X_1, X_2) = \rho\left(PrX_1^{-1}, PrX_2^{-1}\right)$, but is now only a pseudometric. Clearly $\rho(X_1, X_2) \leq \alpha(X_1, X_2)$. Dudley (1968) proved that $\gamma(P_1, P_2) = \gamma(P_2, P_1)$, provided $P_1(S) = P_2(S)$ and there are no restrictions on closed F's, conditions satisfied here. The Lévy metric λ on $P(R)$ is the Prohorov metric restricted to closed sets of the form $(-\infty, x]$. Clearly $\lambda \leq \rho$, but it is well known that λ characterises weak convergence (e.g. Loève (1960), p. 215).

The supremum metric σ is defined by

$$\sigma(P_1, P_2) = \sup\{|P_1(A)-P_2(A)|, A \in B(S)\} . \tag{5.5}$$

σ generates a stronger topology on P than ρ. We shall also refer to the restriction of σ on $P(R)$ to sets of the form $(-\infty, x]$ as the supremum metric ν. Now $\lambda \leq \nu$, but ν also metrizes weak convergence to those P in $P(R)$ with continuous distribution functions.

Dudley (1966) demonstrates that the dual bounded Lipschitz metric β also induces the weak convergence topology. A function $f : (S, d) \to R$ is Lipschitz if

$$\|f\|_1 = \sup_{x \neq y} \{|f(x)-f(y)|/d(x, y)\} < \infty . \tag{5.6}$$

For any such f we may associate the norm

$$\|f\| = \|f\|_1 + \|f\|_\infty , \tag{5.7}$$

where $\|f\|_\infty = \sup\{|f(x)|, x \in S\}$. Then β is defined by

$$\beta(P_1, P_2) = \sup\left\{\left|\int f dP_1 - \int f dP_2\right|, \|f\| \leq 1\right\} . \tag{5.8}$$

Also, according to Dudley (1968), $\beta \leq 2\rho$, so that we may summarize the inter-relationship between the metrics (regarded as pseudo-metrics on RV) by

$$\begin{array}{ccccc} \alpha & & & & \beta/2 \\ & \geq & & \leq & \\ & & \rho & & \\ & \leq & & \geq & \\ \sigma & & & & \lambda \\ & \geq & & \leq & \\ & & \nu & & \end{array} .$$

Closely related to weak convergence is vague convergence, a concept of great importance to point processes: a sequence $\{N_n\}$ of measures $N_n \in N$ converges vaguely to $N \in N$ if

$$\int f dN_n \to \int f dN \quad (n \to \infty) \tag{5.9}$$

for every $f \in C_K$, the set of continuous functions $f : R \to R$ with compact support. This convergence generates the vague topology on N , a basis for which is given by

$$\left\{N \in N : \left|\int f_j dN - \int f_j dN_1\right| < \varepsilon,\ 1 \le j \le n\right\}, \tag{5.10}$$

$n = 1, 2, \ldots$, $f_j \in C_K$, $N_1 \in N$, $\varepsilon > 0$. The vague topology is metrizable (by the Prohorov metric) and renders N separable. Weak convergence of probability measures on $(N, B(N))$ may then be studied (Jagers (1974)). The vague topology is a natural topology for N , since, if $B(N)$ denotes the Borel σ-field generated by the open sets of the vague topology, $B(N) = \sigma(N)$ by Proposition 1.1 of Jagers (1974).

Weak convergence may also be studied on the space $C[0, 1]$ of continuous functions on $[0, 1]$, where $C[0, 1]$ has the uniform topology (U) induced by the metric ρ (no confusion with the Prohorov metric ρ should arise)

$$\rho(x, y) = \sup_{0 \le t \le 1} |x(t)-y(t)| \ , \quad x, y \in C \ . \tag{5.11}$$

More appropriate to the study of point processes is the space $D[0, 1]$ of functions on $[0, 1]$ which are right-continuous and have left-hand limits. D will be endowed with the J_1-topology induced by the Skorokhod metric (using $I \equiv I(t) = t$ and $y \circ \lambda(t) = y(\lambda(t))$)

$$d(x, y) = \inf_{\lambda \in \Lambda} \{\max(\rho(\lambda, I), \rho(x, y \circ \lambda))\} \tag{5.12}$$

where Λ consists of all continuous strictly increasing maps of $[0, 1]$ onto itself. We need note only that $d(x, y) \le \rho(x, y)$, $x, y \in D$, but $d(x_n, x) \to 0 \Rightarrow \rho(x_n, x) \to 0$ $(n \to \infty)$ if x is continuous.

Weak convergence of probability measures on C and D is characterized by convergence of the finite-dimensional distributions and tightness (Billingsley (1968), Theorems 8.1 and 15.1). Necessary and sufficient conditions for tightness of a sequence $\{P_n\}$ of probability measures on C

may be established via the Arzelà-Ascoli theorem in terms of the modulus of continuity of $x \in C$,

$$w_x(\delta) = \sup_{|s-t|<\delta} |x(s)-x(t)| \ . \tag{5.13}$$

The following elementary inequality exists between w and ρ (Billingsley (1968), p. 220),

$$|w_x(\delta)-w_y(\delta)| \leq 2\rho(x, y) \ , \quad x, y \in D[0, 1] \ . \tag{5.14}$$

Weak convergence on $C[0, 1]$ or $D[0, 1]$ will be referred to as functional convergence. Jagers ((1974), Proposition 3.3) relates functional and weak convergence of a sequence $\{P_n\}$ of point processes converging to a point process P . In this case tightness is unnecessary (Straf (1972), Whitt (1975)), provided $D[0, 1]$ has the J_1-topology if P is orderly, or the M_1-topology (see Skorokhod (1956)) if P is non-orderly. Hence p.g.fls. may be used to characterize weak convergence of point processes (Westcott (1972)).

CHAPTER 2

LIMIT LAWS FOR THE $G/G/\infty$ QUEUE

2.1. Introduction

The $G/G/\infty$ queue may be regarded as a point process (the process of arrival times) subjected to independent random displacements. An important process associated with the $G/G/\infty$ queue is the number of servers $\phi(t)$ busy at any time t. Central limit theorems for the accumulated traffic time $\int_0^t \phi(s)ds$ have been obtained in the case of the $M/G/\infty$ queue with bulk arrivals by Rao (1966), under the assumption of finite third service time moment, and the $GI/G/\infty$ queue with bulk arrivals by Narasimham (1968), for finite second moment. Iglehart and Kennedy (1970) generalised the model to the case in which the service times may be mutually dependent, and, using functional weak convergence techniques, demonstrated a functional central limit theorem for $\int_0^t \phi(s)ds$, which, specializing to i.i.d. service times, required finite $(2+\delta)$-th moment for the service time distribution. By confining ourselves to the $G/G/\infty$ queue in which the arrival time point process is stationary, we obtain in Section 2.5 a central limit theorem for $\int_0^t \phi(s)ds$ when the service times have finite second moment.

Brown and Ross (1969) studied almost sure convergence of the $M/G/\infty$ queue with bulk arrivals, and in particular examined the strong law of large numbers for this case. In Section 2.3 we look at stationarity, ergodicity an[d] mixing of the $\phi(s)$ process, and in Section 2.4 derive conditions under whic[h] it obeys the strong law of large numbers. We defer investigation of the law of the iterated logarithm to Chapter 3, where it will be discussed in conjunction with cluster processes. Finally, in Sections 2.6 and 2.7, we enquire into rates of convergence (both ordinary and functional) of the accumulated traffic time to normality.

In each of the above limit theorems our assumptions will be too general to allow the use of specialized techniques, predominant in weak and a.s. convergence applications, such as (with particular reference to the $G/G/\infty$ queue) renewal theory, the Skorkhod representation theorem (in Section 2.7), or the probability generating functional (Vere-Jones (1968), Westcott (1970), Section 3.6). Indeed, in accordance with the philosophy outlined in the Preface to this thesis, we will always ask, in a way to be defined more strictly later, that the limit law we desire of the service expended $\int_0^t \phi(s)ds$ be obeyed by the component processes - for other samples of this attitude, see Iglehart and Kennedy (1970) and Loulou (1973). In particular, we will retain the assumption of dependence between service times. This particular chapter was precipitated by a paper of Kaplan (1974), in which the various limit theorems (except rates of convergence) were obtained for the $GI/G/\infty$ queue by methods strongly depending on the independence assumptions of that system.

Many of the theorems of this chapter have been stated as assuming stationarity, a condition which seems necessary in Sections 2.3 and 2.4, but may (see Problem 6.3.4) be removed in later sections. In some of the theorems of this chapter, two almost identical results exist, and here we have indicated the differences of the second from the first in brackets in the statement of the theorem.

Infinite server queues have been widely applied. The pure birth-and-death process is an $M/M/\infty$ queue (Feller (1968), p. 460). Amongst early investigations were those by C. Palm, A.K. Erlang and T.C. Fry in telephone trunking problems (see Feller (*ibid.*) for references). Benes (1957) obtained the distribution of traffic time average in the $M/M/\infty$ case.

T. Lewis (1961) and Nelsen and Williams ((1968), (1970)) consider regular events under arbitrary i.i.d. translations, i.e., the $D/G/\infty$ queue. Smith (1958) has used the generating function of $\phi(s)$ in the $GI/G/\infty$ queue to illustrate infinite products occurring in renewal theory. Thedeen (1969) discusses the queue as a model for traffic on a long road with free overtaking. For applications to textile research, see Rao (1966) and references there. The queue arises in engineering applications as the "jitter" process (Beutler and Leneman (1966)). Other works on the $G/G/\infty$ queue will be referred to in Section 6.2.

2.2. Notation and preliminaries

It will be convenient to regard the arrival process as a probability measure P_1 on $\left(N, \sigma(N)\right)$, although we will also refer to it as a random non-negative integer-valued measure $\eta : (\Omega, F, Pr) \to \left(N, \sigma(N)\right)$, so that $P_1 = Pr\eta^{-1}$. As always, we may define measurable mappings $t_j : N \to R$ giving the arrival times $t_j(N)$ for any realisation of the arrival process N.

The process of service times may be viewed as a probability measure P_2 on $\left(R_+^Z, B\left(R_+\right)^Z\right)$, where Z is the set of integers; alternatively, (in Section 2.7 and Chapter 4), we will specify it as a random vector $\{V_j\}$ of service times from a fixed space (Ω, F, Pr) to $\left(R_+^Z, B\left(R_+\right)^Z\right)$, so that $P_2 = Pr\{V_j\}^{-1}$.

Finally, we form the probability triple $\left(N\times R_+^Z, \sigma(N)\times B(R_+)^Z, P_1\times P_2\right)$; i.e., we regard the arrival process and service process as independent. We will denote vectors $\{x_j\} \in R_+^Z$ as x if no ambiguity will arise.

Let $\overline{Z}_+$ denote $Z_+ \cup \{\infty\}$. Then $\overline{Z}_+^R$ will mean the set of non-negative integer or infinite valued functions on R. Let $\sigma(\overline{Z}_+)^R$ denote the product σ-field on $\overline{Z}_+^R$, where $\sigma(\overline{Z}_+)$ consists of all subsets of $\overline{Z}_+$.

We can then define the 'number of busy servers' process via a mapping $\phi = \phi_1 \circ \phi_2$ from $N \times R_+^Z \to \overline{Z}_+^R$ by

(i) $\phi_2 : N \times R_+^Z \to R^Z \times R^Z$

$$\phi_2(N, \mathrm{x}) = \left(\mathrm{t}(N), \mathrm{x}\right)$$

(ii) $\phi_1 : R^Z \times R^Z \to \overline{Z}_+^R$

$$\phi_1\left(\{w_j\}, \{x_j\}\right)(t) = \sum_{j=-\infty}^{+\infty} 1_{[w_j, w_j+x_j)}(t). \tag{2.1}$$

It is clear that for any joint realization (N, x) of the arrival and server processes, ϕ will record 1 if an arrival occurs up to and including t and is still being served after t, 0 otherwise and then will add these numbers, therefore corresponding to the number of busy servers at t, for each $t \in R$. To prove measurability of a mapping to a product space, we need only prove that each coordinate mapping is measurable. Hence it is immediately clear that

THEOREM 2.2.1. *ϕ is measurable with respect to* $\sigma(N) \times \mathcal{B}(R_+)^Z$. □

Throughout this chapter we assume that the number of busy servers is finite, i.e., that $\phi \in Z_+^R$, $P_1 \times P_2$ - a.s. We then define a probability measure P_{ns} (corresponding to the number of busy servers process) on $\left(\overline{Z}_+^R, \sigma(\overline{Z}_+)^R\right)$ by

$$P_{ns} = \left(P_1 \times P_2\right)\phi^{-1},$$

such that $P_{ns}\left(Z_+^R\right) = 1$. From now on we will refer to P_{ns} as if it were acting on Z_+^R. We may similarly define a stochastic process

$m : (\Omega, F, Pr) \to (\overline{Z}_+^R, \sigma(\overline{Z}_+)^R)$ such that $Prm^{-1} = P_{ns}$. We will refer to P_{ns}, ϕ , or m as the "number of busy servers" process.

We will frequently employ the following decompositions.

$$\int_0^\tau \phi(s)ds = \sum_{i=1}^{N(0,\tau]} \min(x_i,\ \tau-t_i) + \sum_{i=-\infty}^{0} [\min(x_i+t_i,\ \tau)]^+$$

$$\equiv \Phi_1(\tau) + \Phi_2(\tau) \ , \tag{2.2}$$

say. Here we have split the integral into two parts, corresponding to those arriving up to time zero and those arriving later. Also, for some sequence $c(\tau) \le \tau$, define

$$\Phi'(\tau) \equiv \left|\Phi_1(\tau) - \sum_{i=1}^{N(0,\tau]} x_i\right| \tag{2.3}$$

$$= \sum_{i=1}^{N(0,\tau]} (x_i+t_i-\tau)^+$$

$$= \sum_{i=1}^{N(0,\tau-c(\tau)]} (x_i+t_i-\tau)^+ + \sum_{i=N(0,\tau-c(\tau)]+1}^{N(0,\tau]} (x_i+t_i-\tau)^+$$

$$\le \sum_{i=1}^{N(0,\tau]} (x_i-c(\tau))^+ + \sum_{i=N(0,\tau-c(\tau)]+1}^{N(0,\tau]} x_i$$

$$\equiv \Phi_0'(\tau) + \Phi_1'(\tau) \ , \tag{2.4}$$

say, where we have used a device due to Iglehart and Kennedy (1970). We will decompose $\int_0^\tau m(s)ds$ similarly as

$$\int_0^\tau m(s)ds = M_1(\tau) + M_2(\tau) \ , \tag{2.5}$$

$$M'(\tau) \le M_0'(\tau) + M_1'(\tau) \ . \tag{2.6}$$

Note that (2.2) and (2.4) (similarly with (2.5) and (2.6)) give

$$\left|\int_0^\tau \phi(s)ds - \sum_{i=1}^{N(0,\tau]} x_i\right| \le \Phi_0'(\tau) + \Phi_1'(\tau) + \Phi_2(\tau) \ . \tag{2.7}$$

Thus we have bounded the difference between the quantity we are interested in, and another quantity in which the structure of the component processes is more evident.

2.3. Stationarity, ergodicity and mixing

To examine stationarity of P_{ns}, we must introduce certain translation operators T_t $(t \in R)$, S_k $(k \in Z)$ and S_u $(u \in R)$ as follows.

$$T_t : N \to N \quad \text{where} \quad T_t N(A) = N(A+t)\ ,\ \text{all}\ A \in \mathcal{B}(R)\ , \tag{3.1}$$

$$S_k : R^Z \to R^Z \quad \text{where} \quad S_k\{x_j\} = \{x_{k+j}\}\ , \tag{3.2}$$

$$S_u : Z_+^R \to Z_+^R \quad \text{where} \quad S_u\{x(t)\} = \{x(t+u)\}\ . \tag{3.3}$$

The processes P_1, P_2 and P_{ns} are called stationary if $P_1\left(T_t^{-1}B\right) = P_1(B)$, $P_2\left(S_k^{-1}B\right) = P_2(B)$, $P_{ns}\left(S_u^{-1}B\right) = P_{ns}(B)$ for B within the appropriate σ-algebra, where $T_t^{-1}B$, $S_k^{-1}B$ and $S_u^{-1}B$ are the corresponding inverse image sets. Mixing and related concepts can be defined in a way analogous to (3.4) of Chapter 1 for any stationary process whose translation operator is indexed by Z or R. Before proving our next theorem, we note that, as is well-known from ergodic theory, it is sufficient to prove ergodicity for a class of sets $\mathcal{C}$ (i.e. for $C, D \in \mathcal{C}$) such that $\left\{\left\{\bigcup_{i=1}^{n} C_i\right\}, C_i \in \mathcal{C}, n \in Z_+\right\}$ is an algebra $\mathcal{A}$ which generates the σ-field we require.

THEOREM 2.3.1. *(a) If the arrival process* P_1 *and the service time process* P_2 *are stationary, then so is the "number of servers" process* P_{ns} .

(b) If P_1 *is ergodic and* P_2 *is mixing, then* P_{ns} *is ergodic.*

(c) If P_1, P_2 *are mixing, then so is* P_{ns} .

Proof. We define $T'_y : N \times R_+^Z \to N \times R_+^Z$, $y \in R$, by

$$T'_y(N, x) = \left(T_y N, S_{N(0,y]}x\right) , \quad y > 0 ,$$
$$= \left(T_y N, S_{N[y,0)}x\right) , \quad y < 0 ,$$
$$= (N, x) , \quad y = 0 . \tag{3.4}$$

For any sets $C, D \in \sigma(Z_+)^R$, we see that

$$P_{ns}\left(C \cap S_u D\right) = P_1 \times P_2\left\{\left(\phi^{-1}C\right) \cap T'^{-1}_u\left(\phi^{-1}D\right)\right\} ,$$

where $\phi^{-1}C \equiv \left\{(N, x) : \phi(N, x) \in C\right\} \in \sigma(N) \times B(R_+)^Z$ by Theorem 2.2.1. Hence stationarity, ergodicity and mixing of P_{ns} will follow from stationarity, ergodicity and mixing of $P_1 \times P_2$ with respect to the transformation T' . We prove only *(b)*: proofs for *(a)* and *(c)* are similar.

Let $C' = C_1 \times C_2$, $D' = D_1 \times D_2 \in \sigma(N) \times B(R_+)^Z$.

Then

$$\left|\tau^{-1}\int_0^\tau P_1{\times}P_2\left(C' \cap T'_u D'\right)du - P_1{\times}P_2(C')P_1{\times}P_2(D')\right|$$

$$= \Big|\tau^{-1}\int_0^\tau \sum_{n=0}^\infty P_1\left(C_1 \cap T_u D_1, \{N(0, u] = n\}\right)\left[P_2\left(C_2 \cap S_n D_2\right) - P_2\left(C_2\right)P_2\left(D_2\right)\right]du$$
$$+P_2\left(C_2\right)P_2\left(D_2\right)\tau^{-1}\int_0^\tau \left[P_1\left(C_1 \cap T_u D_1\right) - P_1\left(C_1\right)P_1\left(D_1\right)\right]du\Big|$$

$$\le 2\tau^{-1}\int_0^\tau P_1\{N(0, u] \le m\}du + \sup_{n>m}\left|P_2\left(C_2 \cap S_n D_2\right) - P_2\left(C_2\right)P_2\left(D_2\right)\right|$$
$$+ \left|\tau^{-1}\int_0^\tau P_1\left(C_1 \cap T_u D_1\right)du - P_1\left(C_1\right)P_1\left(D_1\right)\right| . \tag{3.5}$$

But for any $\varepsilon > 0$, there exists $m(\varepsilon)$ such that for $n > m(\varepsilon)$,

$$\sup_{n>m(\varepsilon)}\left|P_2\left(C_2 \cap S_n D_2\right) - P_2\left(C_2\right)P_2\left(D_2\right)\right| \le \varepsilon ,$$

since P_2 is mixing. Thus $\limsup_{\tau\to\infty} (3.5) \le \varepsilon$, i.e., $\lim_{\tau\to\infty} (3.5) = 0$. By the remark immediately preceding Theorem 2.3.1, this proves *(b)* and completes our proof of the theorem. □

It is clear from the above proof that if we had defined

$T''_y : N \times R_+^Z \to N \times R_+^Z$ by

$$T''_y(N, \mathrm{x}) = \left(T_y N, S_{[y]}\mathrm{x}\right) , \tag{3.6}$$

ergodicity of $P_1 \times P_2$ would have followed from ergodicity of P_1 and weak mixing (Equation (3.3), Chapter 1) of P_2. Spectral theory (see e.g. Mackey (1974), p. 213) suggests that this condition is necessary as well as sufficient for T''_y, but we have been unable to weaken Theorem 2.3.1 *(b)* to include this case for T'_y (but see Problem 6.3.1). Note that the argument above actually becomes slightly more difficult when modified to the independent service times situation, if it is not realised that then P_2 is mixing.

2.4. The strong law of large numbers

A theorem associated with ergodicity is the strong law of large numbers. Before we can prove it, however, we need some new equipment.

We will suppose as in the previous section that our arrival process P_1 is stationary, but now with finite intensity, i.e.,

$$m \equiv E_1 N(0, 1] < \infty ; \tag{4.1}$$

if also P_1 is orderly (for P_1 non-orderly, see discussion after Corollary 2.4.2), then it is known (e.g. Slivnyak (1962), (1966)) that uniquely associated with P_1 is a probability measure P_1^0 (Palm measure) on $(N, \sigma(N))$ (in general, non-stationary) such that the arrival times t satisfy $\ldots < t_{-1} < t_0 = 0 < t_1 < \ldots$ P_1^0 - a.s., and the sequence $\{\Pi_j\} \equiv \{t_{j+1} - t_j\}$ of a.s. positive random variables is stationary, i.e., if we define $\theta_k : N \to N \quad (k \in Z)$ by

$$\theta_k N(\cdot) = N(\cdot + t_k) \tag{4.2}$$

then, if $C \in \sigma_{\text{gen.}}\{\{\Pi_j \in A\}, A \in B(R), j \in Z\} \equiv \sigma_\pi \subset \sigma(N)$,

$$P_1^0\left(\theta_k^{-1}C\right) = P_1^0(C) \ . \tag{4.3}$$

We will require the expectation

$$U(x) \equiv E_1^0 N(0, x] < \infty \tag{4.4}$$

for some finite x . According to Kaplan (1955), this guarantees finiteness for all finite x , and also

$$\sup_{x>0} \{U(x+u) - U(x)\} \le 2U(u) + 1 \ , \tag{4.5}$$

whence $U(x) = O(x)$, $(x \to \infty)$. Under this assumption of orderliness, Daley ((1971), Lemma 9) shows that $U(x) \sim \lambda' x$ for some constant $\lambda' \ge m$, which is stronger. We will tacitly assume when discussing P_1^0 in the rest of this chapter that the P_1 to which it corresponds is orderly. We point out, though, that our arguments involving only P_1 (not P_1^0 or $U(x)$) will still go through without this assumption.

Finally we define various invariant σ-fields: let

$T \equiv \left\{C \in \sigma(N) \times B(R_+)^Z : T_y'^{-1}C = C\right\}$, $T_1 \equiv \left\{C \in \sigma(N) : T_y^{-1}C = C\right\}$,

$T_2 \equiv \left\{C \in B(R_+)^Z : S_k^{-1}C = C\right\}$ and $T_1^0 \equiv \left\{C \in \sigma_\pi : \theta_k^{-1}C = C\right\}$, where we have used the definitions at (3.1), (3.2), (3.3) and (4.2).

We may now prove some ergodic theorems for the "busy server" process.

THEOREM 2.4.1. *If the arrival process is stationary with* $m < \infty$ *, and the service time process is stationary with*

(a) $E_2\{x_1\} < \infty$ *, then*

$$\tau^{-1} \int_0^\tau \phi(s)ds \to E_1 \times E_2\{\phi(1)|T\} \ , \quad P_1 \times P_2 \ - a.s.$$

where, also,

$$E_1 \times E_2\{\phi(1)|T\} = E_1\{N(0,\ 1]|T_1\}E_2\{x_1|T_2\}\ ,\quad P_1 \times P_2\ \text{ - a.s.}$$

(b) $E_2\{(x_1 \log x_1)^+\} < \infty$ *, and* (4.4) *and* (4.5) *are satisfied,*

then

$$\tau^{-1}\int_0^\tau \phi(s)ds \to E_2\{x_1|T_2\}/E_1^0\left\{\Pi_1|T_1^0\right\}\ ,\quad P_1^0 \times P_2\ \text{ - a.s.,}$$

provided the right-hand side is $P_1^0 \times P_2$ *- a.s. defined.*

Proof. *(a)* Clearly the theorem will be proved if we can show

$$\tau^{-1}\left|\int_0^\tau \phi(s)ds - \sum_{i=1}^{N(0,\tau]} x_i\right| \to 0\ ,\quad P_1 \times P_2\ \text{ - a.s.} \tag{4.6}$$

We use the inequality (2.7). Clearly if $c(\tau) = o(\tau)$, which we assume, $\tau^{-1}\Phi_1'(\tau) \to 0$ a.s. Now

$$\limsup_{\tau\to\infty} \tau^{-1}\Phi_0'(\tau) \le \limsup_{k\to\infty}\ \limsup_{\tau\to\infty} \tau^{-1}\sum_{i=1}^{N(0,\tau]} \left(x_i-k\right)^+ .$$

But $(x-k)^+$ is a Borel-measurable function of x , for any $k \in Z_+$, and $E_2\{(x_1-k)^+\} \le E_2\{x_1\} < \infty$, and hence by the ergodic theorem

$$\lim_{\tau\to\infty} \tau^{-1}\sum_{i=1}^{N(0,\tau]} \left(x_i-k\right)^+ = E_1\{N(0,\ 1]|T_1\}.E_2\{(x_1-k)^+|T_2\} .$$

But $E_2\{(x_1-k)^+|T_2\} \to 0$ a.s. $(k \to \infty)$, since $(x_1-k)^+ \downarrow 0$ a.s. as $k \to \infty$ (Breiman (1968), Proposition 4.24). To complete the proof of *(a)*, we note only that since $P_1 \times P_2$ is stationary with respect to T_y' , by the ergodic theorem

$$\tau^{-1}\int_0^\tau \phi(s)ds \to E_1 \times E_2\{\phi(1)|T\}\ ,\quad P_1 \times P_2\ \text{ - a.s.,}$$

and

$$E_1 \times E_2\{E_1 \times E_2[\phi(1)|T]\} = E_1\{N(0,\ 1]\}.E_2\{x_1\} . \tag{4.7}$$

Now, since $\left|\Phi_1(\tau) - \sum_{i=1}^{N(0,\tau]} x_i\right| \to 0$ a.s.,

$$\liminf_{\tau\to\infty} \tau^{-1} \int_0^\tau \phi(s)ds \geq E_1\{N(0,\ 1]|T_1\}.E_2\{x_1|T_2\}$$

and, clearly,

$$E_1 \times E_2\{E_1[N(0,\ 1]|T_1].E_2[x_1|T_2]\} = E_1\{N(0,\ 1]\}.E_2\{x_1\}\ . \tag{4.8}$$

(4.7) and (4.8) are consistent only if $\Phi_2(\tau) \to 0$ a.s.

(b) Again we prove $\tau^{-1}\left|\Phi_1(\tau) - \sum_{i=1}^{N(0,\tau]} x_i\right| \to 0$ a.s. in exactly the same way. Now we show $\Phi_2(\tau) \to 0$ a.s. Let F denote the distribution function of the service times, i.e., $F(l) = P_2\{x_i \leq l\}$. Let $r_k \equiv [c^{k-1}]$ for some fixed $c > 1$. For a given $\varepsilon > 0$, suppose $\Phi_2(\tau) \geq \varepsilon\tau$ for some τ . Let r_k be the first term of the geometric subsequence larger than τ . Then $\Phi_2(r_k) \geq \varepsilon r_{k-1}$. Hence we need only prove $\Phi_2(\tau) \to 0$ a.s. for a geometric subsequence. But

$$\begin{aligned}
&P_1^0 \times P_2\left\{r_k^{-1}\Phi_2(r_k) \geq \varepsilon\right\} \\
&\quad \leq E_1^0 \times E_2\{\Phi(r_k)\}/\varepsilon r_k \\
&\quad = \int_0^\infty dU(u) \int_0^\infty \max[\min(l-u,\ r_k),\ 0]\, dF(l)/\varepsilon r_k \\
&\quad = (\varepsilon r_k)^{-1} \int_{r_k}^\infty dF(l) \int_{l-r_k}^{l} (l-u)dU(u) + \varepsilon^{-1} \int_{r_k}^\infty dF(l) \int_0^{l-r_k} dU(u) \\
&\qquad + (\varepsilon r_k)^{-1} \int_0^{r_k} dF(l) \int_0^{l} (l-u)dU(u)\ ,
\end{aligned}$$

so that

$$\varepsilon \sum_{k=1}^{\infty} P_1^0 \times P_2\left\{r_k^{-1}\Phi_2(r_k) \geq \varepsilon\right\}$$

$$\leq 2B(\log c)^{-1} \int_1^{\infty} l \log l dF(l) + (\log c)^{-1} \int_1^{\infty} \log l dF(l)$$

$$+ \left[2 \sum_{k=1}^{\infty} r_k^{-1} + B\left(2c^2+2\right)\right] \int_0^{\infty} l dF(l) + \left(2c^2+1\right) < \infty ,$$

where we have used $U(l) \leq Bl + 1$ for some constant B, bounded the sums by integrals and exchanged the order of integration. Hence $\Phi_2(\tau) \to 0$ a.s. by the Borel-Cantelli lemma. The proof is now completed by observing that

$$t_i/i \xrightarrow{\text{a.s.}} E_1^0\left\{\Pi_1 | T_1^0\right\} ,$$

and employing $\{N(0, y] < j\} = \left\{t_j > y\right\}$ and the monotonicity of $N(0, \tau]$. □

We would anticipate that Theorem 2.4.1 *(b)* is sub-optimal: $E_2\{x_1\} < \infty$ should suffice (see Problem 6.3.2). We deduce immediately from Theorem 2.4.1 the strong law of large numbers.

COROLLARY 2.4.2. *If the arrival process is stationary with* $m < \infty$ *(and* (4.4) *and* (4.5) *hold), and satisfies* $N(0, n]/n \to m$, $P_1\left(P_1^0\right)$ *- a.s., and if the service time process is stationary with* $\sum_{i=1}^{n} x_i/n \to E_2\{x_1\} < \infty$, P_2 *- a.s.,* (*and* $E_2\{(x_1 \log x_1)^+\} < \infty$), *then*

$$\tau^{-1} \int_0^{\tau} \phi(s)ds \to mE_2\{x_1\} , \quad P_1 \times P_2 \ \text{- a.s.} \quad \left(P_1^0 \times P_2 \ \text{- a.s.}\right). \qquad \square$$

We should compare Corollary 2.4.2 with Theorem 2.3.1 *(b)* and the comments following Theorem 2.3.1: it would appear (although we have no example) that the strong law of large numbers could hold for $\phi(s)$ without the process being ergodic - simple examples of this behaviour in other contexts are easy to construct (Breiman (1968), pp. 110 and 113, or Hannan (1973), p. 163).

Functional strong laws can be proved for this process in the manner of Iglehart (1971b).

Because queues with bulk arrivals are given much attention in the literature, we wish to make a few points in this context about the Palm measure corresponding to a non-orderly stationary point process. Slivnyak in fact goes further than indicated in the opening paragraphs of this section.

Consider the set $R_{\uparrow}^{Z}$ of subsets $\mathbf{t}'$ of R without finite limit points satisfying

$$\ldots \leq t'_{-1} \leq t'_{0} = 0 \leq t'_{1} \leq \ldots .$$

Note that if, for a given $\mathbf{t}'$, there are k points at the origin, then there will be $(k-1)$ other members of $R_{\uparrow}^{Z}$, which will constitute only a relabelling of $\mathbf{t}'$. Let $\sigma(\uparrow)$ denote the minimal σ-algebra generated by sets of the form $\{\mathbf{t}' : t'_k \leq y\}$, $y \in R$, $k \in Z$.

Slivnyak then proves that uniquely coupled with a stationary point process P of finite intensity m is a probability measure Pr^0 on $\left(R_{\uparrow}^{Z}, \sigma(\uparrow)\right)$ such that $\{\Pi'_j\} \equiv \{t'_{j+1}-t'_j\}$ is stationary, i.e., if we define S_k as in (3.2),

$$Pr^0\left(S_k^{-1}D\right) = Pr^0(D) \ , \quad k \in Z \ ,$$

where $D \in \sigma_{\text{gen.}}\left\{\{\Pi'_j \in B\},\, B \in \mathcal{B}(R),\, j \in Z\right\} \subset \sigma(\uparrow)$. Note that a mapping $\eta : R_{\uparrow}^{Z} \to N$ may be defined analogously to (2.4) of Chapter 1, and will be measurable, e.g. for $u, v > 0$, $k \in Z_+$,

$$\{\mathbf{t}' : \eta(u, u+v] \geq k\} = \bigcup_{i=1}^{\infty} \left\{\mathbf{t}' : t'_i \leq v,\ t'_{i+1} > u,\ t'_{i+k} \leq u+v\right\} .$$

Therefore Pr^0 induces a probability measure $P^0 = Pr^0\eta^{-1}$ on $\left(N, \sigma(N)\right)$ such that if $N_0 = \{N \in N : N(\{0\}) > 0\}$, $P^0\left(N_0\right) = 1$. However, as is readily seen, several $\mathbf{t}'$'s may map to one N, so that an inverse mapping is undefined, i.e., it is no longer meaningful to refer to

$t_j(N) = t_j'(\mathbf{t}')$ in an obvious notation, unless P is a.s. orderly. However, had we adopted Slivnyak's approach, our theorems would hold, with some modifications, for non-orderly processes too.

Let Pr_1^0 on $R_\uparrow^Z$ correspond to a given stationary arrival process. Despite the aforementioned difficulties, it is natural to identify the $\mathbf{t}'$ with arrival epochs. Hence define $\phi' : R_\uparrow^Z \times R_+^Z \to \overline{Z}_+^R$ in a way similar to (2.1). The quantity naturally thrown up from $\int_0^\tau \phi'(s)ds$ corresponding to $\sum_{j=1}^{N(0,\tau]} x_j$ is $\sum_{t_j' \in (0,\tau]} x_j$, an expression not in a suitable form for application of random change of time lemmas. However,

$$\sum_{j=\eta(\{0\})}^{\eta(0,\tau]} x_j \le \sum_{t_j' \in (0,\tau]} x_j \le \sum_{j=1}^{\eta[0,\tau]} x_j ,$$

and by showing (use geometric subsequences, (4.9), and Chebyshev's inequality) that for any $\delta > 0$,

$$\tau^{-\delta}\left|\sum_{j=1}^{\eta[0,\tau]} x_j - \sum_{j=\eta(\{0\})}^{\eta(0,\tau]} x_j\right| \to 0 , \quad Pr_1^0 \times P_2 \quad \text{- a.s.,} \quad (\tau \to \infty) ,$$

we could employ the same techniques as for the orderly case to establish the same results. However, we do not consider the notational inconvenience is worth the extra generality.

We would also require generalisations of (4.4) and (4.5). Recall that a finite intensity and stationarity imply that the finite rate λ exists (Chapter 1, (3.20)). Again, Kaplan shows that if

$$U(x) \equiv E^0\eta(0, x] \equiv E_1^0 N(0, x] < \infty$$

for a finite x , then $U(x)$ is finite for all x , and

$$\sup_{x>0} \{U(x+v)-U(x)\} \le 2U(v) + m/\lambda , \tag{4.9}$$

whence $U(x) = O(x)$ $(x \to \infty)$. Our theorems now go through as before.

2.5. The central limit theorem

Before we prove the central limit theorem, we need the following simple lemma (which is possibly known, but we could locate no statement or proof). The lemma may also be used to weaken Theorem 3 of Daley (1972).

LEMMA 2.5.1. *Let* $\{Y_n\}$, $\{Z_n\}$ *and* $\{N_n\}$ *be sequences of random variables defined on the same fixed probability space such that* $Y_n \xrightarrow{\mathcal{D}} Y$ *and* $Z_n \xrightarrow{\mathcal{D}} Z$ *for some random variables* Y *and* Z *defined on (not necessarily) another probability space. If* $\{N_n\}$ *consists of non-negative integer-valued random variables such that* $N_n \xrightarrow{P} \infty$ *and* $\{(N_n, Z_n)\}$ *is independent of* $\{Y_n\}$ *, then*

$$(Y_{N_n}, Z_n) \xrightarrow{\mathcal{D}} (Y, Z) .$$

Proof. Let (y, z) be a continuity point of (Y, Z) . Then clearly y is a continuity point of Y , and z a continuity point of Z . Hence, for any arbitrary $\varepsilon > 0$ and for $k \geq k_0(y, \varepsilon) \equiv k_0$,

$$|Pr\{Y_k \leq y\} - Pr\{Y \leq y\}| \leq \varepsilon .$$

(For the purposes of this proof we assume all random variables on a single space (Ω, F, Pr)). Thus

$$Pr\{Y_{N_n} \leq y, Z_n \leq z\} \leq Pr\{N_n < k_0\} + \sum_{k=k_0}^{\infty} Pr\{Y_k \leq y\} Pr\{Z_n \leq z, N_n = k\}$$

$$\leq Pr\{N_n < k_0\} + Pr\{Y \leq y\} Pr\{Z_n \leq z\} + \varepsilon ,$$

i.e., $\limsup_{n\to\infty} Pr\{Y_{N_n} \leq y, Z_n \leq z\} \leq Pr\{Y \leq y, Z \leq z\}$.

Also,

$$Pr\{Y_{N_n} \le y,\ Z_n \le z\}$$

$$\ge \sum_{k=k_0}^{\infty} Pr\{Y_k \le y\} Pr\{Z_n \le z,\ N_n = k\}$$

$$\ge Pr\{Y \le y\} Pr\{Z_n \le z\} + Pr\{Y \le y\} Pr\{N_n \ge k_0\}$$

$$- Pr\{Y \le y\} Pr\{(N_n \ge k_0) \cup (Z_n \le z)\} - \varepsilon$$

$$\ge Pr\{Y \le y\} Pr\{Z_n \le z\} + Pr\{Y \le y\} Pr\{N_n \ge k_0\} - Pr\{Y \le y\} - \varepsilon .$$

Hence

$$\liminf_{n\to\infty} Pr\{Y_{N_n} \le y,\ Z_n \le z\} \ge Pr\{Y \le y,\ Z \le z\} ,$$

i.e., $\lim_{n\to\infty} Pr\{Y_{N_n} \le y,\ Z_n \le z\} = Pr\{Y \le y,\ Z \le z\}$. □

We also require the following lemma, which, conforming to a general theme of this chapter (and this thesis), proves that the difference between the quantity we require (the traffic time average) and a quantity which is easy to handle $\left(\sum_{i=1}^{N(0,\tau]} x_i\right)$ is small under suitable circumstances.

LEMMA 2.5.2. *If the arrival process is stationary with* $m < \infty$ *(and* (4.4) *and* (4.5) *hold), and the sequence of service times is stationary with finite* $(2-\delta+\upsilon)$*-th moment, any* $\upsilon > 0$, $0 < \delta \le 2$, *or under any circumstances if* $\delta > 2$, *then*

$$\tau^{-\delta} \left| \int_0^\tau \phi(s)ds - \sum_{i=1}^{N(0,\tau]} x_i \right| \xrightarrow{P_1 \times P_2 (P_1^0 \times P_2)} 0 , \quad (\tau \to \infty) .$$

Proof. We confine our proof to the $P_1 \times P_2$ case. Arguments for the $P_1^0 \times P_2$ case use techniques similar to those in Theorem 2.4.1 *(b)*, and account has to be taken of arrivals at 0 , but no new ideas are involved. Now by (2.2) and (2.3),

$$\left| \int_0^\tau \phi(s)ds - \sum_{i=1}^{N(0,\tau]} x_i \right| = \Phi_2(\tau) + \Phi'(\tau) .$$

For any $\varepsilon > 0$,

$$P_1 \times P_2\left\{\tau^{-\delta}\Phi_2(\tau) \geq \varepsilon\right\} \leq E_1 \times E_2\{\Phi_2(\tau)\}/\varepsilon\tau^{\delta}$$

$$= \varepsilon^{-1}\tau^{-\delta} \int_0^\infty dF(l) \int_0^l \min(l-u,\ \tau)du$$

$$\leq \varepsilon^{-1}\tau^{-\upsilon} \int_0^\tau l^{2-\delta+\upsilon}dF(l) + \varepsilon^{-1} \int_\tau^\infty l^{2-\delta}dF(l) \quad (5.1)$$

where $F(l) \equiv P_2\{\mathrm{x} : x_i \leq l\}$ is the service time distribution (independent of i since P_2 is stationary). Hence (5.1) $\to 0$, $(\tau \to \infty)$. Similarly, Chebyshev's inequality yields exactly the same upper bound for $P_1 \times P_2\{\tau^{-\delta}\Phi'(\tau) \geq \varepsilon\}$. □

This leads to a simple proof of the central limit theorem. We denote convergence in distribution by $\mathcal{D}_1$, $\mathcal{D}_1^0$, $\mathcal{D}_2$, $\mathcal{D}_1 \times \mathcal{D}_2$, $\mathcal{D}_1^0 \times \mathcal{D}_2$ in an obvious notation.

THEOREM 2.5.3. *If the arrival process is stationary with* $m < \infty$ (*and* (4.4) *and* (4.5) *hold*), *and satisfies*

$$\tau^{-\frac{1}{2}}\{N(0,\ \tau]-m\tau\} \xrightarrow{\mathcal{D}_1(\mathcal{D}_1^0)} X_1 , \quad (5.2)$$

where X_1 *is some random variable, and the service times form a stationary sequence with finite* $(3/2 + \delta)th$ *moment, and satisfy*

$$n^{-\frac{1}{2}}\left\{\sum_{j=1}^n x_j - nE_2\{x_1\}\right\} \xrightarrow{\mathcal{D}_2} X_2 , \quad (5.3)$$

where X_2 *is some other random variable defined on the same probability space as* X_1, *then*

$$\tau^{-\frac{1}{2}}\left\{\int_0^\tau \phi(s)ds - mE_2\{x_1\}\tau\right\} \xrightarrow{\mathcal{D}_1\times\mathcal{D}_2\left(\mathcal{D}_1^0\times\mathcal{D}_2\right)} E_2\{x_1\}X_1 + m^{\frac{1}{2}}X_2 . \quad (5.4)$$

Proof. We will only prove the $\mathcal{D}_1 \times \mathcal{D}_2$ convergence: $\mathcal{D}_1^0 \times \mathcal{D}_2$ only requires writing $\mathcal{D}_1^0$ for $\mathcal{D}_1$ and P_1^0 for P_1 . All of the random variable in the following indexed by τ are on the space $N \times R_+^Z$, so Lemma 2.5.1 may be applied. For this proof, set $\mu = E_2\{x_1\}$, and $N_\tau = N(0, \tau]$. According to Lemma 2.5.2,

$$\tau^{-\frac{1}{2}}\left|\left\{\int_0^\tau \phi(s)ds - m\mu\tau\right\} - \left\{\sum_{i=1}^{N_\tau} x_i - m\mu\tau\right\}\right| = \tau^{-\frac{1}{2}}\left|\int_0^\tau \phi(s)ds - \sum_{i=1}^{N_\tau} x_i\right| \xrightarrow{P_1 \times P_2} 0 .$$

But

$$\tau^{-\frac{1}{2}}\left\{\sum_{i=1}^{N_\tau} x_i - m\mu\tau\right\} = \tau^{-\frac{1}{2}} \sum_{i=1}^{N_\tau} (x_i - \mu) + \tau^{-\frac{1}{2}}\mu\{N_\tau - m\tau\} . \tag{5.5}$$

By (5.2), $(N_\tau+1)/m\tau \xrightarrow{P_1} 1$, and hence by Theorem 4.4 of Billingsley (1968), and the continuous mapping theorem (Section 1.5)

$$(N_\tau - m\tau)/(N_\tau+1)^{\frac{1}{2}} \xrightarrow{\mathcal{D}_1} X_1/\sqrt{m} .$$

In Lemma 2.5.1 (or a continuous version of it), put

$$Z_\tau \equiv \mu(N_\tau - m\tau)/(N_\tau+1)^{\frac{1}{2}} , \quad Y_\tau \equiv \sum_{i=1}^{[\tau]} (x_i - \mu)/([\tau]+1)^{\frac{1}{2}} ,$$

where $[\tau]$ is the integer part of τ, and $N_\tau \equiv N(0, \tau]$; then, applying simultaneously the continuous mapping theorem,

$$(N_\tau+1)^{-\frac{1}{2}}\left\{\sum_{i=1}^{N_\tau} (x_i - \mu) + \mu(N_\tau - m\tau)\right\} \xrightarrow{\mathcal{D}_1 \times \mathcal{D}_2} X_2 + \mu X_1/\sqrt{m} .$$

Again using Theorem 4.4 of Billingsley (1968) and the continuous mapping theorem,

$$(5.5) \xrightarrow{\mathcal{D}_1 \times \mathcal{D}_2} \sqrt{m}X_2 + \mu X_1 . \qquad \square$$

In Theorem 2.5.3, we have avoided the question of norming (see Problem 6.3.3)). Hence, writing $\mu = E_2\{x_1\}$ again,

LEMMA 2.5.4. *If the arrival process is stationary, and the process of service times is stationary with* $E_2\left\{x_i^2\right\} < \infty$, *and*

$$\sum_{i=1}^{\infty} |\rho_i| < \infty , \tag{5.6}$$

where $\rho_i = E_2\{(x_0-\mu)(x_i-\mu)\}$, *and*

(a) $\mathrm{Var}_{P_1} N(0, u] \sim \lambda_1 u$, *then*

$$\mathrm{Var}_{P_1 \times P_2}\left\{\sum_{i=1}^{N(0,\tau]} x_i\right\} \sim \left(m \ \mathrm{Var}_{P_2}(x_1)+m\rho+\lambda_1\mu^2\right)\tau , \quad (\tau \to \infty) ,$$

where $\rho \equiv 2 \sum_{i=1}^{\infty} \rho_i$.

(b) $\mathrm{Var}_{P_1^0} N(0, u] \sim \lambda_2 u$, *and* $U(\tau)/\tau \to \lambda' < \infty$, *then*

$$\mathrm{Var}_{P_1^0 \times P_2}\left\{\sum_{i=1}^{N(0,\tau]} x_i\right\} \sim \left(\lambda' \ \mathrm{Var}_{P_2}(x_1)+\lambda'\rho+\lambda_2\mu^2\right)\tau , \quad (\tau \to \infty) .$$

Remark. If the arrival process is orderly, then it is known (Daley (1971)) that $U(\tau)/\tau \to \lambda' < \infty$.

Proof. Again we prove only *(a)*. Write $N_\tau \equiv N(0, \tau]$. Easily

$$\mathrm{Var}\left\{\sum_{i=1}^{N_\tau} x_i\right\} = m \ \mathrm{Var}\{x_1\}\tau + \mu^2 \ \mathrm{Var}\ N_\tau + 2E_1\left\{\sum_{j=1}^{N_\tau} (N_\tau-j)\rho_j\right\} .$$

Clearly, by (5.6),

$$\left|E_1 \sum_{j=1}^{N_\tau} (N_\tau-j)\rho_j\right| \le m \sum_{j=1}^{\infty} |\rho_j| < \infty .$$

Now,

$$\left|E_1\left\{\sum_{j=1}^{N_\tau} \tau^{-1}(N_\tau-j)\rho_j\right\} - m \sum_{j=1}^{\infty} \rho_j\right|$$

$$\le \left|E_1\left\{\left(\tau^{-1}N_\tau-m\right) \sum_{j=1}^{N_\tau} \rho_j\right\}\right| + \tau^{-1}E_1\left\{\sum_{j=1}^{N_\tau} j|\rho_j|\right\} + mE_1\left\{\sum_{j=N_\tau+1}^{\infty} |\rho_j|\right\} . \tag{5.7}$$

By dominated convergence and the fact that $N_\tau \to \infty$ a.s. as $\tau \to \infty$, the final term of (5.7) $\to 0$ $(\tau \to \infty)$. The second term has an upper bound of

$$\tau^{-1}E_1\left\{\sum_{i=1}^{N_\tau}\sum_{j=i}^{\infty}|\rho_j|\right\}$$

$$= \tau^{-1}\int_{\{N_\tau>k\}}\sum_{i=1}^{N_\tau}\sum_{j=i}^{\infty}|\rho_j|dP_1 + \tau^{-1}\int_{\{N_\tau\le k\}}\sum_{i=1}^{N_\tau}\sum_{j=i}^{\infty}|\rho_j|dP_1 \ . \tag{5.8}$$

For an arbitrary $\varepsilon > 0$, let $k \equiv k(\varepsilon)$ be such that $\sum_{j=k}^{\infty}|\rho_j| < \varepsilon$.

Then clearly the second term of (5.8) has an upper bound of

$$k\sum_{j=1}^{\infty}|\rho_j|P_1\{N_\tau \le k\}/\tau \quad \big(\to 0,\ (\tau\to\infty)\big) \ ,$$

and the first term is bounded by

$$\sum_{i=1}^{k-1}\sum_{j=i}^{\infty}|\rho_j|P_1\{N_\tau > k\}/\tau \qquad\qquad \big(\to 0,\ (\tau\to\infty)\big)$$

$$+ \ \varepsilon\tau^{-1}\int_{\{N_\tau>k\}}N_\tau dP_1 \quad \big(\le m\varepsilon,\ (\tau\to\infty)\big) \ .$$

A similar style of argument applied to the first term of (5.8) yields an asymptotic upper bound of $2m\varepsilon$ for this term: here we need the fact that

$$\int_{\{N_\tau>k\}}\big(N_\tau/\tau-\mu\big)dP_1 = -\int_{\{N_\tau\le k\}}\big(N_\tau/\tau-\mu\big)dP_1 \ .$$

Hence we conclude that (5.7) $\to 0$ $(\tau\to\infty)$. □

This central limit theorem has been proved here under second moment conditions: however, the functional case (Iglehart and Kennedy (1970)) seems to require that the service times have finite $(2+\delta)$-th moment. The generalisation to non-stationary arrival processes seems difficult (see Problem 6.3.4).

2.6. Rates of convergence

Let X be a random variable on a space $\big(\Omega_1, F_1, Pr_1\big)$, and Y a random variable on another (not necessarily distinct) space $\big(\Omega_2, F_2, Pr_2\big)$.

Then we will define the supremum metric for these two random variables by

$$\nu(X, Y) \equiv \sup_{-\infty<x<\infty} |Pr_1\{X \le x\} - Pr_2\{Y \le x\}| \ .$$

Then we say that $X_n \xrightarrow{\mathcal{D}} X$ at rate (at least) $c_n \downarrow 0$ $(n \to \infty)$ if $\nu(X_n, X) \le c_n$. Note that X, $\{X_n\}$ do not have to be on the same space, and hence, in searching for rates of convergence results for the traffic time average for the $G/G/\infty$ queue, we can (and will) retain the same notation as in the previous sections. We will require the following result, which is proved as Lemma 1 of Tomko (1972).

LEMMA 2.6.1. *Let* $X_t = Y_t + \delta_t$, $t \in R$, *be the sum of two random variables, and* N *be a random variable with the normal distribution. If* $\delta_t \to 0$ *in probability as* $t \to \infty$ *and* $Pr\{|\delta_t| \ge \varepsilon_t\} \le \beta_t$, *then*

$$\nu(X_t, \mathrm{N}) \le \nu(Y_t, \mathrm{N}) + \varepsilon_t + \beta_t \ . \qquad \square$$

It is clear that a result similar to Lemma 2.5.1 may be obtained: however, our metric ν, unlike some closely related metrics (discussed later) is not continuous under suitably smooth mappings (Whitt (1974b)) and hence we state more specifically:

LEMMA 2.6.2. *Let* $\{Y_\tau\}$, $\{Z_\tau\}$, $\{N_\tau\}$ *be sequences of random variables on the same space* (Ω, F, Pr) *such that* $\nu(Y_\tau, Y) \le \psi(\tau)$, $\nu(Z_\tau, Z) \le \varphi(\tau)$ *where* Y, Z *are continuous and are on a common space (not necessarily* Ω *). If* $N_\tau \xrightarrow{P} +\infty$ *and* $\{(N_\tau, Z_\tau)\}$ *is independent of* $\{Y_\tau\}$, *then*

$$\nu(Y_{N_\tau} + Z_\tau, Y+Z) \le \varphi(\tau) + \int_0^\infty \psi(x) dPr\{N_\tau \le x\} \ .$$

Proof. Omitted. Similar to the proof of Lemma 2.6.12. $\square$

In our applications of this lemma, we will suppose always that Y and Z have a standard normal distribution. As such it can be supplemented by (denoting standard normal random variables by N)

LEMMA 2.6.3. *Suppose* $(N_\tau - \kappa\tau)/\sigma\sqrt{\tau} \xrightarrow{\mathcal{D}} \mathrm{N}_1$ *at rate* $\beta(\tau)$, *where* N_τ

is a positive random variable for each $\tau > 0$ *, and* σ, κ *are some constants, and, not necessarily independently,* $X_\tau/\sqrt{\tau} \xrightarrow{\mathcal{D}} N_2$ *at rate* $\alpha(\tau)$ *; then*

$$Pr\left\{\left|X_\tau/\left(\kappa^{-1}N_\tau\right)^{\frac{1}{2}} - X_\tau/\tau^{\frac{1}{2}}\right| \geq 2\kappa(\log\tau)/\sigma\sqrt{\tau}\right\} \leq A_1 \max\left(\alpha(\tau), \beta(\tau), 1/\sqrt{\tau}\right)$$

for some constant A_1 .

Proof. Let $\Phi(x) \equiv \int_{-\infty}^{x} e^{-y^2/2}dy/\sqrt{2\pi}$; then, for some $y_\tau > 0$,

$$Pr\left\{\left|X_\tau/\left(\kappa^{-1}N_\tau\right)^{\frac{1}{2}}-X_\tau/\tau^{\frac{1}{2}}\right| \geq \varepsilon_\tau\right\}$$

$$\leq Pr\left\{\left|X_\tau/\tau^{\frac{1}{2}}\right| > y_\tau\right\} + Pr\left\{\left|\left(\tau/\kappa^{-1}N_\tau\right)^{\frac{1}{2}}-1\right| \geq \varepsilon_\tau/y_\tau\right\}$$

$$\leq 2\alpha(\tau) + 2\left(1-\Phi\left(y_\tau\right)\right) + Pr\left\{\left|N_\tau-\kappa\tau\right|/\sigma\tau \geq \kappa\varepsilon_\tau/2\sigma y_\tau\right\}$$

$$\leq 2\{\alpha(\tau)+\beta(\tau)\} + 2\left\{1-\Phi\left(y_\tau\right)\right\} + 2\left\{1-\Phi\left(\kappa\varepsilon_\tau\sqrt{\tau}/2\sigma y_\tau\right)\right\} .$$

To optimise the rate of convergence, we choose $y_\tau = \kappa\varepsilon_\tau\sqrt{\tau}/2\sigma y_\tau$.

According to Feller (1968), p. 175, $1 - \Phi(x) < e^{-\frac{1}{2}x^2}\left[(2\pi)^{\frac{1}{2}}x\right]^{-1}$, $x > 0$. Hence, since we will require ε_τ and $1 - \Phi\left(y_\tau\right)$ to converge to zero at about the same rate, we choose $y_\tau = (\log\tau)^{\frac{1}{2}}$, and obtain

$$1 - \Phi\left(y_\tau\right) < (2\pi\tau)^{-\frac{1}{2}}(\log\tau)^{-1}$$

and $\varepsilon_\tau = 2\kappa(\log\tau)/\sigma\sqrt{\tau}$. Note that we really only require y_τ such that $e^{-\frac{1}{2}y_\tau^2}/y_\tau = 1/\sqrt{\tau}$, and finer solutions are possible, e.g. $y_\tau = [\log(\tau/\log\tau)]^{\frac{1}{2}}$. □

We will require a statement of the rate at which $\tau^{-\frac{1}{2}}\int_0^\tau \phi(s)ds$ and $\sum_{i=1}^{N(0,\tau]} x_i$ approach each other. Since we will make several such estimates

(Lemmas 2.6.7, 2.6.8 and 2.6.10), we propose

DEFINITION 2.6.4. Suppose

$$P_1 \times P_2\left\{\tau^{-\frac{1}{2}}\left|\int_0^\tau \phi(s)ds - \sum_{i=1}^{N(0,\tau]} x_i\right| \geq \varepsilon_\tau\right\} \leq \beta_\tau \ ;$$

then we set $\chi(\tau) = \max(\varepsilon_\tau, \beta_\tau)$.

With this definition we have a simple estimate of rate of convergence to normality of the 'number of busy servers' process.

THEOREM 2.6.5. *Suppose there are constants* m, μ, σ_1, σ_2 *such that the arrival process of the* $G/G/\infty$ *queue satisfies* $(N(0, \tau]-m\tau)/\sigma_1\sqrt{\tau} \xrightarrow{\mathcal{D}_1} \mathrm{N}_1$ *at rate* $\varphi(\tau)$ *, and* $\sum_{i=1}^{n} (x_i-\mu)/\sigma_2\sqrt{n} \xrightarrow{\mathcal{D}_2} \mathrm{N}_2$ *at rate* $\psi(\tau)$ *, then*

$$\left(\int_0^\tau \phi(s)ds-m\mu\tau\right)/\sigma\sqrt{\tau} \xrightarrow{\mathcal{D}_1\times\mathcal{D}_2} \mathrm{N}$$

where $\mathrm{N} \equiv (\mu\sigma_1\mathrm{N}_1+\sqrt{m}\sigma_2\mathrm{N}_2)/\sigma$ *at rate*

$$A_2 \max\left[\varphi(\tau), \sum_{k=0}^{\infty} \psi(k)P_1\{N(0, \tau] = k\}, \chi(\tau), (\log \tau)/\sqrt{\tau}\right] \qquad (6.1)$$

for some constant A_2 *, where* $\sigma^2 = \mu^2\sigma_1^2 + m\sigma_2^2$.

Proof. Write $N_\tau = N(0, \tau]$. A's will denote constants. We firstly note $P_1\left\{(N_\tau+1)^{\frac{1}{2}}/N_\tau^{\frac{1}{2}} \geq \tau^{-\frac{1}{2}}\right\} \leq A_3\tau^{-\frac{1}{2}}$ for some constant A_3 . By Lemmas 2.6.3 and 2.6.1,

$$\nu\left(\sqrt{m}(N_\tau-m\tau)/\sigma_1(N_\tau+1)^{\frac{1}{2}}, \mathrm{N}_1\right) \leq A_4 \max(\varphi(\tau), (\log \tau)/\sqrt{\tau}) \ ;$$

hence, by Lemma 2.6.2,

$$\nu\left[\mu\sigma_1\sigma^{-1}\sqrt{m}(N_\tau-m\tau)/\sigma_1(N_\tau+1)^{\frac{1}{2}}+\sqrt{m}\sigma_2\sigma^{-1}\sum_{i=1}^{N_\tau} (x_i-\mu)/\sigma_2(N_\tau+1)^{\frac{1}{2}}, \mathrm{N}\right]$$

$$\leq A_5 \max\left[\varphi(\tau), \sum_{k=0}^{\infty} \psi(k)P_1\{N(0, \tau] = k\}, (\log \tau)/\sqrt{\tau}\right] .$$

where we have defined $\mathrm{N} \equiv \left(\mu\sigma_1\mathrm{N}_1+\sqrt{m}\sigma_2\mathrm{N}_2\right)/\sigma$. Thus we obtain by Lemma 2.6.3,

$$\nu\left[\left(\sum_{i=1}^{N_\tau} x_i - m\mu\tau\right)/\sigma\sqrt{\tau},\ \mathrm{N}\right] \le A_6 \max\left[\varphi(\tau),\ \sum_{k=0}^{\infty} \psi(k)P_1\{N(0,\ \tau] = k\},\ (\log\tau)/\sqrt{\tau}\right] .$$

By Definition 2.6.4 and Lemma 2.6.1 the conclusion follows. □

A valid enquiry is into the technical handling of the second term in (6.1). Since many rates of convergence results have a $\psi(k)$ of the form $\psi(k) = \{\log(1+k)\}^\gamma/(1+k)^\delta$, $\gamma \ge 0$, $\delta > 0$ we prove

LEMMA 2.6.6. *(a) If* $\psi(k) = \{\log(1+k)\}^\gamma/(1+k)^\delta$ *, some* $\gamma \ge 0$, $\delta > 0$ *and* $E_1\{N(0,\ n]\} = O(n)$ *, then*

$$E_1\{\psi(N(0,\ n])\} = O\left(\max[\{\log(1+n)\}^\gamma\varphi(n),\ \psi(n)]\right) .$$

(b) If also, $\limsup_{n\to\infty} E_1|N(0,\ n]-mn|^{1+\delta}/n < \infty$ *, then* $E_1\{\psi(N(0,\ n])\} = O(\psi(n))$.

Remark. If $0 < \delta \le 1$, condition *(b)* holds if $\limsup_{n\to\infty} \operatorname{Var} N(0,\ n]/n < \infty$, true for many point processes.

Proof. Write $N_n = N(0,\ n]$, $C_n \equiv \{N(0,\ n] \ge nm/2\}$. Then

$$\int_{C_n} \frac{\{\log(1+N_n)\}^\gamma}{(1+N_n)^\delta}\, dP_1 \le (2/nm)^\delta E_1\{\log(1+N_n)\}^\gamma$$

$$\le A_7\{\log(1+n)\}^\gamma/n^\delta$$

by Jensen's inequality, since $\{\log(1+x)\}^\gamma$ is a concave function. Also to prove *(a)*,

$$\int_{C_n^c} \frac{\{\log(1+N_n)\}^\gamma}{(1+N_n)^\delta}\, dP_1 \le A_7'\{\log(1+n)\}^\gamma P_1(C_n^c)$$

$$\le A_7'\{\log(1+n)\}^\gamma P_1\{|N_n/n-m| > m/2\} \tag{6.2}$$

$$\le A_7'(\{\log(1+n)\}^\gamma/n^\delta)\{n^\delta[1-\Phi(\sqrt{n}/2\sigma)]\} + \{\log(1+n)\}^\gamma\varphi(n)$$

as in the proof of Lemma 2.6.2. But

$$n^\delta\{1-\Phi(\sqrt{n}/2\sigma)\} \le (2\sigma/\sqrt{2\pi})n^{\delta-\frac{1}{2}}e^{-n/2\sigma^2} \to 0 \quad (n \to \infty)\ .$$

(b) follows by applying Chebyshev's inequality to (6.2). □

We now turn to estimation of $\chi(\tau)$. Our first result follows easily from Chebyshev's inequality: since all of our results seem to go through most naturally with stationary arrival processes, we make this assumption explicitly from now on.

LEMMA 2.6.7. *If the arrival process is stationary with* $m \equiv E_1N(0, 1] < \infty$ *and the sequence of service times is stationary with finite second moment, then* $\chi(\tau) = O(\tau^{-\frac{1}{4}})$. □

In our next two estimates we assume the service times to be i.i.d.

LEMMA 2.6.8. *If the arrival process is stationary and orderly with* $m \equiv E_1N(0, 1] < \infty$ *and* (4.4) *and* (4.5) *hold and the service times are i.i.d. with finite third moment, then* $\chi(\tau) = O(\tau^{-1/3})$.

Proof. Once again we use Chebyshev. Following Daley (1972), Theorem 4, we assert that from the stationarity and orderliness of P_1 ,

$$du|dU(v)| = P_1\{N(u, u+du] = 1, N(u+v, u+v+dv] = 1\} + o(dudv) \tag{6.3}$$

(where $m \equiv 1$), so that, using the decompositions (2.2) and (2.3),

$$E_1 \times E_2\{\Phi'(\tau)^2\} = \int_0^\tau du \int_0^\infty [(u+l-\tau)^+]^2 dF(l) \tag{6.4}$$

$$+ \int_0^\tau du \int_0^u dU(v) \int_0^\infty dF(l) \int_0^\infty (u+l-\tau)^+(u-v+m-\tau)^+ dF(m) \tag{6.5}$$

$$+ \int_0^\tau du \int_0^{\tau-u} dU(v) \int_0^\infty dF(l) \int_0^\infty (u+l-\tau)^+(v+m-\tau)^+ dF(m) \tag{6.6}$$

where F is the service time distribution function. Easy manipulation then yields $\int_0^\infty l^3 dF(l) < \infty$ as an upper bound for (6.4), whereas (6.5) and (6.6) have upper bounds of $A_8 \int_0^\infty l^3 dF(l) \int_0^\infty l dF(l)$. The term $E_1 \times E_2\left\{\Phi_2(\tau)^2\right\}$ has similar bounds. □

Unless the arrival process is renewal, the technique used in Lemma 2.6.8 cannot readily be extended to higher moments. However, we make

CONJECTURE 2.6.9. If the arrival process is stationary and orderly with $m \equiv E_1 N(0, 1] < \infty$ and (4.4) and (4.5) hold, and the service times are i.i.d. with finite qth moment, then $\chi(\tau) = O\left(\tau^{-\frac{1}{2}(q-1)/q}\right)$. □

By making some rough approximations, we can achieve a lower bound for $\chi(\tau)$. Here, as is often done, we ignore the contributions of arrivals before time zero to the traffic average.

LEMMA 2.6.10. *If the arrival process is stationary with finite p_1th moment, and the service times are i.i.d. with finite qth moment, then, setting $p_2 \equiv (\sqrt{1+q}-1)$, $p \equiv \min(p_1, p_2)$, $r = \min(p_1, 2)$,*

$$\chi(\tau) = O\left(\tau^{-\frac{1}{2}\left(1-pq^{-1}\right)p(p+1)^{-1}}\right), \quad p > 2,$$

$$= O\left(\tau^{-\frac{1}{2}\left(1-2q^{-1}\right)r(r+1)^{-1}}\right), \quad p \le 2.$$

Proof. We use the decomposition (2.4) and the inequality of von Bahr and Esseen (1965) and Dharmadhikari, Fabian and Jogdeo (1968) for

$\{X_n : n \geq 1\}$ a sequence of i.i.d. random variables with $EX_1 = \nu < \infty$: if $l \geq 1$, and $S_k = \sum_{i=1}^{k} X_i$, then

$$E|S_k - k\nu|^l \leq A_9\ k^{\max(1,\frac{1}{2}l)} E|X_1 - \nu|^l ,$$

for some constant A_9 . We also require that $M_n(u)^{1/n} \equiv \left\{E_1 N(0,\ u]^n\right\}^{1/n}$ is a sub-additive function of u for stationary arrival processes P_1 (easily deduced as in Daley (1971)), and hence that for some constant A_{10} , $E_1 N(0,\ u]^n \leq A_{10} u^n$.

Thus, by Chebyshev's inequality, with $\mu = E_2\{x_1\}$ as usual,

$$\begin{aligned} P_1 \times P_2\{\Phi_1'(\tau) \geq \varepsilon_\tau\} & \\ &\leq E_1 \times E_2\left\{\Phi_1'(\tau)^\delta\right\}/\varepsilon_\tau^\delta \\ &\leq 2^\delta\left\{E_1\{N(\tau - c(\tau),\ \tau]^\delta\}\mu^\delta + E_1\{N(\tau - c(\tau),\ \tau]^{\max(1,\frac{1}{2}\delta)}\}E_2|x_1 - \mu|^\delta\right\}/\varepsilon_\tau^\delta \\ &\leq A_{11}\ c(\tau)^\delta/\varepsilon_\tau^\delta . \end{aligned}$$

Also,

$$\begin{aligned} P_1 \times P_2\{\Phi_0(\tau) \geq \varepsilon_\tau\} &\leq A_{12}\tau^{\max(1,\frac{1}{2}\delta)} E_2\left\{\left|(x_1 - c(\tau))^+ - E_2(x_1 - c(\tau))^+\right|^\delta\right\}/\varepsilon_\tau^\delta \\ &\qquad + A_{13}\tau^\delta\left[E_2(x_1 - c(\tau))^+\right]^\delta/\varepsilon_\tau^\delta \\ &\leq A_{14}\left[\tau^{\max(1,\frac{1}{2}\delta)}/\varepsilon_\tau^\delta\right]E_2\left\{\left[(x_1 - c(\tau))^+\right]^\delta\right\} && (6.7) \\ &\qquad + A_{15}\left(\tau^\delta/\varepsilon_\tau^\delta\right)\left[E_2(x_1 - c(\tau))^+\right]^\delta . && (6.8) \end{aligned}$$

In (6.8), if we set $c(\tau) = \tau^\alpha$, $0 < \alpha \leq \frac{1}{2}$, then

$$\tau c(\tau)^{-1} E_2(x_1 - c(\tau))^+ \leq \tau c(\tau)^{-1} \int_{c(\tau)}^{\infty} l\, dF(l) \leq \int_{\tau^\alpha}^{\infty} l^{1/\alpha} dF(l) ,$$

and hence, provided $\int_0^\infty l^{1/\alpha} dF(l) < \infty$,

$$(6.8) \le A_{16}\left(\tau^{\alpha}/\varepsilon_{\tau}\right)^{\delta} .$$

Similarly,

$$(6.7) \le A_{17}\left(\tau^{\alpha}/\varepsilon_{\tau}\right)^{\delta}$$

provided $\int_0^{\infty} l^{(\delta/2\alpha)} dF(l) < \infty \quad (\delta \ge 2)$, or $\int_0^{\infty} l^{1/\alpha} dF(l) < \infty \quad (\delta < 2)$.

Hence

$$P_1 \times P_2\{\Phi'(\tau)/\sqrt{\tau} \ge \varepsilon_{\tau}\} \le A_{18}\left(\tau^{-(\frac{1}{2}-\alpha)}/\varepsilon_{\tau}\right)^{\delta} ,$$

i.e., $\chi(\tau) = \max\left[\varepsilon_{\tau}, \left(\tau^{-(\frac{1}{2}-\alpha)}/\varepsilon_{\tau}\right)^{\delta}\right]$. Choosing ε_{τ} optimally gives $\chi(\tau) = \tau^{-(\frac{1}{2}-\alpha)[\delta/(\delta+1)]}$.

Thus the problem becomes to maximize $f(\alpha, \delta) = (\frac{1}{2}-\alpha)[\delta/(\delta+1)]$ subject to the constraints $0 < \delta \le \min(p_1, q)$, $(1/\alpha) \le q$ and $(\delta/2\alpha) \le q$ (for $\delta \ge 2$). If $p_1 < 2$, the answer clearly is to choose $\alpha = 1/q$ and $\delta = q_1$. If $p_1 \ge 2$, the situation is not so simple. Drawing an (α, δ) graph including the constraint regions, and observing that $f(\alpha, \delta) \uparrow$, $\alpha \downarrow$ and $f(\alpha, \delta) \uparrow$, $\delta \uparrow$, one sees that the optimal (α, δ) pair lie on the line $\delta = 2\alpha q$, where $2 \le \delta \le \min(p_1, q)$. Thus the problem evolves into maximising $g(\delta) = \left(\frac{1}{2}-(\delta/2q)\right)\delta(\delta+1)^{-1}$ subject to this constraint. It is easily demonstrated that $g(\delta)$ has two turning points only, at $\delta = \pm\sqrt{q+1} - 1$. The one which concerns us, $\sqrt{q+1} - 1$, is always a maximum. Thus, writing $p_2 = \sqrt{q+1} - 1$, $p = \min(p_1, p_2)$, we find $\chi(\tau) = \tau^{-g(p)}$ unless $p_2 < 2$, i.e., $q < 8$ when it becomes $\chi(\tau) = \tau^{-g(2)}$. □

A point to note here is that as $p_1, q \to \infty$ (provided $p_1 \approx q$) $\chi(\tau) \to \tau^{-\frac{1}{2}}$, which one would hope for. Indeed, for large q , $\chi(\tau) \sim \tau^{-\frac{1}{2}(1-1/\sqrt{q})}$.

We focus attention now on $\varphi(\tau)$. For point processes specified in terms of counting measures, we will assume $\varphi(\tau)$ can be calculated; however, many point processes on the real line are more easily defined via their inter-arrival times, and for these we provide

LEMMA 2.6.11. *Suppose the inter-arrival times* $\{\Pi_i\}$ *of a point process satisfy, for some constants* λ *and* σ ,

$$\sup_y \left| P\left\{ \left(\sum_{i=0}^{n} (\Pi_i - \lambda) \right) / (\sigma\sqrt{n}) \le y \right\} - \Phi(y) \right| \le \varphi(n) \ ;$$

then, for any $\delta < 1$,

$$\sup_y |P\{ (N(0,\ \tau] - \lambda^{-1}\tau) / (\lambda^{-3/2}\sigma\sqrt{\tau}) \le y\} - \Phi(y)|$$

$$= O\left(\max[\varphi(\lambda^{-1}\delta\tau), (\log\tau)/\sqrt{\tau}]\right), \quad (6.9)$$

where Φ *is the cumulative distribution function of the normal distribution.*

Remark. For the purposes of this proof, we set $\Pi_0 = t_1$.

Proof. We divide the real line into four segments.

(i) $y > (\log\tau)^{\frac{1}{2}}$: Let $\gamma_\tau \equiv \lambda^{-1}\tau + \lambda^{-3/2}\sigma(\tau \log\tau)^{\frac{1}{2}}$, and $[x]$ denote the smallest integer less than x ; then

$$|P\{ (N(0,\ \tau] - \lambda^{-1}\tau) / (\lambda^{-3/2}\sigma\sqrt{\tau}) \le y\} - \Phi(y)| \quad (6.10)$$

$$= |P\{ (N(0,\ \tau] - \lambda^{-1}\tau) / (\lambda^{-3/2}\sigma\sqrt{\tau}) > y\} - [1 - \Phi(y)]|$$

$$\le P\{N(0,\ \tau] > \gamma_\tau\} + e^{-\frac{1}{2}\log\tau}/(2\pi \log\tau)^{\frac{1}{2}}$$

$$= P\left\{ \left(\sum_{i=0}^{[\gamma_\tau]} (\Pi_i - \lambda) \right) / \left(\sigma\sqrt{[\gamma_\tau]}\right) \le \left(\tau - \lambda([\gamma_\tau]+1)\right) / \left(\sigma\sqrt{[\gamma_\tau]}\right) \right\} + (2\pi\tau \log\tau)^{-\frac{1}{2}}$$

$$\le \varphi([\gamma_\tau]) + \Phi\left((\tau - \lambda\gamma_\tau)/(\lambda\sqrt{\gamma_\tau})\right) + (2\pi\tau \log\tau)^{-\frac{1}{2}}$$

$$\le \varphi(\lambda^{-1}\tau) + \Phi\left(-\left(\lambda^{-1}\tau(\log\tau)/\gamma_\tau\right)^{\frac{1}{2}} \right) + (2\pi\tau \log\tau)^{-\frac{1}{2}} \quad (6.11)$$

remembering that $\tau - \lambda\gamma_\tau$ is negative. In the algebra we have used an inequality on p. 175 of Feller (1968), which also yields an upper bound for the middle term of (6.11) of

$$\left\{\gamma_\tau/\left(\lambda^{-1}\tau \log \tau\right)\right\}^{\frac{1}{2}}\exp\left\{-\tfrac{1}{2}\left(\lambda^{-1}\tau(\log \tau)/\gamma_\tau\right)\right\}. \tag{6.12}$$

Let us call the exponential factor in (6.12) $f(\tau)$. Then

$$\log\left(\sqrt{\tau}f(\tau)\right) = \tfrac{1}{2}(\log \tau)\left\{\lambda^{-\frac{1}{2}}\sigma(\tau \log \tau)^{\frac{1}{2}}/\gamma_\tau\right\} \to 0 , \quad (\tau \to \infty) ,$$

so that $f(\tau) = O\left(\tau^{-\frac{1}{2}}\right)$. Since $\gamma_\tau = O(\tau)$, (6.12) is $O(\tau \log \tau)^{-\frac{1}{2}}$.

Hence (6.10) is

$$O\left(\max\left[\varphi\left(\lambda^{-1}\tau\right), (\tau \log \tau)^{-\frac{1}{2}}\right]\right) .$$

(ii) $y < -(\log \tau)^{\frac{1}{2}}$: By similar methods to those of (i)

$$(6.10) = O\left(\max\left[\varphi\left(\lambda^{-1}\delta\tau\right), \tau^{-\frac{1}{2}}\right]\right)$$

for any $\delta < 1$.

(iii) $0 \le y \le (\log \tau)^{\frac{1}{2}}$: Let $\delta_\tau \equiv \delta_\tau(y) \equiv \lambda^{-1}\tau + \lambda^{-3/2}\sigma\tau^{\frac{1}{2}}y$, and set $\Phi^*(y) = 1 - \Phi(y)$. Then

$$P\left\{\left(N(0, \tau]-\lambda^{-1}\tau\right)/\left(\lambda^{-3/2}\sigma\sqrt{\tau}\right) \le y\right\}$$

$$= P\left\{\left(\sum_{i=0}^{[\delta_\tau]} (\Pi_i-\lambda)\right)/\left(\sigma\sqrt{[\delta_\tau]}\right) > \left(\tau-\lambda\left([\delta_\tau]+1\right)\right)/\left(\sigma\sqrt{[\delta_\tau]}\right)\right\} \tag{6.13}$$

$$\le \varphi\left([\delta_\tau]\right) + \Phi^*\left[\left(\tau-\lambda\left(\delta_\tau+1\right)\right)/\left(\sigma\lambda^{-\frac{1}{2}}\sqrt{\tau}\right)\right] , \tag{6.14}$$

remembering that $\tau - \lambda\left([\delta_\tau]+1\right)$ is negative. Now $|\Phi(x)-\Phi(y)| \le |x-y|/\sqrt{2\pi}$, so that by adding and subtracting $\Phi^*\left[\left(\tau-\lambda\delta_\tau\right)/\left(\sigma\lambda^{-\frac{1}{2}}\sqrt{\tau}\right)\right]$ to the last term in (6.14), we obtain an upper bound of $\left(\text{N.B. } \Phi^*(-y) = \Phi(y)\right)$

$$\varphi\left(\lambda^{-1}\tau\right) + \left(2\pi\lambda\sigma^2\tau\right)^{-\frac{1}{2}} + \Phi(y) .$$

In the other direction, (6.13) has a lower bound of

$$-\varphi([\delta_\tau]) + \Phi^*\big((\tau-\lambda\delta_\tau)/\sigma\sqrt{\delta_\tau}\big)$$

$$\geq -\varphi(\lambda^{-1}\tau) + \Phi^*\big(-y\,(1+\lambda^{-\frac{1}{2}}\sigma y/\sqrt{\tau})^{-\frac{1}{2}}\big)$$

$$\geq -\varphi(\lambda^{-1}\tau) + \Phi^*(-y) - (|y|/\sqrt{2\pi})\,|(1+\lambda^{-\frac{1}{2}}\sigma y/\sqrt{\tau})^{-\frac{1}{2}}-1|$$

$$\geq -\varphi(\lambda^{-1}\tau) + \Phi(y) - (\log\tau)^{\frac{1}{2}}(\lambda^{-\frac{1}{2}}\sigma(\log\tau)^{\frac{1}{2}}/\sqrt{\tau})/\sqrt{2\pi}\ .$$

Hence

$$(6.10) = O\big(\max[\varphi(\lambda^{-1}\tau),\ (\log\tau)/\sqrt{\tau}]\big)\ .$$

(iv) $-(\log\tau)^{\frac{1}{2}} \leq y \leq 0$: By methods similar to those of (iii),

$$(6.10) = O\big(\max[\varphi(\lambda^{-1}\delta\tau),\ (\log\tau)/\sqrt{\tau}]\big)$$

for any $\delta < 1$. □

It would interesting to know if the $(\log\tau)/\sqrt{\tau}$ term in (6.9) can be replaced by $1/\sqrt{\tau}$. Applying our result to the $GI/G/\infty$ queue, we obtain (see Problem 6.3.5)

COROLLARY 2.6.12. *If the inter-arrival times and the service times of the $GI/G/\infty$ queue have finite third moments, then*

$$\left(\int_0^\tau \phi(s)ds - mE_2\{x_1\}\tau\right)/\sigma\sqrt{\tau} \xrightarrow{\mathcal{D}_1\times\mathcal{D}_2} \mathrm{N}$$

at rate $O(\tau^{-1/3})$. □

As remarked earlier, the metric used up till now is not continuous under Hölder continuous or Lipschitz mappings (Whitt (1974b)). Two suitable metrics are the Prohorov metric ρ and the dual bounded Lipschitz metric β (see (5.4) and (5.8) of Chapter 1). Our theorems (except possibly Lemma 2.6.11 and Corollary 2.6.12) hold with these metrics if we prove

LEMMA 2.6.13. *Suppose* $\beta(Y_\tau, Y) \leq \psi(\tau)$ $[\rho(Y_\tau, Y) \leq \psi(\tau)]$, *and* $\beta(Z_\tau, Z) \leq \varphi(\tau)$ $[\rho(Z_\tau, Z) \leq \varphi(\tau)]$. *Then if* N_τ *is a sequence of positive random variables such that* $N_\tau \to +\infty$ *in probability, and* $\{(N_\tau, Z_\tau)\}$ *is independent of* $\{Y_\tau\}$, *then* *(a)* $\beta\left[\left(Y_{N_\tau},\ Z_\tau\right),\ (Y,\ Z)\right] \le \varphi(\tau) + \int_0^\infty \psi(x)dPr\{N_\tau \le x\}$,

and, if in addition,

$\psi(\tau) = (\log \tau)^\gamma/\tau^\gamma$, $\gamma > 0$, $\delta > 0$ *and*

$\limsup_{n\to\infty} E|N_\tau - m\tau|^{1+\delta}/\tau < \infty$ *for some constant* m *, then*

(b) $\rho\left[\left(Y_{N_\tau},\ Z_\tau\right),\ (Y,\ Z)\right] = O\left[\max\left(\varphi(\tau),\ \psi(\tau)\right)\right]$.

Proof. *(a)* Let f denote an arbitrary bounded real-valued Lipschitz function on R^2 with norm ≤ 1 ; i.e.,

$$\|f\| \equiv \sup_{\substack{x\neq y \\ (x,y\in R^2)}} \left(|f(x)-f(y)|\right)/\left(\left[\sum_{i=1}^{2}\left(x_i-y_i\right)^2\right]^{\frac{1}{2}}\right) + \sup\left\{|f(x)|,\ x \in R^2\right\} \le 1 .$$

Then the functions $g_1(x) = f\left(x,\ x_2\right)$, x_2 fixed, and $g_2(x) = f\left(x_1,\ x\right)$, x_1 fixed, are clearly bounded real-valued Lipschitz functions on R^1 with norm ≤ 1 , and hence $\left(\int \text{ will mean } \int_{-\infty}^{+\infty}\right)$

$$\iint f(y,\ z)Pr\left\{\left(Y_{N_\tau},\ Z_\tau\right) \in (dy,\ dz)\right\}$$

$$= \int_0^\infty \iint [f(y,\ z)Pr\{Y_l \in dy\}]Pr\{Z_\tau \in dz,\ N_\tau \in dl\}$$

$$\le \int_0^\infty \psi(l)Pr\{N_\tau \in dl\} + \int\left[\int f(y,\ z)Pr\{Z_\tau \in dz\}\right]Pr\{Y \in dy\}$$

$$\le E\left\{\psi\left(N_\tau\right)\right\} + \varphi(\tau) + \iint f(y,\ z)Pr\{(Y,\ Z) \in (dy,\ dz)\} .$$

A similar inequality may be found in the other direction.

(b) Let F be any closed set in $\mathcal{B}\left(R^2\right)$. Then $F_x \equiv \left\{x_1 : \left(x_1,\ x\right) \in F\right\}$ is a closed set in $\mathcal{B}(R)$, and hence

$$Pr\{(Y_{N_\tau}, Z_\tau) \in F\}$$

$$= \int_0^\infty \int Pr\{Y_l \in F_x\} Pr\{Z_\tau \in dx, N_\tau \in dl\}$$

$$\leq \int_0^\infty \psi(l) Pr\{N_\tau \in dl\} + \int_0^\infty \int Pr\left\{Y \in F_x^{\psi(l)}\right\} Pr\{Z_\tau \in dx, N_\tau \in dl\} \ . \qquad (6.15)$$

Let $\theta < 1$; then splitting the $\int_0^\infty$ in the second term of (6.15) into $\int_{(0,m\theta\tau]}$ and $\int_{(m\theta\tau,\infty)}$ we obtain an upper bound for it of

$$Pr\{N_\tau \leq m\theta\tau\} + Pr\left\{(Y, Z_\tau) \in F^{\psi(m\theta\tau)}\right\} \qquad (6.16)$$

$$\leq A_{19}/\tau^\delta + Pr\left\{(Y, Z_\tau) \in \overline{F^{\psi(m\theta\tau)}}\right\}$$

$$\leq A_{19}(\log \tau)^\gamma/\tau^\delta + \varphi(\tau) + Pr\left\{(Y, Z) \in \left(\overline{F^{\psi(m\theta\tau)}}\right)^{\varphi(\tau)}\right\}$$

$$\leq A_{19}\psi(\tau) + \varphi(\tau) + Pr\left\{(Y, Z) \in F^{2\psi(m\theta\tau)+\varphi(\tau)}\right\}$$

using Chebyshev's inequality for the first term in (6.16). For our $\psi(\tau)$, $\psi(m\theta\tau) = O(\psi(\tau))$, and the first term in (6.15) is $\leq A_{20}\psi(\tau)$, by Lemma 2.6.6, so that the conclusion follows. □

As usual, in this theorem it has been tacitly understood that Y_τ, N_τ and Z_τ are all on the same space $(\Omega_1, \mathsf{F}_1, Pr_1)$ and Y and Z are on a possibly different space $(\Omega_2, \mathsf{F}_2, Pr_2)$. In (6.16), we have assumed Y and Z_τ on the same space: to be rigorous, one should work the proof through on the product space with measure $Pr_1 \times Pr_2$. However, we still need not demand that the limit variables be on the same space as the others in the final analysis.

2.7. Rates of functional convergence

Ideas related to speeds of convergence have received a boost since Rosencrantz' illuminating paper in 1968. The results of this section represent a generalisation of Section 3 of Kennedy (1972b), which is partly based on Rosencrantz' work. According to the usual procedure (and Kennedy's), our random functions will take values in the space $D[0, 1]$ of right-continuous functions on $[0, 1]$ with left-hand limits. We will consider rates of convergence in terms of the metric

$$\alpha(X, Y) \equiv \inf\{\varepsilon \geq 0 : Pr\{\rho(X, Y) \geq \varepsilon\} \leq \varepsilon\} , \tag{7.1}$$

where $\rho(x, y) \equiv \sup_{0 \leq t \leq 1} |x(t)-y(t)|$, $x, y \in D[0, 1]$. Note that this definition is only meaningful if X and Y are on the same space (Ω, F, Pr) , say, a fact which we will assume from now on. α corresponds to convergence in probability (Section 1.5) in the uniform metric on $D[0, 1]$ (see Problem 6.3.6). Observe that if $D[0, 1]$ is endowed only with the Skorokhod topology (Billingsley (1968), p. 111), it is easy to prove $\{\rho(X, Y) \geq \varepsilon\} \in F$, or, more generally, that $\rho(X, Y)$ is a random variable on (Ω, F) . Hence, the difficulty referred to in Section 2 of Whitt (1974b) of D being non-separable with the uniform topology (Billingsley (1968), p. 150) is avoided.

In this section we will regard the arrival process as a point process η and the service times as a sequence $\{V_i\}$ of non-negative random variables on (Ω, F, Pr) , and also suppose that standard Brownian motions W_1 and W_2 are defined on Ω , such that (η, W_1) and $(\{V_i\}, W_2)$ are independent. We will require the following random functions in $D[0, 1]$:

$$A_n(t) \equiv \big(\eta(0, nt]-mnt\big)/\big(\sigma_1\sqrt{n}\big) , \tag{7.2}$$

$$S_n(t) \equiv \left(\sum_{i=1}^{[nt]} (V_i-\mu)\right)/\big(\sigma_2\sqrt{n}\big) , \tag{7.3}$$

for some appropriate constants m, μ, σ_1 and σ_2 , and

$$Q_n(t) \equiv \left(\sum_{i=1}^{\eta(0,nt]} V_i-m\mu nt\right)/(\sigma\sqrt{n}) \tag{7.4}$$

where $\sigma^2 = \mu^2\sigma_1^2 + m\sigma_2^2$.

We first require a random change of time lemma.

LEMMA 2.7.1. *If for the* $G/G/\infty$ *queue,* $\alpha(A_n, W_1) \le \varphi(n)$ *and* $\alpha(S_n, W_2) \le \psi(n)$ *, then*

$$\alpha(Q_n, W) \le C_1 \max[\varphi(n), \psi(n), (\log n)^{3/4}/n^{\frac{1}{4}}] \tag{7.5}$$

for some constant C_1 *, where* $W(\cdot) = (\mu\sigma_1 W_1(\cdot)+\sigma_2 W_2(m\cdot))/\sigma$ *is also a standard Brownian motion.*

Remark. We will choose a scale such that $m < 1$. For another remark see Problem 6.3.7.

Proof. Let γ_n be a sequence $\downarrow 0$ $(n \to \infty)$, and set $C_2 = \max[2m\sigma_2/\sigma, 16\sigma_1/\sigma]$. Then

$$Pr\{\rho(Q_n, W) \ge C_2\gamma_n\}$$

$$\le Pr\left\{\rho(A_n, W_1) \ge \tfrac{1}{2}C_2\sigma(\mu\sigma_2)^{-1}\gamma_n\right\} + Pr\left\{\rho(S_n(m\cdot), W_2(m\cdot)) \ge \tfrac{1}{4}C_2\sigma\sigma_1^{-1}\gamma_n\right\} \tag{7.6}$$

$$+ Pr\left\{\rho\left(\sum_{i=1}^{\eta(0,nt]} (V_i-\mu)/\sigma_1\sqrt{n}, S_n(m\cdot)\right) \ge \tfrac{1}{4}C_2\sigma\sigma_1^{-1}\gamma_n\right\} . \tag{7.7}$$

Clearly, for $\delta_n \downarrow 0$ $(n \to \infty)$, and provided $\varepsilon_n \equiv \delta_n + n^{-1} \le 1 - m$ (true for $n \ge$ some n_1)

$$(7.7) \le Pr\left\{\sup_{0\le t\le 1} \left|\frac{\eta(0,nt]}{n} - mt\right| \ge \delta_n\right\}$$

$$+ Pr\left\{\sup_{|s-t|<\varepsilon_n} \left|\left(\sum_{i=[ns]+1}^{[nt]} (V_i-\mu)\right)/(\sigma_1\sqrt{n})\right| \ge 4\gamma_n\right\}$$

$$= Pr\{\rho(A_n, 0) \ge \sqrt{n}\delta_n/\sigma_2\} + Pr\{w_{S_n}(\varepsilon_n) \ge 4\gamma_n\}$$

$$\le Pr\{\rho(A_n, W_1) \ge \sqrt{n}\delta_n/2\sigma_2\} + Pr\{\rho(W_1, 0) \ge \sqrt{n}\delta_n/2\sigma_2\}$$

$$+ Pr\{\rho(S_n, W_2) \ge \gamma_n\} + Pr\{w_{W_2}(\varepsilon_n) \ge \gamma_n\} \tag{7.8}$$

where $w_x(\cdot)$ is the modulus of continuity, and we have used inequality (5.14) of Chapter 1. We choose $\delta_n = \left(2\sigma_2^2 \log n\right)^{\frac{1}{2}}/\sqrt{n}$. Then by equation

(2.4) of Kennedy (1972b), the second term of (7.8) has an upper bound of $4/\left(\pi n^2 \log n\right)^{\frac{1}{2}}$. But $n^{\frac{1}{2}}\delta_n \to \infty$ $(n \to \infty)$, and hence, for some n_2, the first term of (7.8) has an upper bound of $\varphi(n)$, $n \ge n_2$. The final term of (7.8) has, by Lemma 2.4 of Kennedy (1972b), an upper bound of

$$24(2\pi)^{-\frac{1}{2}}\gamma_n^{-1} \exp\left\{-\gamma_n^2/\left(18\varepsilon_n\right)\right\}/\varepsilon_n^{\frac{1}{2}}$$

for $n \ge$ some n_3. Take $\gamma_n \ge 6\delta_n^{\frac{1}{2}}\theta_n$ $\left(\ge \sqrt{18}\varepsilon_n^{\frac{1}{2}}\theta_n \text{ for } n \ge \text{some } n_4\right)$ for some θ_n. Then the upper bound becomes (go back to (7.8) and reapply Kennedy's Lemma 2.4) $4\pi^{-\frac{1}{2}} \exp\left\{-\theta_n^2\right\}/\varepsilon_n\theta_n$. Hence we take $\theta_n = (\log n)^{\frac{1}{2}}$, and obtain $(4/\sqrt{\pi})(\log n)^{-1}n^{-\frac{1}{2}}$ for this rate, but only if $\gamma_n \ge 6\sqrt{2}\sigma_2(\log n)^{3/4}/n^{\frac{1}{4}}$. Hence if

$$\gamma_n = \max\left[\varphi(n),\ \psi(n),\ 6\sqrt{2}\sigma_2(\log n)^{3/4}/n^{\frac{1}{4}}\right],$$

we obtain an upper bound for (7.6) and (7.7) of

$$2\left(\varphi(n)+\psi(n)\right) + C_3(\log n)^{3/4}/n^{\frac{1}{4}}$$

for some constant C_3, provided $n \ge \max\left[n_1, n_2, n_3, n_4\right]$. (7.5) then follows, where the constant C_1 may possibly need to be increased to account for the initial terms. □

For our main result we need to define a random function derived from the number of servers process. Let

$$M_n(t) \equiv \left(\int_0^{nt} m(s)ds - m\mu nt\right)/(\sigma\sqrt{n}) \tag{7.9}$$

where $m(s)$ is defined in Section 2.2, and $\sigma^2 = \mu^2\sigma_1^2 + m\sigma_2^2$.

We will also explicitly make the assumption now that the service times are stationary (but see Problem 6.3.4) so that $\mu = EV_i$, but we will not assume the arrival process is stationary. Now we can prove

THEOREM 2.7.2. *If for the* $G/G/\infty$ *queue* $\alpha(A_n, W_1) \le \varphi(n)$ *and* $\alpha(S_n, W_2) \le \psi(n)$ *, and in addition* $E\eta(-u, u) = O(u)$ *and* $E\left(V_i^{2p_2}\right) < \infty$ $(p_2 > 1)$ *, then*

$$\alpha(M_n, W) \le C_4 \max\left[\varphi(n), \psi(n), n^{-(p_2-1)/(2p_2+1)}, (\log n)^{3/4}/n^{\frac{1}{4}}\right],$$

where W *is as defined in Lemma* 2.7.1.

Proof. Clearly, for $\gamma_n \downarrow 0$, and $C_5 = \max[2C_1, 64, 8/\sigma]$,

$$Pr\{\rho(M_n, W) \ge C_5\gamma_n\} \le Pr\{\rho(M_n, Q_n) \ge C_5\gamma_n/2\} + Pr\{\rho(Q_n, W) \ge C_5\gamma_n/2\}$$

$$\le Pr\left\{\rho\left(\left|\int_0^{nt} m(s)ds - \sum_{i=1}^{\eta(0,nt]} V_i\right|, 0\right) \ge C_5\sigma\gamma_n\sqrt{n}/2\right\}$$

$$+ C_1 \max[\varphi(n), \psi(n), (\log n)^{3/4}/n^{\frac{1}{4}}]$$

provided $\gamma_n \ge \max[\varphi(n), \psi(n), (\log n)^{3/4}/n^{\frac{1}{4}}]$. Now the first term has an upper bound (for some $c_n \uparrow \infty$ $(n \to \infty)$) of

$$Pr\{M_2(n) \ge \tfrac{1}{4}C_5\sigma\gamma_n\sqrt{n}\} + Pr\{M_0'(n) \ge C_5\sigma\gamma_n\sqrt{n}/8\} \tag{7.10}$$

$$+ Pr\left\{\rho\left(\sum_{i=\eta(0,nt-c_n]+1}^{\eta(0,nt]} V_i/\sigma\sqrt{n}, 0\right) \ge C_5\gamma_n/8\right\} \tag{7.11}$$

in the notation of (2.5) and (2.6). If $E\left(V_i^2\right) < \infty$, then the first term of (7.10) is $O\left(n^{-\frac{1}{4}}\right)$ if $n^{-\frac{1}{4}} = O(\gamma_n)$ (*cf*. Lemma 2.6.7), which we require. Also,

$$(7.11) \le Pr\{w_{Q_n}(c_n/n) \ge C_5\gamma_n/16\} + Pr\{c_n/\sqrt{n} \ge C_5\gamma_n/16\}$$

$$\le Pr\{\rho(Q_n, W) \ge \gamma_n\} + Pr\{w_W(c_n/n) \ge \gamma_n\} \tag{7.12}$$

provided $\gamma_n \ge c_n/\sqrt{n}$. By Chebyshev's inequality, the second term of (7.10) has an upper bound of

$$C_6 n^{\frac{1}{2}}\gamma_n^{-1} \int_{c_n}^{\infty} l\,dF(l) , \tag{7.13}$$

where F is the distribution function of the service times. Hence, for this term to have an upper bound of γ_n, we require

$$\limsup_{n\to\infty} n^{\frac{1}{2}}\gamma_n^{-2} \int_{c_n}^{\infty} l\,dF(l) < \infty ,$$

a sufficient condition for which is

$$\limsup_{n\to\infty} n^{3/2} c_n^{-2} \int_{c_n}^{\infty} l\,dF(l) < \infty .$$

If $c_n = n^{\alpha}$, some $\alpha > 0$, this occurs if $E\left[V_i^{(3/2\alpha)-1}\right] < \infty$, and hence $c_n = n^{\frac{1}{2}-(p_2-1)/(2p_2+1)}$, i.e., we require $\gamma_n \geq n^{-(p_2-1)/(2p_2+1)}$. Finally, the second term in (7.12) has an upper bound of $C_7 n c_n^{-3/2} \exp\{-c_n/18\}$, using Lemma 2.4 of Kennedy (1972b), which converges to zero at a rate much faster than $c_n/\sqrt{n}$ and hence γ_n. □

Applying these result to the $GI/G/\infty$ queue, we obtain

COROLLARY 2.7.3. *If the arrival process to the* $G/G/\infty$ *queue is renewal with* $E\left[\Pi_i^{2p_1}\right] < \infty$, *where* Π_i *is an inter-arrival time, and the service times are i.i.d. with* $E\left(V^{2p_2}\right) < \infty$ $(p_1, p_2 > 1)$, *then if* $p = \min(p_1, p_2)$,

$$\alpha(M_n, W) \leq C_8\{(\log n)^p / n^{\min(p-1,p/2)}\}^{(2p+1)^{-1}} \equiv C_8 g(n, p) .$$

Proof. Heyde (1969) establishes $\psi(n) = O\big(g(n, p_2)\big)$ and Kennedy (1972b), Lemma 3.4, that $\varphi(n) = O\big(g(n, p_1)\big)$. □

Note that by using Lemmas 2.2 and 2.3 of Kennedy (1972b), and employing the techniques of his Lemmas 3.1 and 3.4 and proceeding as in (7.8) here, we can establish $g(n, p_1)$ as the relevant rate of convergence for the arrival process without recourse to Kennedy's Skorokhod representations: we only require the Skorokhod representation to give us the rate for

$\sum_{i=1}^{[nt]} \left(\Pi_i-\lambda\right)/\sigma\sqrt{n}$ (where $\lambda = E\{\Pi_i\}$, $\sigma^2 = \mathrm{Var}\{\Pi_i\}$) as per Heyde (1969).

This is a conceptually simpler approach.

CHAPTER 3

FUNCTIONAL LIMIT LAWS FOR CLUSTER POINT PROCESSES

3.1. Introduction

The cluster process has appeared extensively in the literature, a tribute to its practical interest as well as its theoretical accessibility. It has modelled many processes with some sort of regular "triggering" mechanism: early examples are contagion problems in ecology (Thompson (1955)), and the spatial distribution of galaxies (Neyman and Scott (1952), (1958)), followed by failure patterns in computers (Lewis (1964a), (1964b)), and the occurrence times and energies of earthquakes and aftershocks (Vere-Jones (1970)). It has also been used to investigate "bunching" in traffic flow (Bartlett (1963)). Recently, Hawkes and Oakes (1974) have demonstrated its close alliance to the Hawkes process. Finally, Matthes and others (e.g. Matthes (1963b), Kerstan, Matthes and Mecke (1974)), Goldman (1967) and Lee (1967), (1968) have studied these processes in relation to infinitely divisible point processes, since Poisson cluster processes are equivalent to regular infinitely divisible point processes. Closely related models also exist, such as Neyman and Scott's (1964) branching type model for epidemics. As can be seen, multidimensional cluster processes (e.g. galaxies, epidemics) are clearly important, but in this thesis we confine ourselves to one dimension.

This chapter is concerned with functional limit laws for stationary cluster processes. Weak convergence has come into vogue particularly since the publication of Billingsley's (1968) text on the subject, although the theory in its present form has been available since 1956 (Prohorov (1956), Skorokhod (1956)). Applications of the theory of weak convergence have been too diverse to detail - for a review and references, see Iglehart (1974). We mention only that Jagers (1974) has a theory of weak convergence of random measures and point processes on Polish spaces. The equivalence of functional central limit theorems for counting processes and the partial sum processes derived from inter-event times has been established in Iglehart and Whitt (1971) and Verwaat (1972). A summary of the theory of superpositions of point processes, including weak convergence problems, is given in Çinlar (1972). Whitt (1973) investigates rates of convergence of superposed processes to a Poisson process. Thinning of point processes is the subject of a paper by Jagers and Lindvall (1974) and also more generally in Kallenberg (1974). In these latter contexts (superpositions, thinning) in which a sequence of point processes converge to a point process, weak convergence and convergence of finite-dimensional distributions coincide, as was first pointed out by Straf (1972) and more generally by Whitt (1975). This equivalence has been generalized by Saunders (1975) to the space M of finite non-negative measures on a complete σ-compact metric space, where M is endowed with the topology of weak convergence.

The functional version of the law of the iterated logarithm for i.i.d. random variables was presented by Strassen in his famous paper in 1964. It has since been extended in various directions, for example in Heyde and Scott (1973) and Wichura (1973), and references there. Functional strong laws were introduced by Iglehart (1971b).

Section 2 of this chapter defines and examines existence of cluster processes with dependent clusters, as well as establishing notation needed to prove the limit theorems of Sections 3, 4 and 5. As with $G/G/\infty$ queue, we ask for limit theorems for the cluster process in terms of the same properties holding for the components. Also in Section 5, we investigate two processes related to cluster processes, by way of looking at limit laws for doubly stochastic Poisson processes, and the law of the iterated logarithm for the $G/G/\infty$ queue.

3.2. Definition and existence of cluster processes

A cluster point process η^* (compare Daley (1972)) is generated by two independent components, the centre process η consisting of points $\{t_j\}$ $(j = 0, \pm 1, \ldots)$, say, each of which initiates a subsidiary process η_j which is a.s. finite and independent of η . The full process consists of the superposition of the η_j's , i.e.,

$$\eta^*(A) = \sum_{\text{all } j} \eta_j\left(A-t_j\right) \text{ , bounded } A \in \mathcal{B}(R) \text{ .} \tag{2.1}$$

However, as in the arrival process for the $G/G/\infty$ queue, we mainly refer to the centre process as a probability measure P_1 on $\left(N, \sigma(N)\right)$. The subsidiary processes may be considered a probability measure P_2 on $\left(N^Z, \sigma(N)^Z\right)$, so that again assuming that the centre process and process of subsidiaries are independent we form the probability triple $\left(N \times N^Z, \sigma(N) \times \sigma(N)^Z, P_1 \times P_2\right)$.

Let N_∞ denote the set of all non-negative integer or infinite valued measures on R with σ-field generated by the sets $\{N \in N_\infty : N(A) \leq m\}$, $A \in \mathcal{B}(R)$, $m \in Z_+ \cup \{\infty\}$. This is the smallest σ-field such that the mappings $\phi_A : N_\infty \to [0, \infty]$ defined by $\phi_A(N) = N(A)$ are measurable, all $A \in \mathcal{B}(R)$.

We will define our cluster process via the mapping $\eta_c = \phi_3 \circ \phi_2 \circ \phi_1$ from $N \times N^Z \to N_\infty$ defined by

$$N \times N^Z \xrightarrow{\phi_1} R^Z \times N^Z \xrightarrow{\phi_2} N^Z \xrightarrow{\phi_3} N_\infty \tag{2.2}$$

(1) $(N, \mathrm{N}) \xmapsto{\phi_1} \left(\{t_j(N)\}, \mathrm{N}\right)$

(2) $\left(\{\beta_j\}, \{N_j\}\right) \xmapsto{\phi_2} \{N_j(\cdot - \beta_j)\}$

(3) $\{M_j\} \xmapsto{\phi_3} \sum_{j=-\infty}^{+\infty} M_j$,

where the product spaces in (2.2) are equipped with the product σ-fields, and $\{t_j\}$ are defined as in (2.1) of Chapter 1. It should be clear that (2.2) performs the same function as (2.1), namely, sums the contributions of each subsidiary process N_j centred at t_j to any Borel set A to give the number of cluster points in A .

If η_c is measurable, we can define a probability measure P_c on $(N_\infty, \sigma(N_\infty))$ by

$$P_c = (P_1 \times P_2)\eta_c^{-1} . \tag{2.3}$$

If $P_c(N) = 1$, then P_c is referred to as the cluster point process. Since proving measurability of ϕ_2 seems to be surprisingly difficult, we give a complete proof of

THEOREM 3.2.1. *(a) The mapping* η_c *defined around* (2.2) *is measurable.*

(b) $\sigma(N_\infty) \cap N = \sigma(N)$.

Proof. *(a)* Measurability of ϕ_1 follows from that of the sequence of maps $N \mapsto t_j(N)$, $j \in Z$, which in turn follows from (2.2) of Chapter 1. Let $\phi_A : N_\infty \to [0, \infty]$ be defined by $\phi_A(N) = N(A)$, $A \in B(R)$. To prove ϕ_3 is measurable we have to show that $\phi_A \circ \phi_3$ is measurable for each Borel set A . We also have the factorization $\phi_A \circ \phi_3 = \psi \circ \psi_A$ where

$$\{M_j\}_{j\in Z} \xmapsto{\psi_A} \{M_j(A)\}_{j\in Z} \text{ , i.e., } \psi_A : N^Z \to [0, \infty]^Z ,$$

$$\{\beta_j\}_{j\in Z} \xmapsto{\psi} \sum_{j=-\infty}^{+\infty} \beta_j \quad \text{, i.e., } \psi : [0, \infty]^Z \to [0, \infty] .$$

We prove that both ψ and ψ_A are measurable. $[0, \infty]^Z$ takes the product σ-field, so that measurability of ψ_A follows immediately from measurability of the map $M_j \mapsto M_j(A)$ from N to $[0, \infty]$. Since $\psi(\{\beta_j\}_{j\in Z}) = \lim_{n\uparrow\infty} \sum_{|j|\le n} \beta_j$, and each of the maps $\{\beta_j\} \mapsto \sum_{|j|\le n} \beta_j$ is continuous (when $[0, \infty]^Z$ is given the product topology) and the product σ-field is precisely the Borel σ-field induced by the product topology, measurability follows.

Measurability of ϕ_2 will follow from measurability of the coordinate map $(\{\beta_j\}, \{N_j\}) \mapsto N_j(\cdot-\beta_j)$, and since this map only depends on (β_j, N_j), from the measurability of $(\beta, N) \mapsto N(\cdot-\beta)$, or from the (joint) measurability of $(\beta, N) \mapsto N(A-B)$ for any $A \in \mathcal{B}(R)$. For this purpose, define

$$g(\beta, N) = \int h(x-\beta)dN(x) ,$$

where $h(x) \in C_K$, the class of continuous functions with compact support. Without loss of generality N may be endowed with the vague topology (see (5.9) and (5.10) of Chapter 1). Now, for any $\varepsilon > 0$, there exists a δ such that $|h(x-\beta)-h(x-\beta_0)| < \varepsilon$ for $|\beta-\beta_0| < \delta$. Let $K = \{x : |x-y| \le \delta, y \in (\text{supp } h)+\beta_0\}$. Then clearly $|h(x-\beta)-h(x-\beta_0)| < \varepsilon 1_K$ for $|\beta-\beta_0| < \delta$. So on the neighbourhood $\{N : N(K) \le N_0(K)+1\}$ of N_0,

$$\left|\int [h(x-\beta)-h(x-\beta_0)]N(dx)\right| \le \varepsilon N(K) \le \varepsilon[N_0(K)+1] .$$

Also, as β_0 is fixed, $\left\{N : \left|\int h(x-\beta_0)N(dx) - \int h(x-\beta_0)N_0(dx)\right| < \varepsilon\right\}$ is a neighbourhood of N_0 in the vague topology on N. Thus

$$|g(\beta, N)-g(\beta_0, N_0)| \le \varepsilon[N_0(K)+2]$$

on the intersection of these two neighbourhoods $\times\{\beta : |\beta-\beta_0| < \delta\}$. Thus $g(\beta, N)$ is a jointly continuous function of (β, N). According to Proposition 1.1 of Jagers (1974), $\sigma(N) = \mathcal{B}(N)$, the Borel σ-field generated by the open sets of the vague topology. Since the product σ-field on $R \times N$ is the Borel σ-field generated by the product topology $\{$(usual topology on R) $\times$ (vague topology on N)$\}$, it follows that g is also (jointly) measurable in this product σ-field.

Let G be the class of bounded measurable functions $g : R \to R$ for which $\int f(x-\beta)g(x-\beta)dN(x)$ is a jointly measurable function of (β, N) for every $f \in C_K$. Then every bounded continuous function is in G, since $f(x-\beta)g(x-\beta)$ is then continuous and has support a closed subset of a compact set, i.e., compact support, and G is closed under uniformly bounded pointwise limits. Hence G contains all bounded Borel measurable functions, and, in particular, 1_A for every Borel set A. Thus $\int f(x-\beta)1_A(x-\beta)dN(x)$ is measurable. Taking $f_n(x-\beta) \uparrow 1$, we obtain by the monotone convergence theorem that $\int 1_A(x-\beta)N(dx) = N(A-\beta)$ is measurable. *(b)* follows from Halmos (1950), Theorem E, p. 25. □

The existence of the cluster process $\left(\text{i.e., the condition } P_c(N) = 1\right)$ is a vital question, but a non-stochastic sufficient condition may be given as follows:

THEOREM 3.2.2. *If the subsidiary processes of a cluster process are uniformly bounded in the sense* $P_2\{N : N_i(I) > 0\} \le F(I)$ *(all* $i \in Z$*, bounded intervals* I *) for some set function* F*, then the cluster process exists if*

$$\int F(I-v)dM_1(v) < \infty \tag{2.4}$$

for all bounded intervals I*, provided the centre process has finite first moment measure* M_1.

Proof.

$P_c\{N : N(I) < \infty\} = 1$

$$\Longleftrightarrow \int_N P_2\left\{N : \sum_{j=-\infty}^{+\infty} N_j\big(I-t_j(N)\big) < \infty\right\} P_1(dN) = 1 \ ,$$

by (2.3) and Fubini's Theorem,

$$\Longleftrightarrow P_2\left\{N : \sum_{j=-\infty}^{+\infty} N_j\big(I-t_j\big) < \infty\right\} = 1 \ , \quad P_1 \ \text{- a.s.}$$

$$\Longleftrightarrow P_2\{N : N_i\big(I-t_i\big) > 0 \text{ i.o.}\} = 0 \ , \quad P_1 \ \text{- a.s.,}$$

since $N_i(R) < \infty$, P_2 - a.s.,

$$\Leftarrow \sum_{i=-\infty}^{+\infty} P_2\{N : N_i\big(I-t_i\big) > 0\} < \infty \ , \quad P_1 \ \text{- a.s.,}$$

by the Borel-Cantelli Lemma,

$$\Leftarrow \sum_{i=-\infty}^{+\infty} F\big(I-t_i\big) < \infty \ , \quad P_1 \ \text{- a.s.,} \tag{2.5}$$

$$\Longleftrightarrow \int_{-\infty}^{+\infty} F(I-v)dN(v) < \infty \ , \quad P_1 \ \text{- a.s.,}$$

$$\Leftarrow \int_{-\infty}^{+\infty} F(I-v)dM_1(v) < \infty \ . \qquad \square$$

It is clear that the reverse implications at (2.5) and the preceding step are also necessary if the clusters are i.i.d. (see also Westcott (1971); in fact the condition immediately following (2.5) was already familiar to the German school (see e.g. Kestan, Matthes and Mecke (1974), Chapters 5 and 6 and references there): their derivation was via the Borel-Cantelli Lemma). In fact, (2.4) is known to be necessary for a stationary Poisson cluster process with i.i.d. clusters (Matthes (1963b), Westcott (1971)). Westcott's techniques and Matthes' formulation do not seem to readily extend to non-independent subsidiaries.

Let us say that the subsidiary processes are stationary if with

$S_k : N^Z \to N^Z$ defined by $S_k\{N_j\} = \{N_{j+k}\}$,

ADDENDUM (at examiner's request)

We will follow Daley (1972) in employing the decomposition of η_c into a 'coherent' mapping η^+ with remainder terms $\eta^=$ and η^-. Specifically, we write

$$\begin{aligned}\eta_c(0, x] &= \tilde{\eta}(0, x] + \eta^=(0, x] \\ &= \eta^+(0, x] + \eta^=(0, x] - \eta^-(0, x] \end{aligned} \tag{2.7}$$

where

$$\begin{aligned}
\eta^+(0, x] &\equiv \sum_{j=1}^{N(0,x]} N_j(R) \,, \\
\tilde{\eta}(0, x] &\equiv \sum_{j=1}^{N(0,x]} N_j(-t_j, x-t_j] \,, \\
\eta^=(0, x] &\equiv \left(\sum_{j=-\infty}^{0} + \sum_{j=N(0,x]+1}^{\infty} \right) N_j(-t_j, x-t_j] \,, \\
\eta^-(0, x] &\equiv \sum_{j=1}^{N(0,x]} N_j\big(R \backslash (-t_j, x-t_j]\big) \,.
\end{aligned} \tag{2.8}$$

$$P_2\left(S_k^{-1}C\right) = P_2(C) \tag{2.6}$$

for any $C \in \sigma(N)^Z$, all $k \in Z$. Then just as in Theorem 2.3.1 *(a)*, we can prove

THEOREM 3.2.3. *If the centre process and process of subsidiaries of a cluster process are both stationary, then the cluster process itself is stationary.* □

In the remainder of this chapter, we will require, for notational convenience, that the centre process be stationary with finite intensity m, and the process of subsidiaries be stationary with finite first moment, i.e., $\mu \equiv E_2\{N_1(R)\} < \infty$. Under these circumstances it is easy to prove, as in Daley (1972), equation (26), that $E_c N(0, x] = \mu m x$. We will write $m_c = \mu m$. Also it is clear that the cluster process exists in the sense of Theorem 3.2.2.

Clarified in Addendum opposite.

We will follow Daley (1972) in employing the decompositions into a 'coherent' mapping η^+:

$$\eta_c(0, x] = \tilde{\eta}(0, x] + \eta^{=}(0, x], \quad \tilde{\eta}(0, x] = \eta^+(0, x] - \eta^-(0, x] \tag{2.7}$$

where

$$\tilde{\eta}(0, x] = \sum_{j=1}^{N(0,x]} N_j\left(-t_j, x-t_j\right], \quad \eta^+(0, x] = \sum_{j=1}^{N(0,x]} N_j(R). \tag{2.8}$$

Again the device due to Iglehart and Kennedy (1970) will be very useful. Let $c_n \uparrow \infty$ $(n \to \infty)$: then, as in (2.4) of Chapter 2,

$$\sup_{0 \le t \le s} \eta^-(0, nt]$$

$$\le \sum_{j=1}^{N(0,ns]} N_j\left(c_n, \infty\right) + \sum_{j=1}^{N(0,ns]} N_j\left(-\infty, -t_j\right] + \sup_{0 \le t \le s} \eta^+\left(nt-c_n, nt\right]. \tag{2.9}$$

Similarly,

$$\sup_{0\le t\le s} \eta^{=}(0,\ nt]$$

$$\le \sum_{j=1}^{N(0,ns]} N_j\left(-\infty,\ -c_n\right] + \eta^{=}(0,\ ns] + \sup_{0\le t\le s} \eta^{+}\left(nt,\ nt+c_n\right] \ . \quad (2.10)$$

3.3. A functional central limit theorem for cluster point processes

In their summary of the theory of point processes, Daley and Vere-Jones (1972) ask for a functional central limit theorem for cluster point processes. We show in Chapter 4 that ϕ-mixing theorems, and some strong mixing theorems are inapplicable in this context, but in this section we demonstrate that a different proof may be readily assembled from known sources. The proof follows mainly from techniques employed by Iglehart and Kennedy (1970), and by modifying Daley's (1972) proof for one dimension. In the process we remove a second moment condition on the centre process, and find an alternative to a first moment condition on the subsidiaries, but invoke one extra constraint not required in Daley (1972), namely that the weak limit of the normed centre process should have a.s. continuous sample paths (i.e., is in $C[0,\ \infty)$; see Whitt (1970)). We shall use $\Rightarrow$ to denote weak convergence in $D[0,\ \infty)$, the space of functions on $[0,\ \infty)$ which are right continuous with left-hand limits. The technique is to prove the weak convergence theorem on $D[0,\ s]$, each $s > 0$, and thus on $D[0,\ \infty)$ (Lindvall (1973), Whitt (1971)).

Before embarking on the main theorem itself, we will prove the following lemma, which is of independent interest. Let

$$L(t) = P_2\{N_1(-\infty,\ t] > 0\} \ , \quad R(t) = P_2\{N_1(t,\ \infty) > 0\} \ . \qquad (3.1)$$

LEMMA 3.3.1. *If the centre process* P_1 *of finite intensity* m *and process of subsidiaries* P_2 *are both stationary,*

$$\int_0^\infty L(-u)du < \infty \Rightarrow \sum_{j=1}^{\infty} N_j\left(-\infty,\ -t_j\right] < \infty \ , \quad P_1 \times P_2 \ - a.s.,$$

$$\int_0^\infty R(u)du < \infty \Rightarrow \sum_{j=-\infty}^{0} N_j\left(-t_j, \infty\right) < \infty \ , \quad P_1 \times P_2 \quad \text{- a.s.}$$

Proof. The proof seems more natural in the reverse direction. ($\Rightarrow$ in this proof means 'implies'.)

$$\sum_{j=-\infty}^{0} N_j\left(-t_j, \infty\right) < \infty \ , \quad P_1 \times P_2 \quad \text{- a.s.}$$

$$\Longleftrightarrow P_2\{N : N_j\left(-t_j, \infty\right) > 0 \ \text{ i.o., } j \le 0\} = 0 \ , \quad P_1 \quad \text{- a.s.}$$

$$\Leftarrow \sum_{j=-\infty}^{0} P_2\{N_j\left(-t_j, \infty\right) > 0\} < \infty \ , \quad P_1 \quad \text{- a.s.,}$$

by the Borel-Cantelli Lemma,

$$\Longleftrightarrow \int_{-\infty}^{0} P_2\{N_1(-v, \infty) > 0\}dN(v) < \infty \ , \quad P_1 \quad \text{- a.s.}$$

$$\Leftarrow \int_0^\infty R(v)dv < \infty$$

where we have taken expectations in the last step, and effected the change of variable $v' = -v$. The other half of the proof is similar. □

We aim for weak convergence to the Wiener process $W(\cdot)$ of the process

$$Z_n(t) \equiv \left(\eta_c(0, nt]-m_c nt\right)/\sqrt{n} \tag{3.2}$$

for $t \in [0, \infty)$. We also require the processes $X_n : N \to D[0, \infty)$,

$Y_n : N^Z \to D[0, \infty)$ defined by

$$X_n(t) \equiv \left(N(0, nt]-mnt\right)/\sqrt{n} \ , \tag{3.3}$$

$$Y_n(t) \equiv \sum_{j=1}^{[nt]} \left(N_j(R)-\mu\right)/\sqrt{n} \ . \tag{3.4}$$

THEOREM 3.3.2. *If the centre process* P_1 *is stationary, and*

$$P_1 X_n^{-1} \Rightarrow PrX^{-1} \quad . \tag{3.5}$$

where $Pr\{X(\cdot) \in C[0, \infty)\} = 1$ *, and the process of subsidiaries* P_2 *is stationary with the functions* (3.1) *satisfying*

$$n^2\{L(-n)+R(n)\} \to 0 \quad (n \to \infty) \tag{3.6}$$

and

$$P_2 Y_n^{-1} \Rightarrow PrY^{-1} \tag{3.7}$$

then, writing $Z(t) = \mu X(t) + Y(mt)$,

$$\left(P_1 \times P_2\right) Z_n^{-1} \Rightarrow PrZ^{-1} , \tag{3.8}$$

and X *and* Y *are independent.*

Remark. We choose a scale such that $m < 1$.

Proof. We use the decompositions (2.8), (2.9) and (2.10). For random functions θ_n, θ , we will abbreviate $\left(P_1 \times P_2\right)\theta_n^{-1} \Rightarrow Pr\theta^{-1}$ as $\theta_n \Rightarrow \theta$. Then

$$\left(\eta^+(0,\ nt]-m_c nt\right)/\sqrt{n} \Rightarrow \mu X(t) + Y(mt) \tag{3.9}$$

in the Skorokhod topology on $D[0,\ s]$, by Lemma 1 of Iglehart and Kennedy (1970). Let ρ_s and d_s respectively denote the uniform and Skorokhod metrics for $D[0,\ s]$ (*cf*. (5.11) and (5.12) of Chapter 1). To prove that the limit (3.8) occurs, it is sufficient to show that $\rho_s\left(\left(\eta^-(0,\ n\cdot]+\bar{\bar{\eta}}(0,\ n\cdot]\right)/\sqrt{n},\ 0\right) \Rightarrow 0$ by Theorem 4.1 of Billingsley (1968), since $\rho_s \geq d_s$. Provided $c_n = o(\sqrt{n})$ $(n \to \infty)$, then

$$\sup_{0 \leq t \leq s} \eta^+\left(nt-c_n,\ nt+c_n\right] \Rightarrow 0 \tag{3.10}$$

by Lemma 2 of Iglehart and Kennedy (1970). Also, for any $\varepsilon > 0$,

$$P_1 \times P_2\left\{\sum_{j=1}^{N(0,ns]} N_j\left(\left(-c_n,\ c_n\right]^c\right) \geq \varepsilon\sqrt{n}\right\} \tag{3.11}$$

$$\leq \int_N N(0,\ ns]P_2\left\{N_1\left(\left(-c_n,\ c_n\right]^c\right) > 0\right\}dP_1(N)$$

$$\leq ns\left\{L\left(-c_n\right)+R\left(c_n\right)\right\} = s\left(nc_n^{-2}\right)c_n^2\left\{L\left(-c_n\right)+R\left(c_n\right)\right\}$$

so that if a sequence $d_n \uparrow \infty$ $(n \to \infty)$ exists such that $d_n m_n \to 0$ $(n \to \infty)$, where $m_n \equiv n^2\{L(-n)+R(n)\}$, then (3.11) $\to 0$ $(n \to \infty)$. Clearly $m_n \to 0$ $(n \to \infty)$ is necessary for this. Taking $d_n = m_n^{-\frac{1}{2}}$ shows that it is also sufficient.

Consider

$$\eta^{=}(0,\ ns] \le \sum_{j=-\infty}^{0} N_j(-t_j,\ \infty) + \sum_{j=N(0,ns]+1}^{\infty} N_j(-\infty,\ ns-t_j] .$$

But by stationarity of $P_1 \times P_2$ with respect to the transformation T'_{ns} defined analogously to (3.4) of Chapter 2,

$$\sum_{j=N(0,ns]+1}^{\infty} N_j(-\infty,\ ns-t_j] \quad \text{and} \quad \sum_{j=1}^{\infty} N_j(-\infty,\ -t_j] \tag{3.12}$$

have the same distribution. Also, the middle term of (2.9) has an upper bound of $\sum_{j=1}^{\infty} N_j(-\infty,\ -t_j]$, so that, since the conditions of Lemma 3.3.1 are satisfied, dividing by $\sqrt{n}$ means that it converges $P_1 \times P_2$ - a.s. to zero as $n \to \infty$. Similarly $\eta^{=}(0,\ ns]$ converges to zero in $P_1{\times}P_2$-probability.□

Unfortunately the above argument contains one step which depends critically upon stationarity (see Problem 6.4.1). If the clusters are only right-handed, this is unimportant, but a more robust technique is of course to estimate the small probabilities by Chebyshev's inequality. This yields

COROLLARY 3.3.3. *Replacing* (3.6) *in Theorem* 3.3.2 *by the pair of conditions*

$$nE_2\{N_1((-n,\ n]^c)\} \to 0 \quad (n \to \infty)$$

$$\frac{1}{\sqrt{n}} \int_0^n E_2\left\{N_1((-u,\ u]^c)\right\} du \to 0 \quad (n \to \infty) \tag{3.13}$$

yields the same conclusions under the same conditions. □

The conditions (3.13) are predictably more severe than those of Daley (1972) for ordinary convergence, who used this technique. It should be clear though that we can find a weaker version of Theorem 3.3.2 for ordinary convergence.

COROLLARY 3.3.4. *If the centre process* P_1 *is stationary, and*

$$X_n(1) \xrightarrow{\mathcal{D}_1} X(1) , \tag{3.14}$$

and the process of subsidiaries P_2 is stationary with the functions (3.1) satisfying

$$\int_0^\infty \{L(-u)+R(u)\}du < \infty \tag{3.15}$$

and

$$Y_n(1) \xrightarrow{\mathcal{D}_2} Y(1) , \tag{3.16}$$

then $X(1)$ and $Y(1)$ are independent and

$$Z_n(1) \xrightarrow{\mathcal{D}_1 \times \mathcal{D}_2} Z(1) . \qquad \square$$

In Theorem 3.3.2 and its Corollaries we have avoided the question of norming. Denote the norming constants for $X_n(\cdot)$, $Y_n(\cdot)$ and $Z_n(\cdot)$ by σ_X, σ_Y and σ respectively; then clearly $\sigma^2 = \mu^2\sigma_X^2 + m\sigma_Y^2$. If the centre process is, say, ϕ-mixing with finite first and second moments and totally finite reduced covariance measure $C(\cdot)$ (see Section 1.3), then $\sigma_X^2 = \int_R C(du)$ (≥ 0, supposed > 0) and $X(\cdot) = \sigma_X W_1(\cdot)$, where $W_1(\cdot)$ is a standard Brownian motion. It is not difficult to prove, via equation A1 of Vere-Jones (Appendix to Daley (1971)) that $\lim_{u\to\infty} \operatorname{Var}\big(\eta(0, u]\big)/u = \int_R C_\eta(du)$ for any suitable weakly stationary point process η . If, also, the clusters are i.i.d. with $\mu_2 = E_2\left\{N_1(R)^2\right\} < \infty$, then $\sigma_Y^2 = \left(\mu_2-\mu^2\right)$, and $Y(\cdot) = \sigma_Y W_2(\cdot)$ for a standard Brownian motion $W_2(\cdot)$. Then $Z_n(t) \Rightarrow W(t)$ for a standard Brownian motion $W(\cdot)$ by Theorem 3.3.2, and also $\eta_c(\cdot)$ has totally finite reduced covariance measure $C_c(\cdot)$ satisfying $\int_R C_c(du) = \sigma^2$ (Theorem 4 of Daley (1972)).

3.4. The law of the iterated logarithm for cluster point processes

3.4.1. SOME DEFINITIONS AND LEMMAS

The law of the iterated logarithm (LIL) has been studied extensively for many years. In the context of point processes, the law has been demonstrated for the split times of branching processes by Athreya and Karlin (1967), and in queueing situations by Iglehart (1971a). The equivalence of the LIL for the counting process and the process of inter-epoch times has been demonstrated by Verwaat (1972), from which results for many point processes are immediate.

Firstly we require some lemmas which are mostly known. Denote by $D^j[0, \infty)$ $\left(D^j[0, s]\right)$ the j-fold product space of $D[0, \infty)$ $\left(D[0, s]\right)$. These spaces will be endowed with the product Skorokhod topology (Whitt (1971)). Generalizing Strassen (1964), we begin with

DEFINITION 3.4.1. Let K^j $\left(K_s^j\right)$ denote the set of absolutely continuous functions x in $D^j[0, \infty)$ $\left(D^j[0, s]\right)$ such that

$$x(0) = 0 \quad \text{and} \quad \int_0^\infty \dot{x}(t)^2 dt \le 1 \quad \left(\int_0^s \dot{x}(t)^2 dt \le 1\right) \tag{4.1.1}$$

where $\dot{x}$ denotes the derivative of x determined almost everywhere with respect to Lebesgue measure, and the square is to be interpreted as inner product. □

It follows immediately from the Schwarz inequality that for $x \in K^j$ and $0 \le a \le b$;

$$|x(b)-x(a)| \le (b-a)^{\frac{1}{2}} . \tag{4.1.2}$$

In the next definition we state formally the meaning of a functional law of the iterated logarithm (FLIL).

DEFINITION 3.4.2. A process $\{X_n(t)\}_{n\ge3}$ defined on a probability space (Ω, F, Pr), whose sample paths lie Pr - a.s. in $D^j[0, \infty)$

$(D^j[0, s])$ satisfies a FLIL if it is a.s. relatively compact in $D^j[0, \infty)$ $(D^j[0, s])$ and has K^j (K_s^j) as its set of limit points. □

Let $\phi(n) = (2n \log \log n)^{\frac{1}{2}}$. If $\{X_i\}$ is a sequence of zero mean, unit variance i.i.d. random variables, then Strassen's (1964) proof is easily adapted to show that $S_n(t) = \sum_{i=1}^{[nt]} X_i/\phi(n)$ satisfies a FLIL in $D[0, s]$, any $s > 0$. Now we quote the continuous mapping theorem for the FLIL (Strassen (1964), Wichura (1973)) in the form we require.

LEMMA 3.4.3. *Let* (Ω, F, Pr) *be a probability space and* S, S' *metric spaces with* K *a compact subset of* S. *Let* $f_n : \Omega \to S$ *be mappings such that* $\{f_n(\omega)\}$, $\omega \in \Omega$ *is a.s. relatively compact in* S *and has* K *as its set of limit points. Let* $\Phi : S \to S'$ *be a continuous mapping; then a.s. the sequence* $\Phi(f_n(\omega))$ *is relatively compact in* S' *and its set of limit points is* $\Phi(K)$. □

We wish to relate the FLIL on $D^j[0, \infty)$ and $D^j[0, s]$, $s > 0$, to each other. This can be done in the same way as Whitt (1971) for weak convergence. Let the continuous mapping $r_s : D^j[0, \infty) \to D^j[0, s]$, $s > 0$, be defined by $r_s(x)(t) = x(t)$, $0 \le t \le s$, for arbitrary $x \in D^j[0, \infty)$. Then we obtain

LEMMA 3.4.4. *A random vector* $\{X_n(t)\}$ *in* $D^j[0, \infty)$ *satisfies a FLIL if and only if the random sequences* $\{r_s(X_n)(t)\}$ *in* $D^j[0, s]$ *satisfy a FLIL for each* $s > 0$.

Proof. The forward direction follows immediately from Lemma 3.4.3 and observing that $r_s(K^j) = K_s^j$. Conversely, the relative compactness of $\{X_n(t)\}$ follows from that of $\{r_s(X_n)(t)\}$, all $s > 0$, by an easy extension of Theorem 2.3 of Whitt (1971) to $D^j[0, \infty)$. That the derived set is K^j ensues easily from the form of K_s^j. □

We require the following simple result in Section 3.4.2. Let $0 < \lambda \le 1$, $\alpha \ge 0$, $\beta \ge 0$, $\lambda\alpha^2 + \beta^2 = 1$ and define $g : D^2[0, s] \to D[0, s]$ by

$$g(y, z) = \alpha y(\lambda \cdot) + \beta z(\cdot) . \tag{4.1.3}$$

LEMMA 3.4.5. $g\left(K_s^2\right) = K_s^1$.

Proof. Let $x \in K_s^1$, and define $z(t) \in D[0, s]$ by $z(t) = \beta x(t)$. Also define $y(t) \in D[0, s]$ by

$$\begin{aligned} y(t) &= \alpha\lambda x(t/\lambda) , \quad 0 \le t \le \lambda s \\ &= \alpha\lambda x(s) \quad , \quad \lambda s \le t \le s . \end{aligned}$$

Then clearly $\big(y(0), z(0)\big) = (0, 0)$, and (y, z) are absolutely continuous with respect to Lebesgue measure. Also

$$\int_0^s \left[\dot{y}(t)^2+\dot{z}(t)^2\right]dt = \int_0^{\lambda s} \alpha^2\lambda^2\left[\frac{dx(u)}{du}\bigg|_{u=t/\lambda}\right]^2 dt + \beta^2 \int_0^s \dot{x}(t)^2 dt$$

$$= \left(\lambda\alpha^2+\beta^2\right) \int_0^s \dot{x}(t)^2 dt \le 1$$

where we have made the change of variable $t' = t/\lambda$ in the final step. Since $x(t) = \alpha y(\lambda t) + \beta z(t)$, $(y, z) \in K_s^2$, $K_s^1 \subset g\left(K_s^2\right)$.

Now to prove $g\left(K_s^2\right) \subset K_s^1$. Let $(y, z) \in K_s^2$. Clearly $g(y, z)(0) = 0$, and $g(y, z)$ is absolutely continuous with respect to Lebesgue measure.

$$\int_0^s \left\{\alpha \frac{dy(\lambda t)}{dt} + \beta\dot{z}(t)\right\}^2 dt$$

$$= \int_0^s \left(\alpha^2+\beta^2\lambda^{-1}\right)\left[\frac{dy(\lambda t)}{dt}\right]^2 + \left(\lambda\alpha^2+\beta^2\right)\dot{z}(t)^2 - \left\{\alpha\sqrt{\lambda}\dot{z}(t) - \beta\lambda^{-\frac{1}{2}}\frac{dy(\lambda t)}{dt}\right\}^2 dt$$

$$\leq \left(\alpha^2+\beta^2\lambda^{-1}\right)\lambda \int_0^{\lambda s} \dot{y}(t)^2 dt + \int_0^s \dot{z}(t)^2 dt$$

$$\leq \int_0^s \left(\dot{y}(t)^2+\dot{z}(t)^2\right)dt \leq 1 . \qquad \square$$

For $x, y \in D[0, s]$, $s > 0$, define the supremum metric by

$$\rho_s(x, y) \equiv \sup_{0\leq t\leq s} |x(t)-y(t)| \tag{4.1.4}$$

and the modulus of continuity of x by

$$w_s(x, \delta) \equiv \sup_{\substack{0\leq t,u\leq s \\ |t-u|\leq\delta}} |x(t)-x(u)| . \tag{4.1.5}$$

In Lemma 3.4.7 and Theorem 3.4.8 we will need

LEMMA 3.4.6. *If a process* $\{X_n(t)\}$ *on* $D[0, \infty)$ *satisfies a FLIL, then for any* $s > 0$ *,*

(a) $\rho_s\left(X_n, K_s^1\right) \to 0$ *a.s.* $(n \to \infty)$ *,*

(b) $\lim_{\delta\downarrow 0} \limsup_{n\to\infty} w_s\left(X_n, \delta\right) = 0$ *a.s.*

Proof. *(a)* Since K_s^1 is the derived set in the Skorokhod topology (metric d_s) on $D[0, s]$, $d_s\left(X_n, K_s^1\right) \to 0$ a.s. Hence for each $\omega \in \Omega$, $\varepsilon > 0$, there exists an $n_0(\omega, \varepsilon)$ such that for each $n \geq n_0(\omega, \varepsilon)$ there is $z_n(\omega, \varepsilon) \in K_s^1$ satisfying

$$d_s\left(X_n, z_n\right) < \varepsilon . \tag{4.1.6}$$

Now if Λ_s is the set of increasing invertible maps $[0, s] \to [0, s]$

$$|X_n(t)-z_n(t)| \leq |X_n(t)-z_n(\lambda_n(t))| + |z_n(\lambda_n(t))-z_n(t)|$$

for any $\lambda_n \in \Lambda$. Since $z_n \in K_s^1$, $|z_n(a)-z_n(b)| \leq (a-b)^{\frac{1}{2}}$, and by

(4.1.6), $\sup_t|\lambda_n(t)-t| \le \varepsilon$ for some λ_n. Hence $\rho_s(X_n, K_s^1) \le \varepsilon + \varepsilon^{\frac{1}{2}}$, i.e., *(a)* holds.

(b) can then be proved as in Iglehart (1971a), Theorem 3.3 (see Problem 6.4.2). For completeness, we give it here. Clearly for $z \in K_s^1$, $w_s(z, \delta) \le \delta^{\frac{1}{2}}$. But for $\varepsilon > 0$, $\omega \in \Omega$, $n \ge n_0(\omega, \varepsilon)$, there exists (as in *(a)*) a $z_n \in K_s^1$ such that

$$w_s(X_n, \delta) \le w_s(z_n, \delta) + 2\rho_s(X_n, z_n) \le \delta^{\frac{1}{2}} + 2\varepsilon$$

using (5.13) of Chapter 1, and the conclusion follows. □

This proof depends strongly on the properties of K_s. It is not simply a result of the relative compactness of processes satisfying the FLIL (i.e., of the Arzelà-Ascoli Theorem (Billingsley (1968), p. 211)), although for processes 'sufficiently like those in $C[0, s]$' it is (since the Skorokhod topology relativized to C coincides with the uniform topology there). For instance if $S_n(t) = \sum_{j=1}^{[nt]} X_j/\sigma\phi(n)$ satisfies a FLIL, where $\{X_j\}$ is stationary, $EX_1 = 0$ and σ is some positive constant, and if we associate with $S_n(\cdot)$ its linear approximation

$$S_n'(t) = \left[\sum_{j=1}^{[nt]} X_j + (nt-k)X_{k+1}\right]/\sigma\phi(n) \;,\quad ks \le nt \le ks+1 \;,\quad 0 \le k \le ns-1 \;,$$

then
$$\rho_s(S_n, S_n') \le \max_{1 \le k \le ns} |X_k|/\sigma\phi(n) \;. \tag{4.1.7}$$

Clearly the RHS of (4.1.7) $\to 0$ a.s. if $|X_n|/\phi(n) \to 0$ a.s. which follows from the Borel-Cantelli Lemma if $EX_1^2 < \infty$ (see Problem 6.4.3). This type of argument may also be applied to $\eta_n(t) = (\eta(0, nt]-mnt)/\sigma\phi(n)$ for a suitable point process η. The point is, though, that if the derived set is not K^1 (e.g., Theorem 4.2 of Iglehart (1971a)) for some process $X_n(t)$ which is relatively compact in D, and X_n is not 'sufficiently like a process in $C[0, s]'$, Lemma 3.4.6 will have to be reproved before it can be used.

Finally we require a 'random change of time' lemma: this result has been observed before by Freedman (1967) and Iglehart (1971b), although not quite in this setting.

LEMMA 3.4.7. *Let* $\{X_n(t)\}_{n\geq 3}$ *and* $\{N_t\}$ $(t \geq 0)$ *be processes defined on a space* (Ω, F, Pr) *such that* $N_t \geq 0$, $N_t \uparrow \infty$ $(t \to \infty)$ *and* $N_t/t \to m$, $0 < m < 1$ *a.s.* $(t \to \infty)$ *; if* $\{X_n(\cdot)\}$ *satisfies a FLIL, then upon setting*

$$\Phi_n(t) \equiv N_{nt}/n \, , \tag{4.1.8}$$

$$\rho_s\big(X_n\big(\Phi_n(\cdot)\big), X_n(m\cdot)\big) \to 0 \quad a.s., \text{ each } s > 0 \, .$$

Remark. This of course means that $X_n\big(\Phi_n(\cdot)\big)$ and $X_n(m\cdot)$ obey the same FLIL on $D[0, s]$, any $s > 0$, and hence on $D[0, \infty)$ by Lemma 3.4.4.

Proof. The process N_t obeys a functional strong law of large numbers (as per Iglehart (1971b), Theorem 3.1); hence, a.s. for each $\omega \in \Omega$, $\delta > 0$, there exists an $n_0 \equiv n_0(\omega, \delta)$ such that for $n \geq n_0$,

$$\rho_s\big(\Phi_n(\cdot), m\cdot\big) \leq \delta \, .$$

Hence, choosing $\delta \leq s(1-m)$, for $n \geq n_0$,

$$\rho_s\big(X_n\big(\Phi_n(\cdot)\big), X_n(m\cdot)\big) \leq w_s\big(X_n, \delta\big)$$

and the conclusion follows from Lemma 3.4.6 *(b)*. □

3.4.2. THE LIL FOR PROCESSES WITH RIGHT HAND CLUSTERS

We are now in a position to prove our theorem for cluster processes with right-hand clusters. We must suppose the centre process and process of subsidiaries to be defined on the same space $N \times N^Z$.

Once again we suppose our processes stationary. Let

$$X_n(t) \equiv \big(N(0, nt]-mnt\big)/\sigma_1\phi(n) \, , \tag{4.2.1}$$

$$Y_n(t) \equiv \sum_{j=1}^{[nt]} \left(N_j(R)-\mu\right)/\sigma_2\phi(n) \tag{4.2.2}$$

for some appropriate constants σ_1, σ_2, where $\phi(n) = (2n \log\log n)^{\frac{1}{2}}$.

Then setting $\sigma^2 = \mu^2\sigma_1^2 + m\sigma_2^2$, we define

$$Z_n(t) \equiv \left(\eta_c(0, nt]-m_c nt\right)/\sigma\phi(n) . \tag{4.2.3}$$

THEOREM 3.4.8. *If* (X_n, Y_n) *jointly satisfy a FLIL, i.e., if* (X_n, Y_n) *is relatively compact in* $D^2[0, \infty)$ $P_1 \times P_2$ *- a.s. and has* K^2 *as its set of limit points, and if either*

$$\int_e^\infty E_2\{N_1(u, \infty)\}/(\log\log u)^{\frac{1}{2}} du < \infty \tag{4.2.4}$$

or, if, in the notation of (3.1),

$$\int_0^\infty uR(u)du < \infty \tag{4.2.5}$$

then $P_1 \times P_2$ *- a.s.* Z_n *is relatively compact in* $D[0, \infty)$ *with limit set* K^1 .

Remarks. The condition on (X_n, Y_n) and (4.2.4) and (4.2.5) will be discussed in Section 3.4.3. We will take $m < 1$, a condition easily achieved by scaling (see Problem 6.4.4).

Proof. Clearly (recalling that $m_c = \mu m$),

$$Z_n(t) = \sigma^{-1}\mu\sigma_1 X_n(t) + \sigma^{-1}\sigma_2 Y_n\left(N(0, nt]/n\right)$$
$$+ \left\{\sum_{j=1}^{N(0,nt]} N_j\left(nt-t_j, \infty\right)/\sigma\phi(n)\right\} + \left\{\sum_{j=-\infty}^{0} N_j\left(-t_j, nt-t_n\right]/\sigma\phi(n)\right\} . \tag{4.2.6}$$

Define a continuous mapping $g_s : D^2[0, s] \to D[0, s]$ by

$$g_s(x, y) = \left(\sigma_1\mu x(\cdot)+\sigma_2 y(m\cdot)\right)/\sigma . \tag{4.2.7}$$

Expressing the first two terms of (4.2.6) in terms of g_s, we find

$$\rho_s\left(g_s\left[X_n(\cdot),\ Y_n\left(N(0,\ n\cdot]/mn\right)\right],\ g_s\left(X_n,\ Y_n\right)\right)$$

$$= \sigma^{-1}\sigma_2\rho_s\left(Y_n\left(N(0,\ n\cdot]/n\right),\ Y_n(m\cdot)\right) \to 0\ ,\quad P_1 \times P_2\ \text{- a.s.}$$

by Lemma 3.4.7. Hence the FLIL for the first two terms of (4.2.6) is the same as for $g_s\left(X_n,\ Y_n\right)$, which, by Lemma 3.4.3, is relatively compact in $D[0,\ s]$ and has $g_s\left(K_s^2\right) = K_s^1$ (Lemma 3.4.5) as its set of limit points. Then, defining $g : D^2[0,\ \infty) \to D[0,\ \infty)$ as in (4.2.7), $g\left(X_n,\ Y_n\right)$ is almost surely relatively compact in $D[0,\ \infty)$ with limit set K^1 (Lemma 3.4.4).

We now prove that the condition (4.2.4) guarantees that the remainder terms in (4.2.6) converge almost surely (i.e., ρ_s (remainder terms, 0) $\to 0$, $P_1 \times P_2$ - a.s.) to zero.

For any arbitrary fixed $s > 0$, define random functions in $D[0,\ s]$ by

$$\gamma_n(t) \equiv \left(\eta^+(0,\ nt]-m_c nt\right)/\sigma\phi(n)$$

$$\theta_n(t) \equiv \sum_{j=1}^{N(0,nt]} N_j\left(c_{mns},\ \infty\right)/\phi(n)\ ,$$

where $c_n \uparrow \infty$ $(n \to \infty)$. Note that $\gamma_n(\cdot)$ represents the first two terms of (4.2.6). Now

$$\rho_s\left(\sum_{j=1}^{N(0,nt]} N_j\left(nt-t_j,\ \infty\right)/\phi(n),\ 0\right)$$

$$\le \theta_n(s) + \rho_s\left(\eta^+\left(nt-c_{mns},\ nt\right],\ 0\right)$$

$$\le \theta_n(s) + \sigma\rho_s\left(\gamma_n(t),\ \gamma_n\left(t-c_{mns}/n\right)\right) + \mu c_{mns}/\phi(n)\ . \quad (4.2.8)$$

We take $c_n = o\left(\phi(n)\right)$, so that the final term of (4.2.8) converges to zero. The second term of (4.2.8) has an upper bound of $\sigma w_s\left(\gamma_n,\ c_{mns}/n\right)$, which, by Lemma 3.4.6, converges to zero almost surely since we have proved $\gamma_n(t)$ satisfies a FLIL. We now tackle the first term. This term can be handled using a geometric subsequence argument, but it seems worthwhile to record the following alternative argument, which yields exactly the same condition. Let

$$\theta_n'(t) \equiv \sum_{j=1}^{[mnt]} N_j\left(c_{mns}, \infty\right)/\phi(n) .$$

Since X_n satisfies a FLIL, there exists, for any $\delta > 0$ and for each $N \in N$ $\left(P_1\right.$ - a.s.$\left.\right)$, a $n_0(N, \delta)$ such that for $n \geq n_0(N, \delta)$,

$$[n(mt-\delta)] \leq N(0, nt] \leq [n(mt+\delta)] , \quad 0 \leq t \leq s .$$

Hence, for $n \geq n_0(N, \delta)$,

$$\rho_s\left(\theta_n, \theta_n'\right) \leq w_s\left(\theta_n', \delta\right) \leq 2\rho_s\left(\theta_n', 0\right) = 2\theta_n'(s)$$

and $\theta_n(s) \to 0$ a.s. if $\theta_n'(s) \to 0$ a.s. But

$$\theta_n'(s) \leq \left\{\sum_{j=1}^{[mns]} N_j\left(c_j, \infty\right)/\phi([mns])\right\}.\{\phi([mns])/\phi(n)\} \tag{4.2.9}$$

and the RHS of (4.2.9) converges to zero if the first factor converges to zero, i.e., if $\sum_{j=1}^{n} N_j\left(c_j, \infty\right)/\phi(j) \to$ a.s. by Kronecker's Lemma. But by monotonicity, we only require a subsequence to converge, i.e.,

$$\sum_{j=1}^{N(0,n]} N_j\left(c_j, \infty\right)/\phi(j) \to P_1 \times P_2 \text{ - a.s.} \tag{4.2.10}$$

We require $c_j = o\left(\phi(j)\right)$, so we take $c_j = m^{-\frac{1}{2}}(j \log\log\log j)^{\frac{1}{2}}$; sincethe LHS of (4.2.10) will converge if it has an asymptotic upper bound wh converges, and since $t_j/j \to m^{-1}$ a.s. (*cf.* Proof of Theorem 2.4.1 *(b)*), we require only the convergence of

$$\sum_{j=1}^{N(0,n]} N_j\left(\sqrt{t_j}, \infty\right)/\phi\left(t_j\right) . \tag{4.2.11}$$

By the sub-martingale convergence theorem (Breiman (1968), p. 89), a sufficient condition is

$$\limsup_{n\to\infty} \int_e^n E_2\{N_1(\sqrt{u}, \infty)\}/(u \log\log u)^{\frac{1}{2}} du < \infty .$$

Changing variables $u' = \sqrt{u}$ gives (4.2.4). Note that a finer condition can be obtained by refining the $\sqrt{t}$ factor in (4.2.11), but since the resultant expression is untidy, and there is a grosser approximation in the argument (see Problem 6.4.5), we have not given it.

The final term of (4.2.6) may be dealt with by a standard geometric sybsequence argument, which we will not elaborate on. It yields the condition (use Chebyshev's inequality, the Borel-Cantelli Lemma, bound sums by integrals, exchange integrals *cf.* proof of Theorem 2.4.1 *(b)*)

$$\int_0^\infty E_2\{N_1(u, \infty)\}/\sqrt{u}\, du < \infty$$

which is weaker than (4.2.4).

To obtain the condition (4.2.5), we now employ a geometric subsequence argument on $\theta_n(s)$. Let $\alpha > 1$, and consider $r_k = [\alpha^{2k}]$, $k \in Z_+$. For an arbitrary fixed $\varepsilon > 0$, writing $d(n) = c_{mns}$, define

$B_k = \left\{ \sum_{j=1}^{N(0,r_kt]} N_j\big(d(r_{k-1}), \infty\big] > \varepsilon\phi(r_{k-1}) \right\}$. Then clearly

$P_1 \times P_2\{\theta_n(s) > \varepsilon\phi(n) \;\; \text{i.o.}\} \le P_1 \times P_2\{B_k \;\; \text{i.o.}\}$, since, if for some l , $\theta_n(l) > \varepsilon\phi(l)$, then, setting $r_k(l)$ to be the next term in the geometric subsequence greater than l , $B_k(l)$ holds (in an obvious notation). We require $P_1 \times P_2\{B_k \text{ i.o.}\}$ to be zero, which occurs if $\sum P_1{\times}P_2\{B_k\} < \infty$ by the Borel-Cantelli Lemma. Arguing as in (3.11), and choosing $d(n) = \sqrt{n}$ (we require $d(n) = o\big(\phi(n)\big)$, and choosing finer $d(n)$'s leads to intractab expressions), this occurs if

$$\sum_{k=3}^{\infty} r_k R\big(r_{k-1}\big) \le \int_3^\infty \alpha^{2\theta} R\big(\alpha^{\theta-3}\big)\, d\theta$$

$$\le \big(\alpha^6/\log\alpha\big) \int_1^\infty uR(u)du < \infty ,$$

where we change variables $u = \alpha^{\theta-3}$ in the last step.

The final term of (4.2.6) converges to zero a.s. under the condition (4.2.5) by Lemma 3.3.1. □

COROLLARY 3.4.9 (Ordinary Law of the Iterated Logarithm). *Under the conditions of Theorem* 3.4.8,

$$\limsup_{n\to\infty} Z_n(1) = +1 \;, \quad \liminf_{n\to\infty} Z_n(1) = -1$$

Proof. As indicated by Strassen (1964). □

3.4.3. COMMENTS

(a) The LIL for processes with double-sided clusters

If the clusters can occur on the left-hand side of their initiating points, we obtain more remainder terms which we must prove converge to zero almost surely. Most of these can be handled as before, but

$$\sum_{j=N(0,n]+1}^{\infty} N_j\left(-t_j,\, n-t_j\right]/\phi(n) \tag{4.3.1}$$

proves difficult. If the left-hand clusters have bound d, then ∞ in (4.3.1) may be replaced by $N(0,\, n+d]$, and this converges to zero a.s. as does $\gamma_n(t)$ in (4.2.8). In the general case, if we replace $N(0,\, n] + 1$ in (4.3.1) by $N\left(0,\, \alpha^2 n\right] + 1$ for some $\alpha > 1$, then we can prove convergence using the standard geometric subsequence argument and Chebyshev's inequality under the condition

$$\int_1^{\infty} E_2\{N_1(-\infty,\, -u]\}/\sqrt{u}\; du < \infty \;.$$

Conditions ensuring that the remainder (the sum from $N(0,\, n] + 1$ to $N\left(0,\, \alpha^2 n\right]$) converge to zero a.s. seem elusive. The reasoning around (3.12) has also defied extension to the a.s. convergence case.

(b) The conditions in Theorem 3.4.8.

It is clear that if the clusters have only one member each, then (4.2.4) is much weaker than (4.2.5). On the other hand, it is easy to construct examples in which (4.2.5) is satisfied but (4.2.4) not, but mostly these involve $N_1(R)$ not having a first moment. The following example indicates that under fairly reasonable circumstances (4.2.4) may be much weaker. A general comparison does not seem feasible.

The cluster members of a Bartlett-Lewis process form a finite renewal process of length S, where the inter-point times are i.i.d. with common distribution F and independent of S. Let

$R_s(r) = \sum_{j=r+1}^{\infty} Pr\{S = j\}$. Then, using Lawrance (1972), Equation (4.3.7), (4.2.4) becomes

$$\int_e^{\infty} \sum_{i=0}^{\infty} R_s(i)\left(1-F^{(i+1)*}(u)\right)/(\log\log u)^{\frac{1}{2}}du < \infty . \qquad (4.3.2)$$

Suppose, for example, that F has regularly varying tails with exponent $\alpha > 0$ (see Feller (1966), p. 268), i.e., as $u \to \infty$,

$$1 - F(u) \sim u^{-\alpha}\hat{L}(u)$$

where $\hat{L}(s) \geq 0$ is of slow variation. Then, using the Corollary, p. 272 of Feller (1966), and provided $E\left(S^2\right) < \infty$, (4.3.2) reduces to

$$\int_e^{\infty} \hat{L}(u)/\left(u^{\alpha}(\log\log u)^{\frac{1}{2}}\right)du < \infty ,$$

so that even if $\hat{L}(u)$ is bounded, we still require $\alpha > 1$, a rather strong condition. However, $R(u) = \sum_{j=0}^{\infty} Pr\{S=j\}\left(1-F^{j*}(u)\right)$, so that (4.2.5) will require $\alpha > 2$, a much stronger condition.

The assumption that X_n and Y_n jointly satisfy a FLIL warrants some attention. Certainly this is true if the centre process is a stationary renewal process (or the superposition of stationary renewal processes) whose second moment exists, and the clusters are i.i.d.: for, if W_1 and W_2 are independent Brownian motions on a common space Ω, and we redefine X_n and Y_n also to be on Ω, we can show $\rho_s\left(X_n, W_n^1\right) \to 0$ a.s. (Iglehart

(1971a), Section 2) and $\rho_s\left(Y_n, W_n^2\right) \to 0$ a.s. (Strassen (1964)), where $W_n^i(t) = W_i(nt)/\phi(n)$, $i = 1, 2$. Hence if ρ_s^2 is the product supremum metric for $C^2[0, s]$ or $D^2[0, s]$,

$$\rho_s^2\left(\left(X_n, Y_n\right), \left(W_n^1, W_n^2\right)\right) \to 0 \quad \text{a.s.} \tag{4.3.3}$$

The result then follows from Strassen (1964), Theorem 1, who employed the Skorokhod representation theorem approach. Alternatively, if the centre process η and process of subsidiaries $\boldsymbol{\eta}$ possess some form of asymptotic independence, say ϕ-mixing, and given this, satisfy the conditions of Corollary 3 of Heyde and Scott (1973), then again the above Skorokhod type approach, as used by Heyde and Scott, will prove that the joint FLIL follows merely from the individual FLIL's and the independence of X_n and Y_n . Alternatively, for η and $\boldsymbol{\eta}$ ϕ-mixing or strongly mixing, we may possibly use Chover's (1967) approach as applied by Oodaira and Yoshihara (1971a, b) to demonstrate this. It is intuitively clear, however, that the joint FLIL will not necessarily follow from merely the individual FLIL's and independenc we need some statement such as (4.3.3) on the 'density' of subsequences converging to particular points in the limit set. Indeed, in order to satisfactorily achieve this, we may need to define the LIL for processes via an appropriate generalization of (4.3.3).

3.5. Related topics

3.5.1. FUNCTIONAL STRONG LAWS FOR CLUSTER POINT PROCESSES

The strong law of large numbers for cluster processes has been investigated by Daley (1972). Here we strengthen his theorem, as well as generalize it to the functional case.

Let $T_y : N \to N$, $S_k : N^Z \to N^Z$ and $T_y' : N \times N^Z \to N \times N^Z$ be defined analogously to (3.1), (3.2) and (3.4) of Chapter 2, and let T_1, T_2 and T

denote their respective invariant σ-fields. We will assume P_1 and P_2 stationary (Theorem 3.2.3).

THEOREM 3.5.1. *If the centre process* P_1 *is stationary with* $E_1N(0, 1] < \infty$ *, and the process of subsidiaries* P_2 *is stationary with* $E_2N_1(R) < \infty$ *, then*

$$\lim_{n\to\infty} \eta_c(0, n]/n = E_1 \times E_2\{\eta_c(0, 1]|T\} \text{ , } P_1 \times P_2 \text{ - a.s.,}$$

where $(E_1 \times E_2)\eta_c(0, 1] = E_1\{N(0, 1]\}E_2\{N_1(R)\} < \infty$ *, and*

$$E_1 \times E_2\{\eta_c(0, 1]|T\} = E_1\{N(0, 1]|T_1\}E_2\{N_1(R)|T_2\} \text{ .} \qquad (5.1)$$

Remark. The LHS and RHS of (5.1) are random variables on $N \times N^Z$ (see Problem 6.4.6).

Proof. Omitted. Similar to the proofs of Theorem 5, Daley (1972) and Theorem 2.4.1 *(a)*. □

COROLLARY 3.5.2. *If the conditions of Theorem* 3.5.1 *are satisfied, and* $\{N(0, n] : n \geq 1\}$ *and* $\{N_i(R) : i \geq 1\}$ *both satisfy the strong law of large numbers* P_1 *- a.s. and* P_2 *- a.s. respectively, then as* $n \to \infty$ *,*

$$\eta_c(0, n]/n \to E_1\{N(0, 1]\}E_2\{N_1(R)\} < \infty \text{ , } P_1 \times P_2 \text{ - a.s.} \qquad \square$$

COROLLARY 3.5.3. *Let* $L_s(t) = E_1\{N(0, 1]|T_1\}E_2\{N_1(R)|T_2\}t$ *,* $t \in [0, s]$ *. Then if* ρ_s *is the supremum metric on* $D[0, s]$ *, and the conditions of Theorem* 3.5.1 *are satisfied,*

$$\rho_s\big(\eta_c(0, n\cdot]/n, L_s\big) \to 0 \text{ , } P_1 \times P_2 \text{ - a.s.}$$

Proof. Repeat the argument in Iglehart (1971b), Theorem 3.1, for each $(N, \mathrm{N}) \in N \times N^Z$ for which convergence of $\eta_c(0, n]/n$ holds. □

3.5.2. THE LAW OF THE ITERATED LOGARITHM FOR THE *G/G/∞* QUEUE

The LIL for the "number of servers process" of the $G/G/\infty$ queue may be proved using similar techniques to those for cluster processes (note that here we only have right-hand 'clusters'). Hence we will give no proofs. Let the arrival process η and the process of service times $\{V_j\}$ be defined on a common probability space (Ω, F, Pr) (we need this formulation in Corollary 3.5.5). Let

$$A_n(t) \equiv \big(\eta(0, nt]-mnt\big)/\sigma_1\phi(n)$$

$$S_n(t) \equiv \sum_{i=1}^{[nt]} \big(V_i-E(V_1)\big)/\sigma_2\phi(n)$$

for some appropriate constants σ_1, σ_2, $m = E\,\eta(0, 1] < \infty$, where $\phi(n) = (2n \log \log n)^{\frac{1}{2}}$.

THEOREM 3.5.4. *Suppose* η *and* $\{V_j\}$ *are stationary with finite second moments, and* (A_n, S_n) *jointly obey a FLIL; then the process*

$$\left(\int_0^{nt} m(s)ds - mE\big(V_1\big)nt\right)/\sigma\phi(n)$$

obeys a FLIL, where $\sigma^2 = (EV_1)^2\sigma_1^2 + m\sigma_2^2$. □

Once again we require $m < 1$. A similar theorem holds if η is replaced by η^0, the arrival process corresponding to P_1^0 (see Section 2.4). The proof again consists of approximating by a coherent mapping, and showing various remainder terms converge a.s. to zero. Geometric subsequence arguments suffice for this latter half of the proof, indeed, seem necessary for the η^0 case, but in the η case neater proofs can be devised via Kronecker's Lemma and monotonicity.

We will say that a process $X(t)$, $t \geq 0$, obeys the ordinary law of the iterated logarithm (OLIL) if, with $W(t)$, $t \geq 0$, a Brownian motion defined on the same space (Ω, F, Pr),

$$|X(t)-\phi(t)^{-1}W(t)| \to 0 \text{ a.s.}$$

The techniques of the proof of Theorem 3.5.4 reveal

COROLLARY 3.5.5 (see Problem 6.4.7). *If the arrival process is* *stationary with* $m < \infty$ *, and* $A_n(1)$ *obeys an OLIL, and the service times are i.i.d. with finite second moment, then*

$$\left(\int_0^t m(s)ds - mE(V_1)t\right)/\sigma\phi(t)$$

obeys the OLIL. □

Note that the FLIL has been proved under weaker moment conditions than the functional central limit theorem. Also, if a non-stationary arrival process η (with appropriate centring and norming) satisfies the OLIL, and $E_\eta(I) = O(|I|)$, $|I| \to \infty$, then so will the "accumulated number of servers" process.

3.5.3. LIMIT LAWS FOR THE DOUBLY STOCHASTIC POISSON PROCESS

Kingman (1964) has shown that a doubly stochastic Poisson process may be represented as a random time transformation of a stationary Poisson process of unit parameter. Hence it is unnecessary to prove functional limit laws for these processes, as they will follow from e.g. Section 17 of Billingsley (1968), or Lemma 3.4.7 here. Ordinary central limit theorems can be deduced from Lemma 2.5.1, rendering unnecessary the characteristic function techniques of Grandell (1971), at least in his cases $0 < k < \infty$ (see Problem 6.4.8).

CHAPTER 4

ASYMPTOTIC INDEPENDENCE OF POINT PROCESSES, PARTICULARLY CLUSTER PROCESSES

4.1. Introduction

The concept of asymptotic independence arose simultaneously with ergodic theory, and its suitability, in the form of strong mixing and ϕ-mixing, as a sufficient condition for central limit theorems has been known since the papers of Rosenblatt (1956), Billingsley ((1956), (1962)) and Ibragimov (1962). Important recent developments are due to Oodaira and Yoshihara ((1971a), (1971b), (1972)) and Heyde (1974). Stronger mixing conditions have also been studied e.g. Philipp (1969).

Our original motivation for considering ϕ-mixing of point processes was the hope that it would be preserved under the clustering operation (independent subsidiaries), and hence give a simple avenue to functional limit theorems for cluster point processes (Daley and Vere-Jones (1972), Theorem 8.6). In fact, Westcott (1973), (Concluding remarks), suggests that limit laws obtained by more direct methods (as in Chapter 3) should be weakened in the presence of a mixing condition. However, our investigations suggest that, unlike weaker forms of asymptotic independence (Westcott (1971), (1972)), ϕ-mixing and strong mixing are only maintained under extra conditions, which are quite severe (bounded clusters) in the case of ϕ-mixing.

Our basic definitions conclude this section, and in Section 4.2 we give two examples of point processes which are not ϕ-mixing; the first of these examples illustrates, in Section 4.3, the problem of characterizing complete ϕ- and strong mixing, and that complete mixing (or ϕ-mixing) on a determining class is not always sufficient to establish complete mixing (respectively ϕ-mixing). Section 4.4 investigates the preservation of various modes of mixing under clustering, and Section 4.5 contains concludir remarks.

We will need the following definitions: let

$$\sigma\big(N(B)\big) \equiv \sigma_{\text{gen.}}\left\{\{N(A) = k\},\ A \in \mathcal{B}(R) \cap B,\ k \in Z_+\right\}\ ,\quad B \in \mathcal{B}(R)\ ,$$

i.e., the smallest σ-field such that the maps $N \longmapsto N(A)$ $\big($for any $A \in \mathcal{B}(R) \cap B\,\big)$ from $N \to R$ are measurable.

We shall introduce the term 'complete mixing' for a stationary point process P satisfying, for a given function $\gamma(\cdot)$,

$$|P(C \cap D)-P(C)P(D)| \le \gamma(\tau)P(C)P(D) \tag{1.1}$$

whenever $C \in \sigma\big(N(-\infty,\ t]\big)$, $D \in \sigma\big(N(t+\tau,\ \infty)\big)$, $t \in R$, $\tau > 0$. Here $\gamma : [0,\ \infty) \to [0,\ 1]$ is a monotone decreasing function satisfying $\lim_{\tau\to\infty} \gamma(\tau) = 0$. This type of mixing appears in at least one paper of Philipp (1969). As a non-trivial example (not a point process), consider a process that is defined as a real-valued function on the state space of a discrete time aperiodic irreducible stationary Markov chain on some finite state space. Such a process is completely mixing (for proof, see Billingsley (1968), pp. 167-8, where the example is used to illustrate the weaker concept of ϕ-mixing).

A stationary point process P is ϕ-mixing for a given function $\phi(\cdot)$ if

$$|P(C \cap D)-P(C)P(D)| \le \phi(\tau)P(C)\ , \tag{1.2}$$

and strong mixing for a given function $\alpha(\cdot)$ if

$$|P(C \cap D)-P(C)P(D)| \le \alpha(\tau)\ , \tag{1.3}$$

where C and D are as before, and $\phi(\tau)$ and $\alpha(\tau)$ have the same properties as $\gamma(\tau)$ (see Problem 6.5.13).

In the following we will assume, as usual, that the centre process is stationary, and that the process of subsidiaries is stationary as well as being independent of the centre process. Also, we will take it for granted that our cluster processes exist (in the sense of Theorem 3.2.2), and the subsidiaries are a.s. finite.

4.2. Two examples of point processes which are not ϕ-mixing

It is well-known that the output of a stationary $M/G/\infty$ queue is a stationary Poisson process, being equivalent to the random translations of points of the original (input) process (see, e.g., Daley (1975) for references).

Let G_δ $(\delta > 0)$ denote a distribution function satisfying

$$\begin{aligned} G_\delta(x) &= 0 \qquad , \quad x \le \delta \ , \\ &= G(x-\delta) \ , \quad x > \delta \ , \end{aligned} \tag{2.1}$$

where $G(x)$ is any arbitrary distribution function on $[0, \infty)$. The following is true:

THEOREM 4.2.1. *Superposing the input and output of a $M/G/\infty$ queue for which $G(x) < 1$ for all finite x results in a stationary point process which is not ϕ-mixing.*

Remark. In this example, and the next, we will take our processes on an arbitrary probability space (Ω, F, Pr), although we could equally as well work from $N \times R_+^Z$.

Proof (see also Problem 6.5.2). We will only consider the case where $G(x) = F_\delta(x)$, some $\delta > 0$, and some F. The stronger result can be proved using the techniques of Example 4.2.2, but the F_δ case is neater.

Let λ be the common parameter of the input (η_1) and output (η_2) processes. The superposed input and output process is denoted by $\eta \ (= \eta_1 + \eta_2)$.

Let h satisfy $\exp(-2\lambda h) > \frac{1}{2}$, $0 < h \le \delta$. Then $\eta_1(-h, 0]$ and $\eta_2(-h, 0]$ are independent Poisson variables, and hence, for all $k \in Z_+$,

$$Pr\{\eta(-h, 0] = k\} = e^{-2\lambda h}(2\lambda h)^k/k! \ . \tag{2.2}$$

Now using Milne (1970), Theorem 2, or from first principles,

$$Pr\{\eta(-h,\ 0] = k,\ \eta(\tau,\ \tau+h] = 0\} = \exp\left(-4\lambda h+2\lambda hp(\tau)\right)\left(2\lambda h\left(1-p(\tau)\right)\right)^{k}/k!\ , \quad (2.3)$$

where $p(\tau) = \int_{-h}^{0} \left[F_\delta(\tau+h-v)-F_\delta(\tau-v)\right]dv/2h$.

Thus from (2.2) and (2.3), we have that

$Pr\{\eta(\tau,\ \tau+h] = 0|\eta(-h,\ 0] = k\} - Pr\{\eta(\tau,\ \tau+h] = 0\}$

$$= \exp(-2\lambda h).\left[\exp\left(2\lambda hp(\tau)\right)\left(1-p(\tau)\right)^{k}-1\right]\ . \quad (2.4)$$

For all finite x , $F_\delta(x) < 1$, so for any given finite τ we can find $\tau^* > \tau$ such that $p(\tau^*) > 0$, and hence $1 - p(\tau^*) < 1$. Thus

$\sup_k |Pr\{\eta(\tau^*,\ \tau^*+h] = 0|\eta(-h,\ 0] = k\} - Pr\{\eta(\tau^*,\ \tau^*+h] = 0\}|$

$$= \exp(-2\lambda h) > \tfrac{1}{2}\ .$$

Let $P_\eta \equiv Pr\eta^{-1}$ be the probability measure on $\left(N,\ \sigma(N)\right)$ corresponding to η. Clearly,

$$\phi(\tau) \equiv \sup_{C,D} |P_\eta(D/C)-P_\eta(D)| \geq \phi(\tau^*) > \tfrac{1}{2}\ ,$$

where the supremum is taken over $C \in \sigma\left(N(-\infty,\ t]\right)$, $D \in \sigma\left(N(t+\tau,\ \infty)\right)$. Hence $\phi(\tau) \not\downarrow 0$ $(\tau \to \infty)$, which shows that the process is not ϕ-mixing. □

Unbounded translations of a ϕ-mixing point process do not necessarily result in a ϕ-mixing point process. Counter-examples can be difficult to establish however, since, as remarked before, it is well-known that Poisson processes are invariant under translation, and the "counting" behaviour of non-Poisson centre processes after translation is in general algebraically intractable (or well-nigh so), with the exception of the translation of compound Poisson processes, as in the following example.

EXAMPLE 4.2.2. Consider a centre process of Poisson doublets (i.e., a stationary Poisson process with each point doubled) with rate 1 , and i.i.d. translations X such that $Pr\{X > t\} = e^{-t}$. Hence if η is the cluster process (see the remark at the end of this section), then by standard techniques we can prove

$$E\{z^{\eta(-h,0]}\} = \exp\left(-\sigma+z\rho+z^2\nu\right) \quad (2.5)$$

where $\rho = 2\left(1-e^{-h}\right)$, $\nu = h - \left(1-e^{-h}\right)$ and $\sigma = \rho + \nu$. (This is the probability generating function of the so-called Hermite distribution (Kemp and Kemp (1965)).)

Similarly we can prove

$$E\left\{z^{\eta(-h,0]}1_{\{\eta(\tau,\tau+h]=0\}}\right\} = \exp\left(-\gamma+z\delta+z^2\nu\right) \tag{2.6}$$

where $\delta = \rho - \left(1-e^{-h}\right)^2 e^{-\tau}$, and $\gamma = 2\sigma - \left(1-e^{-h}\right)^2 e^{-\tau}$ and $1_A(\omega)$, $\omega \in \Omega$, is the indicator function of the set A .

Let $d_n(\alpha, \beta) = \sum_{j=0}^{n} \alpha^{n-j}\beta^{2j}/\left((n-j)!(2j)!\right)$ for any $\alpha, \beta \geq 0$. From (2.5) and (2.6),

$$Pr\{\eta(\tau, \tau+h] = 0 \mid \eta(-h, 0] = 2k\} = \exp\left(\sigma-\left(1-e^{-h}\right)^2 e^{-\tau}\right).d_k(\nu, \delta)/d_k(\nu, \rho) .$$

We assert that $\lim_{k\to\infty} d_k(\nu, \delta)/d_k(\nu, \rho) = 0$. To prove this, choose $\varepsilon \geq 0$ and J such that $\delta^{2j} \leq \varepsilon\rho^{2j}$ for $j > J$. Then for $k > J$,

$$d_k(\nu, \delta)/d_k(\nu, \rho) \leq \varepsilon + \left[\sum_{j=0}^{J} \nu^{k-j}\delta^{2j}/\left((k-j)!(2j)!\right)\right]/d_k(\nu, \rho) . \tag{2.7}$$

Let l be any fixed integer, $0 \leq l \leq J$. Then

$$\left(\nu^{k-l}\delta^{2l}\right)/\left((k-l)!(2l)!d_k(\nu, \rho)\right) = \left(\delta^2/\nu\right)^l/\left((2l)!D_k\right)$$

where $D_k = \sum_{j=0}^{k} (k-l)!\left(\delta^2/\nu\right)^j/\left((k-j)!(2j)!\right)$.

But the $(l+1)$th term of D_k is $(k-l)\left(\delta^2/\nu\right)^{l+1}/(2l+2)!$. Hence $D_k \to \infty$ as $k \to \infty$, and all the terms in the fixed sum of (2.7) have limit 0 . Hence for $C \in \sigma\left(N(-\infty, t]\right)$, $D \in \sigma\left(N(t+\tau, \infty)\right)$,

$$\sup_{C,D} \left|P_\eta(D|C)-P_\eta(D)\right| \geq \exp(-\sigma) ,$$

where P_η is the probability measure corresponding to η , so that η is not ϕ-mixing.

Note that the point process in this last example is the same as an elementary Neyman-Scott cluster process with Poisson centres, and in which each subsidiary consists of 2 points whose positive distances D_1, D_2 from the centre have independent exponential distributions with parameter 1 .

4.3. On characterisations of complete, ϕ- and strong mixing

Westcott (1971) has given a neat characterisation of mixing in terms of the probability generating functional (p.g.fl.) (see also Westcott (1972)). This is very useful for cluster point processes, since, if $G_1[\xi]$ is the p.g.fl. of the centre process, and $G_2(\xi|t)$ is the p.g.fl. of a subsidiary given its centre is at t (we are considering independent subsidiaries here) then the p.g.fl. $G[\xi]$ of the cluster point process is given by

$$G[\xi] = G_1[G_2(\xi|t)] \ .$$

(Our definition of p.g.fls. is over a suitable class V of real-valued measurable functions ξ for which $1 - \xi$ has bounded support, i.e., $\xi(t) = 1$ for t outside some bounded set.) Let $V_I \subset V$ denote those functions ξ in V for which $1 - \xi$ vanishes outside I . Then we can prove the following

THEOREM 4.3.1. *Let* $\xi_1 \in V_{(-\infty,t]}$, $\xi_2 \in V_{(t,\infty)}$, *and thus* $S_\tau\xi(u) \equiv \xi(u-\tau) \in V_{(t+\tau,\infty)}$, $\tau > 0$. *If a stationary point process with p.g.fl.* G *is*

(a) *completely mixing with rate* $\gamma(\tau)$, *then*

$$|G[\xi_1 S_\tau \xi_2]-G[\xi_1]G[\xi_2]| \leq \gamma(\tau)G[\xi_1]G[\xi_2] \ ; \tag{3.1}$$

(b) *ϕ-mixing, then*

$$|G[\xi_1 S_\tau \xi_2]-G[\xi_1]G[\xi_2]| \leq 2\phi(\tau)G[\xi_1] \ . \quad \square \tag{3.2}$$

The proof of *(a)* is almost immediate. For a proof of *(b)*, see Theorem 4.3.3, in which similar, but slightly more complex techniques are used (see

Problem 6.5.7).

The converses to Theorem 4.3.1 do not hold. It is easy to demonstrate that the p.g.fl. of the superposed input and output of the $M/M/\infty$ queue satisfies (3.1) (and thus also (3.2)), with

$$G[\xi] = \exp\left\{-\int_{-\infty}^{+\infty}\left[1-\xi(t)\int_0^\infty \xi(t+x)\mu dx\right]\lambda dt\right\}$$

and

$$\gamma(\tau) = \min\left(1, \exp\left[\lambda\mu^{-1}e^{-\mu\tau}\right]-1\right) \tag{3.3}$$

where λ is the arrival rate, and μ the service time parameter (see Problem 6.5.4).

For orderly point processes P, mixing can be characterized using the zero probability function

$$\phi(B) = P\{N(B) = 0\} \ , \quad B \in \mathcal{B}(R) \ . \tag{3.4}$$

Kurtz (1974) has stated without proof similar characterisations of stationarity and ergodicity.

THEOREM 4.3.2. *If, for an orderly point process* P,

$$\lim_{\tau\to\infty} \phi\left(B_1 \cap \left(B_2+\tau\right)\right) = \phi\left(B_1\right)\phi\left(B_2\right) \tag{3.5}$$

for any $B_1, B_2 \in R(I)$, *the ring of finite unions of intervals, then* P *is mixing.*

Proof. Firstly note that $S \equiv \left\{\{N(B) = 0\}, B \in R(I)\right\}$ is closed under intersections, and that $\sigma_{\text{gen.}}(S) = \sigma(N)$ (Kallenberg (1973)). Then, for any fixed $A_2 \in S$,

$$\mathcal{D}_1 \equiv \{A_1 : \lim_{\tau\to\infty} P\left(A_1 \cap T_\tau A_2\right) = P\left(A_1\right)P\left(A_2\right)\}$$

is a Dynkin system (Ash (1972), p. 168) and hence by the Dynkin system theorem, $\mathcal{D}_1 \supset \sigma_{\text{gen.}}(S)$. Similarly, for any fixed $A_1 \in \sigma(N)$,

$$\mathcal{D}_2 \equiv \{A_2 : \lim_{\tau\to\infty} P\left(A_1 \cap T_\tau A_2\right) = P\left(A_1\right)P\left(A_2\right)\}$$

is a Dynkin system so that $\mathcal{D}_2 \supset \sigma_{\text{gen.}}(S)$. □

The equations for complete mixing and ϕ-mixing analogous to (3.5) follow from (3.1) and (3.2) by taking ξ's of the form $1 - 1_A(t)$. Hence, similar characterisations do not hold for complete mixing and ϕ-mixing. If they did, the converses to Theorem 4.3.1 would be true for orderly point processes.

We turn now to characterisations of strong mixing, aiming at an analogue of Theorem 4.3.1. Firstly we note that we can extend the domain of the probability generating functionals to include complex-valued, measurable functions $\phi \in \Phi$ with $\phi(x) = 1$ on the complement of a bounded set, and satisfying $\sup_{x \in R} |\phi(x)| = 1$ (see, e.g., Fisher (1972)). This, because we anticipate that the RHS of (3.2) will no longer contain $G[\xi_1]$, and the modulus on the LHS can be interpreted more liberally. If, however, we introduce the characteristic functional (ch.fl.)

$$C[\theta] \equiv G[\exp i\theta(t)] \tag{3.6}$$

where $\theta(t)$ has bounded support (i.e., $\theta \in V$), and define

$$\Phi_{(t,t+x]} \equiv \{\phi \in \Phi : 1-\phi \text{ has support within } (t, t+x]\}$$

we can easily show, via the multidimensional maximum modulus principle (e.g., Gunning and Rossi (1965), p. 7) that

$$|G[\phi_1 S_\tau \phi_2]-G[\phi_1]G[\phi_2]| \le |C[\theta_1+S_\tau\theta_2]-C[\theta_1]C[\theta_2]| \tag{3.7}$$

where $\phi_1 \in \Phi_{(-\infty,t]}$, $\phi_2 \in \Phi_{(t,\infty)}$; $\theta_1 \in V_{(-\infty,t]}$, $\theta_2 \in V_{(t,\infty)}$. Hence the ch.fl. is (potentially) more useful than the p.g.fl. in characterising strong mixing.

THEOREM 4.3.3. *If a stationary point process P is strongly mixing with rate $\alpha(\tau)$, then its ch.fl. satisfies*

$$|C[\theta_1+S_\tau\theta_2]-C[\theta_1]C[\theta_2]| \le 4\alpha(\tau) \tag{3.8}$$

where $\theta_1 \in V_{(-\infty,t]}$, $\theta_2 \in V_{(t,\infty)}$.

Proof. Let $I_1, \ldots, I_k \in B(R) \cap (t+\tau, \infty)$, and $J_1, \ldots, J_l \in B(R) \cap (-\infty, t]$. Denote $(I_1, \ldots, I_k)$ by I and

$(J_1, \ldots, J_l)$ by J . Also, put

$$\begin{aligned} \mathrm{P} &\equiv P(\mathrm{I}, \mathrm{n}; \mathrm{J}, \mathrm{m}) \\ &\equiv P\{N(I_1) = n_1, \ldots, N(I_k) = n_k; N(J_1) = m_1, \ldots, N(J_l) = m_l\} \\ &\qquad - P\{N(I_1) = n_1, \ldots, N(I_k) = n_k\}P\{N(J_1) = m_1, \ldots, N(J_l) = m_l\} . \end{aligned}$$

Then

$$\begin{aligned} &|E(\Pi_r\Pi_s \exp i[\theta_r N(I_r)+\psi_s N(J_s)]) - E(\Pi_r \exp[i\theta_r N(I_r)])E(\Pi_s \exp[i\theta_s N(J_s)])| \\ &\qquad = |\Sigma_{\mathrm{n}}\Sigma_{\mathrm{m}} P(\mathrm{I}, \mathrm{n}; \mathrm{J}, \mathrm{m})\exp[i(\boldsymbol{\theta}.\mathrm{n}+\boldsymbol{\psi}.\mathrm{m})]| \qquad (3.9) \end{aligned}$$

where $\boldsymbol{\theta} \equiv (\theta_1, \ldots, \theta_k)$, $\boldsymbol{\psi} \equiv (\psi_1, \ldots, \psi_l)$. Now set $\gamma = \boldsymbol{\theta}.\mathrm{n} + \boldsymbol{\psi}.\mathrm{m}$ and define $D_1 = \{\mathrm{n}, \mathrm{m} : \mathrm{P} \cos \gamma \geq 0\}$, $D_3 = \{\mathrm{n}, \mathrm{m} : \mathrm{P} \sin \gamma \geq 0\}$, and let D_2, D_4 be their respective complements. Clearly

$$\begin{aligned} (3.9) \leq \max(|\Sigma_{D_1} \mathrm{P} \cos \gamma|, |\Sigma_{D_2} \mathrm{P} \cos \gamma|) \\ + \max(|\Sigma_{D_3} \mathrm{P} \sin \gamma|, |\Sigma_{D_4} \mathrm{P} \sin \gamma|) . \qquad (3.10) \end{aligned}$$

Now let $E_1 = \{\mathrm{n}, \mathrm{m} : \cos \gamma \geq 0\}$, $E_3 = \{\mathrm{n}, \mathrm{m} : \sin \gamma \geq 0\}$ and E_2, E_4 be their respective complements. Then

$$|\Sigma_{D_1} \mathrm{P} \cos \gamma| \leq |\Sigma_{D_1 \cap E_1} \mathrm{P} \cos \gamma| + |\Sigma_{D_1 \cap E_2} \mathrm{P} \cos \gamma| \leq 2\alpha(\tau) , \qquad (3.11)$$

and similarly for the other expressions in the right-hand side of (3.10). Hence (3.9) $\leq 4\alpha(\tau)$. If we define

$$\theta_1(u) = \sum_{i=1}^{l} \psi_i 1_{J_i}(u) , \quad \theta_2(u) = \sum_{i=1}^{k} \theta_i 1_{I_i - \tau}(u)$$

then we have

$$|C[\theta_1+S_\tau\theta_2]-C[\theta_1]C[\theta_2]| \leq 4\alpha(\tau) . \qquad (3.12)$$

A measurable function $\theta \in V$ can be approximated pointwise by a sequence of simple functions within the same class, so, since $C[\theta]$ is continuous for $\theta \in V$ (*cf.* Westcott (1972)), (3.8) follows. □

Examples of disjoint distributions with uniformly close characteristic functions are well known in Fourier series theory (for probabilistic examples, see Chung[1] (1968), Ex. 6.3.12, or Dudley (1968), Section 4), so that a converse to Theorem 4.3.3 would clearly require extra conditions. According to Prohorov and Rozanov ((1969), p. 162), convergence "in variation" (which we require) cannot adequately be expressed in terms of characteristic functions, so we abandon this approach in favour of more direct techniques.

Finally, for reference, we quote a known theorem which characterizes complete, ϕ- and strong mixing in terms of rings of events $R(N(I))$ generating the σ-algebras $\sigma(N(I))$ used in the definitions.

THEOREM 4.3.4. *A stationary point process P is*

(a) complete mixing if and only if

$$|P(C \cap D)-P(C)P(D)| \leq \gamma(\tau)P(C)P(D)$$

(b) ϕ-mixing if and only if

$$|P(C \cap D)-P(C)P(D)| \leq \phi(\tau)P(C)$$

(c) strong mixing if and only if

$$|P(C \cap D)-P(C)P(D)| \leq \alpha(\tau)$$

for all $C \in R(N(-\infty, t])$, $D \in R(N(t+\tau, \infty))$, $t \in R$, $\tau > 0$, *where* $\gamma, \phi, \alpha : [0, \infty) \to [0, 1]$ *are monotone decreasing functions satisfying* $\lim_{\tau\to\infty} \gamma(\tau) = \lim_{\tau\to\infty} \phi(\tau) = \lim_{\tau\to\infty} \alpha(\tau) = 0$.

Proof. *(b)* is proved on p. 167 of Billingsley (1968), and *(a)* and *(c)* are proved similarly. □

As far as we know, Theorem 4.3.4 cannot usefully be weakened. For example, $R(N(-\infty, t])$ may, in *(c)*, be replaced by a class S_1 such that

$R(N(-\infty, t]) = S_1 \cup \left\{C^c : C \in S_1\right\}$, and similarly for $R(N(t+\tau, \infty))$, but we know of no stronger reduction.

The characterization Theorem 4.3.4 is as it stands not particularly useful to us. It needs to be supplemented in particular cases by characterizations of the sets in the classes $R(N(-\infty, t])$ and $R(N(t+\tau, \infty))$ so that complete, ϕ- , or strong mixing can be proved on these classes.

[1] I thank Dr C.C. Heyde for this reference.

Examples occur in the next section.

4.4. Strong mixing of cluster processes

In this section we develop a technique which allows us decide whether a cluster process is mixing in some sense if its centre process and the process of subsidiaries are mixing in the same sense. The most interesting results occur in the case in which the centre process is strongly mixing, and hence we will introduce the ideas in that setting.

Let the centre process of a cluster process be a measure P_1 on $(N, \sigma(N))$ and the process of clusters be a measure P_2 on $(N^R, \sigma(N)^R)$. Here $\sigma(N)^A$ for any given $A \in \mathcal{B}(R)$ is a product σ-field on N^R, i.e., the smallest σ-field such that the maps $N^R \to R^k$ $(k \in Z_+)$ defined by $\{N_v\} \to (N_{v_1}(B_1), \ldots, N_{v_k}(B_k))$, $B_i \in \mathcal{B}(R)$, $v_i \in A$, $1 \le i \le k$, are measurable. Also we will denote members of N^R by N.

Let N_∞ be the set of non-negative integer or infinite valued measures on R which may be infinite on bounded Borel sets. Define $\eta_c : N \times N^R \to N_\infty$ by

$$\eta_c(N, \{N_v\}) = \int N_v(\cdot - v)dN(v) . \tag{4.1}$$

Throughout the following, we will assume that $\eta_c \in N$, $P_1 \times P_2$ - a.s., and also that η_c is measurable with respect to the product σ-field $\sigma(N) \times \sigma(N)^R$ (see Problem 6.5.10). We define the cluster process via (4.1) (i.e., as $P_c \equiv (P_1 \times P_2)\eta_c^{-1}$), rather than as in Chapter 3, to avoid substantial indexing problems. Note that our process therefore differs from the usual model, in that if the centre process has a multiple event of size n at x, then the contribution to the RHS of (4.1) is $nN_x(\cdot - x)$, rather than the sum of n different subsidiaries all having the same origin. Our model and the usual correspond if the centre process is orderly. We point out that if the subsidiaries are i.i.d., the theorems of this section may be proved for non-orderly centre processes in the manner of Kerstan, Matthes and Mecke (1974), p. 326. We prefer the simplicity of our formulation (4.1).

As usual, we require that P_1 be stationary with respect to $T_y : N \to N$ defined by $T_y N(\cdot) = N(\cdot + y)$, and also we will take P_2 stationary with respect to $S_x : N^R \to N^R$ defined by $S_x\{N_v\} = \{N_{v+x}\}$. Hence let $T_y \times S_y : N \times N^R \to N \times N^R$ by specified by $T_y \times S_y(N, \mathrm{N}) = (T_y N, S_y \mathrm{N})$. Then clearly $P_1 \times P_2$ is stationary with respect to $T_y \times S_y$, and, for $C \in \sigma(N)$,

$$\begin{aligned} P_c(T_y C) &= P_1 \times P_2\left\{(N, \mathrm{N}) : \int N_v(\cdot - v)dN(v) \in T_y C\right\} \\ &= P_1 \times P_2\left\{(N, \mathrm{N}) : \int N_v(\cdot - v + y)dN(v) \in C\right\} \\ &= P_1 \times P_2\left\{(N, \mathrm{N}) : \int N_{v+y}(\cdot - v)dN(v+y) \in C\right\} \\ &= P_1 \times P_2\{T_y \times S_y(\eta_c \in C)\} \\ &= P_c(C), \end{aligned}$$

so that P_c is also stationary.

Our most relevant results will occur in the case of P_2 having independent increments, i.e., if $C \in \sigma(N)^I$, $D \in \sigma(N)^J$, I, J intervals, $I \cap J = \emptyset$, then $P_2(C \cap D) = P_2(C)P_2(D)$.

Finally, before proving our theorem, we define an event G_τ in $\sigma(N) \times \sigma(N)^R$ as follows: for an arbitrary fixed $t \in R$,

$$G_\tau = \left\{(N, \mathrm{N}) : \int_{-\infty}^{t+2\tau} N_v(t+3\tau-v, \infty)dN(v) + \int_{t+\tau}^{\infty} N_v(-\infty, t-v]dN(v) = 0\right\}. \quad (4.2)$$

This event imposes restrictions on the extent of interactions of the subsidiary processes. Its occurrence is sufficient to allow the properties of the centre process to dominate, since it effectively means that events of the cluster process within $(-\infty, t)$ and $(t+3\tau, \infty)$ only interact via the centre process if P_2 has independent increments (see Problem 6.5.1).

THEOREM 4.4.1. *If* P_1 *is strongly mixing with rate* $\alpha(\tau)$, P_2 *has independent increments, and* $P_1 \times P_2(G_\tau) \to 1$ $(\tau \to \infty)$, *then* P_c *is strongly mixing with rate* $\alpha(\frac{1}{3}\tau) + 6P_1 \times P_2(G^c_{\frac{1}{3}\tau})$.

Proof (see Problem 6.5.13). Let t be an arbitrary but fixed number in R . We will denote by $\sigma^t_{-\infty}$ and σ^∞_t the product σ-field $\sigma(N(I)) \times \sigma(N)^I$ for $I = (-\infty, t]$ and $I = (t, \infty)$ respectively. We will say that $P_1 \times P_2$ is strongly mixing if there exists $\psi : [0, \infty) \to [0, 1]$ such that $\lim_{\tau\to\infty} \psi(\tau) = 0$, and

$$|P_1 \times P_2(C \cap D)-P_1 \times P_2(C)P_1 \times P_2(D)| \leq \psi(\tau) \tag{4.3}$$

if $C \in \sigma^t_{-\infty}$ and $D \in \sigma^\infty_{t+\tau}$.

We first prove that $P_1 \times P_2$ is strongly mixing with rate $\alpha(\tau)$. Consider the class of sets in $\sigma^t_{-\infty}$ of the form $C = \sum_{i=1}^n C_{1i} \times C_{2i}$, $C_{1i} \in \sigma(N(-\infty, t])$, $C_{2i} \in \sigma(N^{(-\infty,t]})$ and $n \in Z_+$, where $\{C_{2i}\}$ are disjoint. Then this class forms a ring, for if $F = \sum_{j=1}^m F_{1j} \times F_{2j}$, then

$$C \cup F = \sum_{i=1}^n \left[C_{1i} \times \left(C_{2i} - \bigcup_{j=1}^m F_{2j}\right)\right] + \sum_{j=1}^m \left[F_{1j} \times \left(F_{2j} - \bigcup_{i=1}^n C_{2i}\right)\right]$$

$$+ \sum_{i=1}^n \sum_{j=1}^m [(C_{1i} \cup F_{1j}) \times (C_{2i} \cap F_{2j})] ,$$

$$C - F = \sum_{i=1}^{n} \sum_{j=1}^{m} [C_{1i} \times (C_{2i}-F_{2j})] + \sum_{i=1}^{n} \sum_{j=1}^{m} [(C_{1i}-F_{1j}) \times (C_{2i} \cap F_{2j})]$$

are in the same class. The same of course holds for sets in $\sigma_{t+\tau}^{\infty}$. By Theorem 4.3.4 (which holds more generally than for point processes), we only need prove $P_1 \times P_2$ strongly mixing for sets in these classes. Hence, with C as above, and $D = \sum_{j=1}^{m} D_{1j} \times D_{2j} \in \sigma_{t+\tau}^{\infty}$,

$$|P_1 \times P_2(C \cap D) - P_1 \times P_2(C)P_1 \times P_2(D)|$$

$$= \left| \sum_{i=1}^{n} \sum_{j=1}^{m} [P_1(C_{1i} \cap D_{1j}) - P_1(C_{1i})P_1(D_{1j})]P_2(C_{2i})P_2(D_{2j}) \right|$$

$$\le \alpha(\tau) \ . \tag{4.4}$$

For any set $F \in \sigma(N)$, let $F_\eta \equiv \{\eta_c \in F\} \in \sigma_{-\infty}^{+\infty}$. Hence if $C \in \sigma((-\infty, t])$, then

$$C_\eta \cap G_\tau = \left\{(N, \mathrm{N}) : \int_{-\infty}^{+\infty} N_v(\cdot - v)d\mathrm{N}(v) \in C\right\} \cap G_\tau$$

$$= \left\{(N, \mathrm{N}) : \int_{-\infty}^{t+\tau} N_v(\cdot - v)d\mathrm{N}(v) \in C\right\} \cap G_\tau$$

$$= C'_\eta \cap G_\tau ,$$

say, where $C'_\eta \in \sigma_{-\infty}^{t+\tau}$. Correspondingly, for $D \in \sigma(N(t+3\tau, \infty))$, $D_\eta \cap G_\tau = D'_\eta \cap G_\tau$, where $D'_\eta \in \sigma_{t+2\tau}^{\infty}$. Thus

$$|P_c(C \cap D)-P_c(C)P_c(D)|$$

$$= |P_1 \times P_2(C_\eta \cap D_\eta)-P_1 \times P_2(C_\eta)P_1 \times P_2(D_\eta)|$$

$$\leq |P_1 \times P_2(C_\eta \cap D_\eta \cap G_\tau)-P_1 \times P_2(C_\eta \cap G_\tau)P_1 \times P_2(D_\eta \cap G_\tau)| + 3P_1 \times P_2(G_\tau^c)$$

$$= |P_1 \times P_2(C'_\eta \cap D'_\eta \cap G_\tau)-P_1 \times P_2(C'_\eta \cap G_\tau)P_1 \times P_2(D'_\eta \cap G_\tau)| + 3P_1 \times P_2(G_\tau^c)$$

$$\leq |P_1 \times P_2(C'_\eta \cap D'_\eta)-P_1 \times P_2(C'_\eta)P_1 \times P_2(D'_\eta)|$$

$$+ P_1 \times P_2[(C'_\eta \cap D'_\eta) \cup G_\tau] + P_1 \times P_2(C'_\eta \cup G_\tau) + P_1 \times P_2(D'_\eta \cup G_\tau)$$

$$- 3P_1 \times P_2(G_\tau) + 3P_1 \times P_2(G_\tau^c)$$

$$\leq \alpha(\tau) + 6P_1 \times P_2(G_\tau^c) \ . \qquad \square$$

It is clear that the same techniques will yield

COROLLARY 4.4.2. *(i) If* P_1 *is ergodic and* P_2 *weakly mixing, then* P_c *is ergodic. Similarly, if* P_1 *is weakly mixing and* P_2 *ergodic,* P_c *is ergodic.*

(ii) P_c *is weakly mixing if both* P_1 *and* P_2 *are.*

(iii) P_c *is mixing if both* P_1 *and* P_2 *are.*

(iv) If P_1 *is strongly mixing with rate* $\alpha(\tau)$ *and* P_2 *is completely mixing with rate* $\gamma(\tau)$ *and* $P_1 \times P_2(G_\tau) \to 1$ $(\tau \to \infty)$ *, then* P_c *is strongly mixing with rate*

$$\alpha(\tfrac{1}{3}\tau) + \gamma(\tfrac{1}{3}\tau) + 6P_1 \times P_2(G^c_{\frac{1}{3}\tau}) \ .$$

The same conclusion holds if P_1 *is completely mixing with rate* $\alpha(\tau)$ *and* P_2 *is strongly mixing with rate* $\gamma(\tau)$.

(v) If P_1 *is* ϕ*-mixing and* P_2 *completely mixing with rate* $\gamma(\tau)$ *, and the subsidiary processes are* P_2 *- a.s. bounded (i.e.,*

$P_2\{N : N_1(K^c) = 0\} = 1$ *for some bounded interval* $K: |K| = \delta$ *), then* P_c

is ϕ_c*-mixing, where, for* $\tau \geq 0$,

$$\phi_c(\tau) = \phi(\tau-\delta) + \gamma(\tau-\delta) \ , \quad \tau \geq \delta \ ,$$
$$= 1 \qquad\qquad\qquad , \quad \tau < \delta \ .$$

Alternatively, if P_1 *is completely mixing* (*rate* $\gamma(\tau)$) *and* P_2 *is* ϕ*-mixing and the subsidiaries are bounded as above, then* P_c *is* ϕ_c*-mixing.*

(vi) If P_1 *is completely mixing with rate* $\gamma_1(\tau)$ *and the subsidiary processes are completely mixing with rate* $\gamma_2(\tau)$ *and are* P_2 *- a.s. bounded as in (v), then* P_c *is completely mixing with rate* $\gamma_c(\tau)$ *, where*

$$\gamma_c(\tau) = \gamma_1(\tau-\delta) + \gamma_2(\tau-\delta) \ , \quad \tau \geq \delta \ ,$$
$$= 1 \qquad\qquad\qquad , \quad \tau < \delta \ .$$

Proof (see Problem 6.5.8). We prove only the first half of *(iv)*, the rest being similar. Clearly all that is needed is an extension of the argument at (4.4). So

$$\begin{aligned}
&|P_1 \times P_2(C \cap D)-P_1 \times P_2(C)P_1 \times P_2(D)| \\
&\quad = \Big|\sum_{i,j}\big[P_1(C_{1i} \cap D_{1j})-P_1(C_{1i})P_1(D_{1j})\big]P_2(C_{2i} \cap D_{2j}) \\
&\qquad\qquad + \sum_{i,j} \big[P_2(C_{2i} \cap D_{2j})-P_2(C_{2i})P_2(D_{2j})\big]P_1(C_{1i})P_1(D_{1j})\Big| \\
&\quad \leq \alpha(\tau) + \gamma(\tau) \ . \qquad \square
\end{aligned}$$

Unfortunately the strong mixing rate in Theorem 4.4.1 involves $P_1 \times P_2(G_\tau)$, a quantity which is very difficult to calculate without further knowledge of the structure of the process. Of course, Chebyshev's inequality yields the crude upper bound for $P_1 \times P_2(G_\tau^c)$ of (with $m \equiv E_1N(0, 1] < \infty$)

$$m\int_\tau^\infty \{E_2N_1(-\infty, -u]+E_2N_1(u, \infty)\}du \ ,$$

but by arguing more precisely we can discover a finer condition involving only the component processes which guarantees $P_1 \times P_2(G_\tau) \to 1 \quad (\tau \to \infty)$.

Define

$$L(x) = P_2\{N : N_1(-\infty, x] > 0\} \ , \quad R(x) = P_2\{N : N_1(x, \infty) > 0\} \ . \tag{4.5}$$

Note that N_1 could be replaced by any N_l because of the stationarity of P_2.

LEMMA 4.4.3. *A sufficient stochastic condition for* $P_1 \times P_2(G_\tau) \to 1$ $(\tau \to \infty)$ *is*

$$\int_a^\infty L(a-v)dN(v) + \int_{-\infty}^a R(a-v)dN(v) < \infty \ , \quad P_1 \ - a.s. \tag{4.6}$$

for at least one $a \in R$. *If* P_2 *has independent increments, and there is a uniform bound on the multiplicity of events of the centre process, then this condition is also necessary.*

Remark. This result overlaps with Lemma 3.3.1, but for clarity, we give all details.

Proof. In (4.2), we may write $G_\tau = G_{1\tau} \cap G_{2\tau}$, where

$$G_{1\tau} = \left\{\int_{-\infty}^{t+2\tau} N_v(t+3\tau-v, \infty)dN(v) = 0\right\} \ , \quad G_{2\tau} = \left\{\int_{t+\tau}^{\infty} N_v(-\infty, t-v]dN(v) = 0\right\} .$$

It is necessary and sufficient for $P_1 \times P_2(G_\tau) \to 1$ $(\tau \to \infty)$ that $P_1 \times P_2(G_{1\tau}) \to 1$ and $P_1 \times P_2(G_{2\tau}) \to 1$ $(\tau \to \infty)$. Let

$$G'_{1\tau} = \left\{\int_{-\infty}^{t-\tau} N_v(t-v, \infty)N(dv) = 0\right\} .$$

By stationarity, $P_1 \times P_2(G'_{1\tau}) = P_1 \times P_2(G_{1\tau})$. So let $t \in R$, $\tau_1 > 0$ be arbitrary and fixed. Then note that, for given $N \in N$,

$$\left\{N : \int_{-\infty}^{t-\tau_1} N_v(t-v, \infty)N(dv) < \infty\right\} = \bigcup_{\tau>\tau_1} \left\{N : \int_{-\infty}^{t-\tau} N_v(t-v, \infty)N(dv) < \infty\right\}$$

$$= \lim_{\tau\to\infty} \left\{N : \int_{-\infty}^{t-\tau} N_v(t-v, \infty)N(dv) = 0\right\} \tag{4.7}$$

and this limit is monotone $\uparrow$. Hence, writing $F(v) = \{N : N_v(a-v, \infty) > 0\}$

$$\int_{-\infty}^{a} R(a-v)N(dv) < \infty \qquad P_1 \text{ - a.s. for any fixed } a \in R$$

$$\Longleftrightarrow \int_{-\infty}^{t-\tau_1} R(t-v)N(dv) < \infty \qquad P_1 \text{ - a.s.}$$

$$\Longleftrightarrow \sum_{t_i(N)\le t-\tau_1} P_2\{F(t_i)\} < \infty \qquad P_1 \text{ - a.s.}$$

$$\Rightarrow P_2\{F(t_i) \text{ i.o., } t_i \le t-\tau_1\} = 0 \qquad P_1 \text{ - a.s.} \tag{4.8}$$

by the Borel-Cantelli Lemma,

$$\Longleftrightarrow P_2\left\{\mathrm{N} : \int_{-\infty}^{t-\tau_1} N_v(t-v, \infty)N(dv) < \infty\right\} = 1 \qquad P_1 \text{ - a.s.,} \tag{4.9}$$

since the N_v's are a.s. finite,

$$\Longleftrightarrow P_2\left\{\mathrm{N}: \int_{-\infty}^{t-\tau} N_v(t-v, \infty)N(dv) = 0\right\} \to 1 \quad (\tau \to \infty) \qquad P_1 \text{ - prob.,} \tag{4.10}$$

by (4.7),

$$\Longleftrightarrow P_1 \times P_2\left\{(N, \mathrm{N}) : \int_{-\infty}^{t-\tau} N_v(t-v, \infty)N(dv) = 0\right\} \to 1 \qquad (\tau \to \infty) , \tag{4.11}$$

by dominated convergence,

$$\Longleftrightarrow P_1 \times P_2\{G_{1\tau}\} \to 1 \qquad (\tau \to \infty) .$$

In the forward direction, we have (4.9) $\Rightarrow$ (4.10) by (4.7), but in the reverse direction we are using the fact that if random variables $X_n \to X$ a.s., and $X_n \to Y$ in probability, then $X = Y$ a.s.. Also (4.11) $\Rightarrow$ (4.10) because in (4.11) we have convergence in first mean. Finally, observe that if the centre process is orderly, and P_2 has independent increments, then the reverse implication in (4.8) is immediate, since $\{F(t_i)\}$ are then independent events. If the centre process is non-orderly (see Problem 6.5.1) then we can find a subsequence $\{t_i'\}$ of $\{t_i\}$ such that $t_i' \ne t_j'$, $i \ne j$ by counting multiple events as one. If there is an upper bound M on the multiplicity of centre points, then

$$M \sum_{t_i' \le t-\tau_1} P_2\{F(t_i')\} \ge \sum_{t_i \le t-\tau_1} P_2\{F(t_i)\} = \infty ,$$

so that the Borel-Cantelli Lemma still applies to give $P_2\{F(t_i') \text{ i.o.}\} = 1$ and hence $P_2\{F(t_i) \text{ i.o.}\} = 1$.

In a similar manner we can handle the other half of (4.6). □

COROLLARY 4.4.4. *A sufficient condition for* $P_1 \times P_2(G_\tau) \to 1$ *is*

$$\int_0^\infty \{L(-t)+R(t)\}dt < \infty \tag{4.12}$$

if the centre process has $m = E_1N(0, 1] < \infty$.

Proof. One proof is to take expectations in (4.6), but a more instructive proof is as follows: clearly

$$P_1 \times P_2(G_\tau^c) \le P_1 \times P_2(G_{1\tau}'^c) + P_1 \times P_2(G_{2\tau}^c) ,$$

$$P_1 \times P_2(G_{1\tau}'^c) = \int_N P_2\left\{\bigcup_{t_i \le t-\tau} F^c(t_i(N))\right\}dP_1(N)$$

$$\le \int_N \int_{-\infty}^{t-\tau} P_2\{F^c(v)\}dN(v)dP_1(N)$$

$$= m \int_\tau^\infty R(v)dv .$$

Similarly for $P_1 \times P_2(G_{2\tau}^c)$. □

It would appear from Lemma 4.4.3 and Corollary 4.4.4 that (4.12) is a fairly fine condition for $P_1 \times P_2(G_\tau) \to 1$. This is confirmed in the next result, in which we interpret (4.12) in special cases (see Lawrance (1972)).

LEMMA 4.4.5. *(a) If the centre process* P_1 *is Poisson, and* P_2 *has independent increments,* (4.12) *is also necessary.*

(b) If the subsidiary structure is Neyman-Scott with F *denoting the distribution of each point from the centre and* S *the number of points per*

subsidiary, (4.12) *becomes*

$$\sum_{k=0}^{\infty} Pr\{S = k\} \int_0^{\infty} \left[1-F(t)^k\right] dt < \infty \, . \tag{4.13}$$

(c) If the subsidiary structure is Bartlett-Lewis with inter-epoch distribution F and S points per subsidiary, (4.12) *becomes*

$$\sum_{k=0}^{\infty} Pr\{S = k\} \int_0^{\infty} \left[1-F^{k*}(t)\right] dt < \infty \, . \tag{4.14}$$

Proof. We prove only *(a)*. The random process $\{V_x\}$ in R^R specified by $V_x = \inf\{t : N_x(t, \infty) = 0\}$ clearly has independent increments. Note that

$$\begin{aligned} G'_{1\tau} &= \left\{\int_{-\infty}^{t-\tau} N_x(t-x, \infty) dN(x) = 0\right\} \\ &= \left\{\int_{-\infty}^{t-\tau} 1_{\{N_x(t-x,\infty)>0\}} dN(x) = 0\right\} \\ &= \left\{\int_{-\infty}^{t-\tau} 1_{\{V_x>t-x\}} dN(x) = 0\right\} , \end{aligned} \tag{4.15}$$

so that calculating $P_1 \times P_2\left(G_{1\tau}\right)$ becomes equivalent to determining the probability of (4.15) for a Poisson process (of rate λ, say) subjected to i.i.d. translations V. This is easily calculated using the techniques of Milne (1970) to be

$$P_1 \times P_2\left(G_{1\tau}\right) = \exp\left(-\lambda \int_{\tau}^{\infty} R(t) dt\right) . \tag{4.16}$$

Similarly, we may calculate

$$P_1 \times P_2\left(G_{2\tau}\right) = \exp\left(-\lambda \int_{\tau}^{\infty} L(-t) dt\right) . \qquad \square \tag{4.17}$$

4.5. Conclusions

Since a Poisson process subjected to i.i.d. translations is also Poisson, the conditions of bounded subsidiaries for ϕ-mixing and complete mixing and of $P_1 \times P_2(G_\tau) \to 1$ for strong mixing are not also necessary for preservation of the mixing property. However, the counter-examples contained in Section 4.2 suggest that the clustering operation may not preserve ϕ-mixing under fairly wide circumstances. The conditioning event $\{N(-h, 0] = n\}$ implies for large n that there are a large number of centres in the neighbourhood of $(-h, 0]$ and hence that $\{N(\tau, \tau+h] = 0\}$ will have a reduced probability of occurrence. Hence we may conjecture that the conditions $R(x) > 0$ for all finite x, and $P_2\{N_1[0, \infty) > 1\} > 0$ may be sufficient for a cluster process not to be ϕ-mixing.

We should also remark on the applicability of Theorem 4.4.1 to limit laws. All such theorems (Oodaira and Yoshihara (1971a, b), (1972), Heyde and Scott (1973), Heyde (1974)) seem to impose on the strong mixing rate $\alpha'(\tau)$ the condition

$$\int_0^\infty \alpha'(\tau)d\tau < \infty \, , \tag{5.1}$$

if not the stronger condition in which the integrand is raised to the power $(\delta/(2+\delta))$, for some $\delta > 0$. As in the proof of Corollary 4.4.4, we have the following upper bound for the strong mixing rate $\alpha_c(\tau)$ of the cluster process:

$$\alpha_c(3\tau) \le \alpha(\tau) + m \int_\tau^\infty [L(-t)+R(t)]dt \, . \tag{5.2}$$

(Write $a(\tau)$ for the second term in the RHS of (5.2). If the centre process is Poisson, $\alpha(\tau) = 0$, and $\alpha_c(3\tau) = O(a(\tau))$, but $\neq o(a(\tau))$, when the subsidiaries are i.i.d., by (4.16) and (4.17).) Thus, (5.1) becomes

$$\int_0^\infty \{\alpha(t)+t[L(-t)+R(t)]\}dt < \infty \, , \tag{5.3}$$

which is stronger than (3.6) or (for processes with right-hand clusters) (4.2.5) of Chapter 3, so that it does not appear promising to prove limit laws for cluster processes via strong mixing theorems. However, we point out that strong mixing theorems are potentially able to provide laws of the iterated logarithm for unbounded double-sided clusters, which Theorem 3.4.8 was unable to do.

CHAPTER 5

RENEWAL CHARACTERIZATIONS OF POINT PROCESSES

5.1. Introduction

Problems of characterization of point process systems arise, amongst other reasons, from the fact that very few classes of point processes can be handled reasonably; renewal processes, after Poisson processes, are prime examples of those which are tractable. Consequently it is pertinent to study characterizations via renewal processes. It should be pointed out, though, that such characterizations are not particularly useful unless complemented by some idea of "robustness". For example, it is known that the superposition of n i.i.d. stationary point processes is approximately Poisson for large n, and it is believed (Conjecture 5.3.1 below) that the superposition of n i.i.d. stationary point processes is a renewal process if and only if all processes are Poisson. Thus, although characterization problems are of mathematical interest, they are not necessarily of such practical importance. A suitable definition of "robustness" (see Problem 6.6.1) may be difficult to come by, however. A possibility is to replace the renewal characterizations here by processes with uncorrelated inter-epoch times (Problems 6.6.2, 6.6.4). Alternatively, we may ask for the Prohorov distance ((5.4) of Chapter 1) between the point process in question and a renewal (or Poisson) process of the same intensity. The context will probably decide the appropriateness or otherwise of any particular concept.

In this chapter we tackle two characterization problems: the first in a queueing situation, and the second an example concerning superpositions of point processes.

5.2. Characterizing the finite capacity $GI/M/1$ queue with renewal output

5.2.1. INTRODUCTION AND SUMMARY

In tandem queueing systems in which the output of one queue becomes the input of the next, it is computationally convenient to know when this output is a renewal process. The problem of characterizing outputs of queueing systems has been of interest since Burke (1956) proved that the stationary $M/M/s$ queue has Poisson output; a summary of some of the work in the problem is available in Daley (1975). The object of this section is to complete the proof of the following results for certain single-server queueing systems with waiting room capacity N, i.e., for certain $GI/G/1/N$ systems.

THEOREM 5.2.1. *The only stationary* $GI/M/1/N$ $(0 \le N \le \infty)$ *queueing systems with renewal output are the* $M/M/1/0$ *and stable (traffic intensity* < 1 *)* $M/M/1/\infty$ *systems.* □

THEOREM 5.2.2. *The only stationary* $GI/M/1/N$ $(0 \le N \le \infty)$ *queueing systems with adjacent departure intervals independent are the* $M/M/1/0$ *and stable* $M/M/1/\infty$ *systems.* □

Theorem 5.2.2 is, of course, stronger than Theorem 5.2.1, which, however, is of more interest to us, in that it demonstrates the non-conservation of a local property (that of being renewal) of a point process when subjected to a sufficiently rough transformation.

Daley ((1968), (1974)) has completely characterized the stationary $GI/G/1/0$ and $GI/M/1/\infty$ systems with renewal output; our proof of Theorem 5.2.1 handles only the intermediate situations $(1 \le N < \infty)$. Complete renewal characterization results have now been obtained for the stationary $M/G/1/N$ $(0 \le N \le \infty)$ (Disney *et al.* (1973)) and $GI/M/1/N$ systems; the solution techniques involve using the imbedded Markov chain of queue lengths at service completions or at arrivals. The general problem of characterizing the $GI/G/1/N$ queue with renewal departure process (Problem

6.6.3) seems very difficult (but see Daley (1975)).

The point of Theorem 5.2.2 is the contrast with the $M/G/1/N$ $(1 \le N < \infty)$ situation, in which Daley and Shanbhag (1975) show that there are service time distribution functions G for which adjacent departure intervals are independent, although the output is not renewal. We will only prove Theorem 5.2.2 for $1 \le N \le \infty$; Daley (1974) handles the $GI/G/1/0$ situation (see the analysis prior to his Theorem 1): adjacent departure intervals are independent if and only if the output is renewal.

Throughout this section we will assume without further comment that our random variables are all defined on a common probability space (Ω, F, Pr) .

Consider a single server queueing system with independent service times $\{S_n\}$, $(n = \ldots -1, 0, 1, \ldots)$, where $Pr\{S_n \le x\} = B(x) = 1 - e^{-\mu x}$, $x \ge 0$ and $\mu > 0$. Potential customers arrive at successive epochs $\{t_i\}$ $(\ldots \le t_0 = 0 < t_1 \le \ldots)$ of a renewal process, with inter-epoch times $\{\Pi_i = t_i - t_{i-1}\}$, $Pr\{\Pi_i \le x\} = A(x)$, $A(0+) = 0$. An arrival finding a queue with N customers waiting does not enter the system, and is not considered here as part of the output (compare Boes, (1969)). The nth served arrival finds Q_n customers in the system (queue and service), waits for a time W_n until service, and is served for a time S_n .

Let q_i denote the number of customers in the system the instant before arrival epoch t_i , and $W^{(i)}$ the waiting time of the arrival at t_i if $q_i \le N$. If $\nu(n)$ is defined to be the index of the nth served arrival, then

$$\nu(0) = \sup\left\{i : t_i + W^{(i)} \le 0,\ q_i \le N\right\}$$

$$\nu(n) = \inf\{i > \nu(n-1) : q_i \le N\} , \quad n > 0 ,$$

$$\nu(n) = \sup\{i < \nu(n+1) : q_i \le N\} , \quad n < 0 .$$

Observe that $\nu(n+1) = \nu(n) + 1$ if $Q_n < N$, while when $Q_n = N$,

$$\nu(n+1) = \inf\left\{i : \sum_{j=\nu(n)+1}^{i} \Pi_j > S'_{n-N}\right\}$$

where S'_{n-N} is the time elapsing from the $\nu(n)$th arrival epoch to the completion of servicing of the customer being serviced. Since the service time distribution function $B(\cdot)$ is exponential, $B(\cdot)$ is also the distribution function of S'_{n-N}.

Denote the stationary state probabilities of the number of customers in the system the instant before an arrival occurs by $\{\pi_j\}$, $0 \le j \le N+1$, i.e., $\pi_j = Pr\{q_i = j\}$. These probabilities are known to exist (e.g., Keilson (1966)). Clearly

$$Pr\{Q_n = j\} = \pi_j/(\pi_0 + \dots + \pi_N) = \pi_j/(1-\pi_{N+1}) , \quad 0 \le j \le N , \qquad (2.1.1)$$

and

$$Pr\{W_n \le x\} = B^{Q_n *}(x) , \quad x \ge 0 , \qquad (2.1.2)$$

where

$$B^{0*}(x) = 1 \quad \text{if} \quad x \ge 0 ,$$
$$= 0 \quad \text{if} \quad x < 0 ,$$

and $B^{j*}(x)$ is the j-fold convolution of B with itself.

Always, then, the time between arrivals of the nth and $(n+1)$th customers to be served is

$$\Pi'_n \equiv \Pi_{\nu(n)+1} + \dots + \Pi_{\nu(n+1)} . \qquad (2.1.3)$$

Let

$$I_{n+1} \equiv \max(0, \Pi'_n - W_n - S_n) \qquad (2.1.4)$$

denote the idle time for the server between the completion of the nth service and the beginning of the $(n+1)$th service. We define the output process of the queueing system via the sequence $\{D_n\}$ of inter-departure times, these being almost surely positive random variables given by

$$D_n = S_n + I_n \qquad (2.1.5)$$

where S_n and I_n are independent, since the arrival process and service times are independent.

5.2.2. THE DISTRIBUTION OF INTER-DEPARTURE TIMES

From the definitions, it is easily seen that

$$\left(1-\pi_{N+1}\right) Pr\{I_{n+1} \le y\} = \sum_{j=0}^{N} Pr\{I_{n+1} \le y \mid Q_n = j\}\pi_j \ . \tag{2.2.1}$$

After some algebra, we find

$$Pr\{I_{n+1} \le y \mid Q_n = j\} = \int_0^\infty A(u+y)dB^{(j+1)*}(u) \ , \quad 0 \le j \le N-1 \tag{2.2.2}$$

$$Pr\{I_{n+1} \le y \mid Q_n = N\}$$

$$= \int_0^\infty dB(u) \int_0^\infty Pr\{\Pi'_n \le y+u+v \mid S'_{n-N} = u\} dB^{N*}(v)$$

$$= \int_0^\infty dB(u) \int_0^\infty dB^{N*}(v) \int_u^{y+u+v} [1-A(y+u+v-w)]dH(w) \tag{2.2.3}$$

where

$$H(x) = \sum_{i=1}^{\infty} A^{i*}(x) = A(x) + \int_0^x A(x-u)dH(u) \tag{2.2.4}$$

is the renewal function of the arrival process. By interchanging the order of integration in (2.2.3), we can replace the renewal function by the Laplace-Stieltjes transform $\alpha(\mu)$ where

$$\alpha(s) = \int_0^\infty e^{-st}dA(t) \ , \quad Re(s) \ge 0 \ . \tag{2.2.5}$$

Thus

$$Pr\{I_{n+1} \le y \mid Q_n = N\} = [1-\alpha(\mu)]^{-1}\left[\int_0^\infty A(u+y)dB^{(N+1)*}(u) - \alpha(\mu)\right] \ . \tag{2.2.6}$$

Putting

$$\begin{aligned} \pi'_j &= \pi_j \ , && 0 \le j \le N-1 \\ \pi'_N &= \pi_N/[1-\alpha(\mu)] = \pi_N + \pi_{N+1} && \end{aligned} \tag{2.2.7}$$

we obtain, for $y \ge 0$,

$$\left(1-\pi_{N+1}\right)Pr\{I_{n+1} \leq y\} = \sum_{j=0}^{N} \pi'_j \int_0^\infty A(u+y)dB^{(j+1)*}(u) - \pi'_N\alpha(\mu) \ . \quad (2.2.8)$$

Using similar techniques, we can obtain the joint distribution of $\left(D_n, I_{n+1}\right)$. If $Q_n > 0$, then $I_n = 0$, since S_{n-1} has not yet been completed, and by (2.1.2), (2.1.3), (2.1.4) and (2.1.5) we can discuss the joint distribution of $\left(D_n, I_{n+1}\right)$ in terms of the arrival of the nth served customer. Otherwise we have to consider the $(n-1)$th served arrival. For $x \geq 0$, $y \geq 0$,

$$\left(1-\pi_{N+1}\right)Pr\{D_n \leq x, I_{n+1} \leq y\}$$

$$= \sum_{j=1}^{N} \pi_j Pr\{D_n \leq x, I_{n+1} \leq y \mid Q_n = j\}$$

$$+ \sum_{j=0}^{N} \pi_j Pr\{D_n \leq x, I_{n+1} \leq y, Q_n = 0 \mid Q_{n-1} = j\}$$

$$= \sum_{j=1}^{N} \pi'_j \int_0^\infty dB^{j*}(u) \int_0^x A(u+v+y)dB(v) - \pi'_N\alpha(\mu)B(x)$$

$$+ \sum_{j=0}^{N} \pi'_j \int_0^\infty dB^{(j+1)*}(u) \int_u^{u+x} dA(v) \int_0^{u+x-v} A(w+y)dB(w) \ . \quad (2.2.9)$$

5.2.3. PROOFS OF THE THEOREMS

A necessary condition for the output to be renewal is that $Pr\{D_n \leq x, D_{n+1} \leq y\} = Pr\{D_n \leq x\}Pr\{D_{n+1} \leq y\}$, for all x, y. By recognising that $\left(D_n, I_{n+1}\right)$ is jointly independent of S_{n+1}, and taking Laplace-Stieltjes transforms, this necessary condition reduces to

$$Pr\{D_n \leq x\}Pr\{I_{n+1} \leq y\} = Pr\{D_n \leq x, I_{n+1} \leq y\} \ . \quad (2.3.1)$$

In particular, for all $\gamma \geq 0$, $\phi \geq -\mu$,

$$\int_0^\infty e^{-\gamma x} dPr\{D_n \leq x\} \int_{0+}^\infty e^{-(\phi+\mu)y} dPr\{I_{n+1} \leq y\}$$

$$= \int_0^\infty e^{-\gamma x} \int_{0+}^\infty e^{-(\phi+\mu)y} d_x d_y Pr\{D_n \leq x, I_{n+1} \leq y\} \ . \quad (2.3.2)$$

Write (2.3.2) as $L(\phi, \gamma) = R(\phi, \gamma)$ say. Let $P_n(\phi, \gamma)$ ($n = 1, 2, 3, 4$) denote functions of ϕ and γ that are in fact polynomials of degree N in ϕ. Then (we omit some algebraic detail; see Section 5.2.4) L and R are expressible in the form

$$\left(1-\pi_{N+1}\right)(-\phi)^{N+1}L(\phi, \gamma) = \alpha(\phi+\mu)P_1(\phi, \gamma) + P_2(\phi, \gamma) , \tag{2.3.3}$$

and

$$\left(1-\pi_{N+1}\right)(\phi-\gamma)(-\phi)^{N}R(\phi, \gamma) = \alpha(\phi+\mu)P_3(\phi, \gamma) + P_4(\phi, \gamma) . \tag{2.3.4}$$

The coefficients of ϕ^i $(0 \le i \le N)$ in each of the polynomials $P_n(\phi, \gamma)$ $(1 \le n \le 4)$ simplify considerably using the properties of the imbedded Markov chain of queue lengths just prior to arrival. Solving the equation $L(\phi, \gamma) = R(\phi, \gamma)$ then leads to the following expression for $\alpha(\phi+\mu)$, valid for $\phi \ge -\mu$, $\gamma > 0$, $\phi \ne \gamma$,

$$\alpha(\phi+\mu) = \frac{-\gamma\psi\pi_{N+1}\mu^{N+1} + \sum_{j=0}^{N-1}(-\phi)^{N-j}\mu^{j+1}\left[\mu\pi_{j+2}(1-\psi)-\gamma\psi\pi_{j+1}\right] + (-\phi)^{N+1}C(\gamma)}{-\gamma\psi\pi_N'\mu^{N+1} + \sum_{j=0}^{N-1}(-\phi)^{N-j}\mu^{j+1}\left[\mu\pi_{j+1}'(1-\psi)-\gamma\psi\pi_j'\right] + (-\phi)^{N+1}D(\gamma)} \tag{2.3.5}$$

where

$$\psi \equiv \psi(\gamma) = E\left(e^{-\gamma D_n}\right) \tag{2.3.6}$$

and $C(\cdot)$ and $D(\cdot)$ are independent of ϕ; we need note only the form of $D(\cdot)$, namely

$$D(\gamma) = -\mu\pi_0\psi + \left(1-\pi_{N+1}\right)\left[(\mu+\gamma)\psi-\mu Pr\{I_n = 0\}\right] . \tag{2.3.7}$$

We see immediately that $C(\gamma) \equiv 0$, since $\alpha(\phi+\mu) \to 0$ as $\phi \to \infty$, but since this yields a complicated functional equation for α, we pass it by. Instead, observe that the term independent of ϕ in the denominator of (2.3.5) cannot vanish for all γ, and that therefore the denominator cannot vanish identically in ϕ.

From (2.2.8) we observe that

$$Pr\{I_n = 0\} = 1 - \pi_0/\left(1-\pi_{N+1}\right) . \tag{2.3.8}$$

We require $\alpha(0) = 1$, which from (2.3.5) is possible only if

$$\psi(\gamma) = m/(m+\gamma) \tag{2.3.9}$$

where (using (2.3.8))

$$m = \mu\left[1-\pi_1/\left(1-\pi_0-\pi_{N+1}\right)\right] . \tag{2.3.10}$$

Taking Laplace-Stieltjes transforms in (2.1.5), and employing (2.3.9), we find

$$Pr\{I_n = 0\} = m/\mu \tag{2.3.11}$$

and (2.3.7) vanishes identically.

Thus, substituting (2.3.9) into (2.3.5), we obtain, for $\phi \geq -\mu$,

$$\alpha(\phi+\mu) = \frac{-m\pi_{N+1}\mu^{N+1}+\sum_{j=0}^{N-1}(-\phi)^{N-j}\mu^{j+1}\left[\mu\pi_{j+2}-m\pi_{j+1}\right]}{-m\pi_N'\mu^{N+1}+\sum_{j=0}^{N-1}(-\phi)^{N-j}\mu^{j+1}\left[\mu\pi_{j+1}'-m\pi_j'\right]} . \tag{2.3.12}$$

Again we set $\alpha(0) = 1$, and (2.3.12) yields

$$m/\mu = \pi_1/\pi_0 . \tag{2.3.13}$$

Substituting (2.3.13) into (2.3.12), we deduce that

$$m/\mu = \pi_{i+1}/\pi_i , \quad 0 \leq i \leq N . \tag{2.3.14}$$

It follows that, for $\phi \geq -\mu$,

$$\alpha(\phi+\mu) = m\pi_{N+1}/\left(m\pi_N'+\phi\pi_{N+1}\right) . \tag{2.3.15}$$

Thus $A(x) = 1 - e^{-mx}$, $x \geq 0$, i.e., the system has a Poisson arrival process.

Finally, we appeal to a known result (Finch (1959)) that the $M/M/1/N$ queueing system $(1 \leq N < \infty)$ never has renewal output, which proves Theorem 5.2.1.

If $N < \infty$, then Theorem 5.2.2 follows immediately from (2.3.15) and Daley and Shanbhag (1975), but Daley (1968) asks for D_n, D_{n+2} to be independent in the $GI/M/1/\infty$ situation. However, a similar analysis to the above yields instead of (2.3.5),

$$\alpha(\phi+\mu) = \frac{\sum_{j=0}^{\infty} (-\mu/\phi)^{j+1}\left[\mu\pi_{j+2}(1-\psi)-\gamma\psi\pi_{j+1}\right]+C(\gamma)}{\sum_{j=0}^{\infty} (-\mu/\phi)^{j+1}\left[\mu\pi_{j+1}(1-\psi)-\gamma\psi\pi_{j}\right]+D(\gamma)}, \tag{2.3.16}$$

valid for $\phi = -\mu$, $\phi > \mu$, $\gamma > \mu$, where

$$D(\gamma) = -\mu\pi_0\psi + \left[(\mu+\gamma)\psi-\mu Pr\{I_n = 0\}\right] . \tag{2.3.17}$$

The same procedure gives rise to $\psi(\gamma) = m/(m+\gamma)$, $(\gamma > \mu)$, and also $m/\mu = \pi_{i+1}/\pi_i$, $i \geq 0$. However, (2.3.16) is now unhelpful, in that the RHS gives $0/0$; similarly for the identity $C(\gamma) \equiv 0$. However, it is not difficult to establish (Daley (1968), Theorem 3) that

$$\psi(\gamma) = \mu(\mu+\gamma)^{-1}[\delta\gamma-\mu(1-\delta)\alpha(\gamma)][\gamma-\mu(1-\delta)]^{-1} ,$$

where δ is the unique root in $0 < \delta < 1$ of $\delta = \alpha\big(\mu(1-\delta)\big)$. Solving for $\alpha(\gamma)$, we obtain

$$\alpha(\gamma) = \left[\gamma^2(m-\mu\delta)-m\mu^2(1-\delta)\right].\left[-\mu^2(m+\gamma)(1-\delta)\right]^{-1}$$

and hence $m = \mu\delta$, since $\alpha(\gamma) \to 0$ as $\gamma \to \infty$. Thus $\alpha(\gamma) = m/(m+\gamma)$, $(\gamma > \mu)$ which establishes $A(\cdot)$ uniquely (Feller (1966), p. 410) and thus the input process must be Poisson. □

5.2.4. SUPPLEMENTS TO SECTION 5.2.3

For reasons of clarity, we outline the transition from (2.3.2) to (2.3.5) in this section, since the steps are purely manipulative. Straightforwardly, for $\phi \geq -\mu$, $\gamma > 0$, $\phi \neq \gamma$,

$$P_3(\phi, \gamma) = -\mu(-\phi)^N\left\{\int_0^{\infty} e^{-\gamma u}F(u, \gamma)dA(u) + \sum_{j=1}^{N} (-\mu/\gamma)^j\pi_j'\right\}$$

$$- (\phi-\gamma) \sum_{j=1}^{N} (-\mu/\gamma)^{j+1}\pi_j' \sum_{k=0}^{j-1} (-\gamma)^{k+1}(-\phi)^{N-k-1} , \tag{2.4.1}$$

where we have used only

$$F(u, \gamma) \equiv \sum_{j=0}^{N} \pi_j' \int_0^{u} e^{\gamma w}dB^{(j+1)*}(w) \tag{2.4.2}$$

and the trivial identity

$$\int_0^t e^{-\mu x}\mu^{j+1}\, x^j/j!\; dx = 1 - e^{-\mu t}\sum_{k=0}^{j} (\mu t)^k/k! \; . \tag{2.4.3}$$

Expanding the final term in (2.4.1) gives

$$-\mu \sum_{j=1}^{N} (-\mu/\gamma)^j \pi_j'(-\phi)^N + \sum_{j=1}^{N} \mu^{j+1}\pi_{j+1}'(-\phi)^{N-j} \tag{2.4.4}$$

so that (2.4.1) simplifies to

$$P_3(\phi,\, \gamma) = -\mu(-\phi)^N \int_0^\infty e^{-\gamma u}F(u,\, \gamma)dA(u) - \sum_{j=1}^{N} \mu^{j+1}\pi_j'(-\phi)^{N-j} \; . \tag{2.4.5}$$

A similar surprising simplification occurs in $P_4(\phi,\, \gamma)$,

$P_4(\phi,\, \gamma) = -\alpha(\gamma+\mu)$ {first term of (2.4.1)}

$$+ (\phi-\gamma) \sum_{j=1}^{N} (-\mu/\gamma)^{j+1}\pi_j' \sum_{k=0}^{j-1} (-\gamma)^{k+1} \int_0^\infty e^{-\mu t} \sum_{l=0}^{k} (-\phi)^{N+l-k-1}\left(t^l/l!\right) dA(t) \tag{2.4.6}$$

and expanding the second term of (2.4.6) we obtain

$$-(\phi/\gamma)^N h(\mu,\, \gamma) + h(\mu,\, \phi)$$

where

$$h(\mu,\, \phi) = \sum_{j=1}^{N}\sum_{k=0}^{j-1} \mu^{j+1}\pi_j'(-\phi)^{N+k-j} \int_0^\infty e^{-\mu t}\left(t^k/k!\right) dA(t) \; . \tag{2.4.7}$$

Denote the transition probabilities of the imbedded Markov chain of the number of people in the queue the instant before arrivals occur by p_{ij} .

Clearly

$$p_{ij} = \int_0^\infty e^{-\mu t}(\mu t)^{i-j+1}/(i-j+1)!\; dA(t) \; , \quad 0 \le i \le N \; , \quad 1 \le j \le i+1 \; ,$$

$$p_{i0} = \int_0^\infty \{1-A(t)\}dB^{(i+1)*}(t) \qquad , \quad 0 \le i \le N \; ,$$

$$p_{N+1,j} = p_{N,j} \qquad , \quad 0 \le j \le N+1 \; .$$

Looking at the coefficient of $(-\phi)^{N-i}$, $1 \le i \le N$, in (2.4.7) we obtain

$$(-\phi)^{N-i} \sum_{j=i}^{N} \pi_j' \mu^{j+1} \int_0^\infty e^{-\mu t}\, t^{j-i}/(j-i)!\, dA(t) = (-\phi)^{N-i} \mu^{i+1} \sum_{j=i}^{N} \pi_j' p_{j,i+1}$$

$$= (-\phi)^{N-i} \mu^{i+1} \pi_{i+1} \,. \qquad (2.4.8)$$

Hence (2.4.6) becomes

$P_4(\phi, \gamma) = -\alpha(\gamma+\mu)$ {first term of (2.4.1)}

$$- (-\phi)^N \sum_{i=1}^{N} \mu^{i+1} \pi_{i+1} \{(-\gamma)^{-i} - (-\phi)^{-i}\} \,. \qquad (2.4.9)$$

With reference to (2.4.5) and (2.4.9) (compare (2.3.7)) note that

$$\mu \int_0^\infty e^{-\gamma u} F(u, \gamma) dA(u) = \sum_{j=0}^{N} \pi_j' \int_0^\infty e^{-\gamma u} dA(u) \int_0^u e^{\gamma w} dB^{(j+1)*}(w)$$

$$= \left(1-\pi_{N+1}\right) \int_{0+}^\infty e^{-\gamma u} dPr\{I_{n+1} \le u\}$$

$$= \left(1-\pi_{N+1}\right) \left[(\mu+\gamma) E\left(e^{-\gamma D_n}\right) - \mu Pr\{I_n = 0\}\right] \,. \qquad (2.4.10)$$

Straightforward manipulation also yields $\left(\text{put } \psi(\gamma) \equiv E\left(e^{-\gamma D_n}\right)\right)$

$$P_1(\phi, \gamma) = \psi(\gamma) \sum_{j=0}^{N} \pi_j' \mu^{j+1} (-\phi)^{N-j}$$

$$P_2(\phi, \gamma) = -\psi(\gamma) \sum_{j=0}^{N} \pi_{j+1} \mu^{j+1} (-\phi)^{N-j}$$

arguing as in (2.4.8). Substituting (2.4.10) and the above expressions for $P_n(\phi, \gamma)$ into (2.3.3) and (2.3.4) then gives (2.3.5), with

$$C(\gamma) = -\alpha(\mu+\gamma)\{\text{first term of } (2.4.1)\} - \mu \sum_{j=1}^{N} (-\mu/\gamma)^j \pi_{j+1} - \mu\psi(\gamma)\pi_1 \,.$$

The expression $C(\gamma) = 0$ only readily yields information on $\alpha(\mu+\gamma)$ for $\gamma > 0$.

It is interesting to note that equation (2.3.5) has its LHS independent of γ, but not its RHS. We have exploited this fact only peripherally in our analysis. A more direct argument is as follows:

There are two possibilities to consider:

(a) the coefficient of ϕ^N in the numerator of (2.3.5) vanishes identically, from which we easily deduce

$$\psi(\gamma) = m/(m+\gamma) \ , \quad m = \mu\pi_2/\pi_1$$

and the analysis proceeds as from (2.3.12);

(b) the coefficient of ϕ^N in the numerator does not vanish identically, and neither therefore can $D(\gamma)$. In this case, since the LHS of (2.3.5) is independent of γ , so is the RHS, and so the ratio of the coefficient of ϕ^N in the numerator to $D(\gamma)$ must be independent of γ , equal to some constant K , say. This leads to

$$\psi(\gamma) = m'/(m'+\gamma) \tag{2.4.11}$$

where $m' = \left[\mu\pi_2+\mu K\left(1-\pi_{N+1}\right)Pr\{I_n = 0\}\right]/\left[\pi_1+\left(1-\pi_{N+1}\right)K\right]$.

But if (2.4.11) is true, then from (2.1.5) and (2.2.8) we find

$$Pr\{I_n = 0\} = m'/\mu = 1-\pi_0/\left(1-\pi_{N+1}\right)$$

and $D(\gamma)$ vanishes identically. Thus case (b) is impossible.

5.3. Point processes whose superposition is a renewal process

In this section we solve a problem in support of the following

CONJECTURE 5.3.1. *n independent and identically distributed stationary point processes η_i $(1 \le i \le n)$ superpose to an orderly stationary renewal process η if and only if all processes are Poisson.* □

Daley (1973a, b) found counter-examples to the corresponding conjecture in the non-identically distributed case, but now conjectures (private communication) only that at least one of the superposed processes must be Poisson (see Problem 6.6.5). The above conjecture, which he proves (1973a) when the summands are renewal (see also Störmer (1969)) is therefore a special case of his. We prove here only that his counter-example (1973a) does not apply to the above conjecture, and we suspect this non-extension anyway since the argument for his Theorem 1 does not reproduce. Since the proof of his Theorem 2 and our theorem are virtually identical, we sketch only enough to point out the differences, and assume complete familiarity with Daley's paper.

THEOREM 5.3.2. *n i.i.d. stationary alternating renewal processes* η_i *superpose to an orderly stationary renewal process* η *if and only if all processes are Poisson.*

Proof. As in Daley (1973a), consider

$$G(x) \equiv E\left[\eta(0,\, x]-mx \mid \eta(\{0\}) > 0\right]$$

$$G(x,\, y) \equiv E\left[\{\eta[-x,\, 0)-mx\}\{\eta(0,\, y]-my\} \mid \eta(\{0\}) > 0\right]$$

where $m = E\eta(0,\, 1]$. Define $G_i(\cdot)$, $G_i(\cdot,\, \cdot)$ for η_i $(1 \le i \le n)$ similarly. It is easily deduced from (9) of Daley (*ibid.*) that

$$G(x,\, y) = G_1(x,\, y) + m\left(1-n^{-1}\right) \int_0^y \left[G_1(x+u)-G_1(u)\right] du \;, \tag{3.1}$$

and immediately it follows that

$$m\left(1-n^{-1}\right) \int_0^y \left[G_1(x+u)-G_1(u)\right] du = K(x)K(y)$$

for some function $K(\cdot)$. Hence Daley's Sections 1, 2, 3 and 5 immediately carry over. Thus if the lifetime distributions of the alternating renewal processes are F_1, F_2 respectively, then with $Q_i(x) \equiv 1 - F_i(x)$, $i = 1, 2$,

$$\text{(i)}\quad Q_i(x) = A_i e^{-c_i' x} + \left(1-A_i\right)e^{-c_i'' x} \;, \quad 0 < c_i' < c_i'' \;,$$

$$0 \le A_i \le c_i''/\left(c_i''-c_i'\right) \;, \quad A_i \ne 1 \;.$$

$$\text{(ii)}\quad Q_i(x) = \left(1+A_i x\right)e^{-c_i x} \;, \quad 0 < A_i \le c_i \;, \tag{3.2}$$

where Q_1, Q_2 cannot both be of the form (ii), nor both of the form (i) if $A_i > 0$, $c_1' = c_2' < c_1'' = c_2''$. If we define

$$R(x) = Pr\{\eta(0,\, x] = 0 \mid \eta(\{0\}) > 0\} \tag{3.3}$$

$$R(x,\, y) = Pr\{\eta[-x,\, 0) = 0,\, \eta(0,\, y] = 0 \mid \eta(\{0\}) > 0\} \;, \tag{3.4}$$

then since η is stationary renewal,

$$R(x, y) = R(x)R(y) \ . \tag{3.5}$$

Hence, if $R_i(\cdot)$, $R_i(\cdot, \cdot)$ are defined similarly,

$$\int_x^\infty \int_y^\infty R_1(u, v)\left[\int_{u+v}^\infty R_1(w)dw\right]^{n-1} dudv$$

$$= \left(m^{n-1}/n^{n+1}\right)\left[\int_x^\infty R_1(u)du \int_y^\infty R_1(v)dv\right]^n \tag{3.6}$$

where $2R_1(x) = Q_1(x) + Q_2(x)$, and $2R_1(x, y) = Q_1(x)Q_2(y) + Q_2(x)Q_1(y)$. If Q_2 is of the form (ii), and Q_1 of (i), then the RHS of (3.6) contains a term in $(xy)^n$, whereas the LHS does not. If Q_1 and Q_2 are both of the form (i), then the LHS consists of sums of terms of the form $\exp\left(-\sum_{l=1}^{n-1} c_{k_l}.(x+y)-c_i x-c_j y\right)$, where c_i is c_1' or c_1'' , and c_j is c_2' or c_2'' , and c_{k_l} is any of c_i', c_i'' , and the RHS is a sum of terms of the form $\exp\left(-\sum_{l=1}^{n} c_{k_l} x - \sum_{s=1}^{n} c_{r_s} y\right)$. But there exists a smallest (or largest) exponential constant c , say, with coefficient A , say. Then the coefficient of $\exp\left(-nc(x+y)\right)$ on the RHS is $(A/2c)^{2n}.\left(m^{n-1}/n^{n+1}\right)$, but zero on the LHS. Since c is the largest or smallest exponential constant, no other sum of n exponential constants can contribute to $\exp\left(-nc(x+y)\right)$. □

CHAPTER 6

UNSOLVED PROBLEMS AND GENERALIZATIONS

6.1. Introduction

In this final chapter we collect together, for easy reference (rather than having them scattered randomly throughout the thesis) some unsolved problems and possible generalizations which relate to the ideas presented in the body of the thesis. Of course, there are many more unsolved problems than there were solved, but only few of those we suggest will fall into that elusive class "interesting and (possibly) resolvable". Some of the rest (probably) defy currently available techniques, while others may be of only borderline interest. The problems from each chapter will be listed approximately in order of occurrence with miscellaneous problems last. Reference to its source will be given at the beginning of any problem whose origin is not specified in its statement. If the formulation is due to another person, then he will be appropriately accredited. We will not attempt to impose our own bias by indicating those we regard as most important. Some partial solutions will be provided.

Before embarking on this project, however, we will present in the next section a problem which is solved, but not in the sense of this thesis, and hence is included in this chapter.

6.2. Identifiability of the cluster structure of a stationary Poisson cluster process

6.2.1. INTRODUCTION

Suppose we know that the centre process of a cluster process is stationary Poisson with finite intensity; can we deduce from a complete record of the cluster process alone the structure of the subsidiary processes? The answer to this question is in the negative when each subsidiary is of unit size, because Poisson processes subjected to independent random translations are themselves Poisson. Milne (1970), Brown (1970) and Ross (1970) have all shown that the service time distribution of the $M/G/\infty$ queue can be identified from a complete input-output record. We answer affirmatively the corresponding question for Poisson cluster processes here, that is, we prove that the structure of a subsidiary process can be identified given realizations of a cluster process and of the Poisson process from which it is derived, without any information being given on the linkage between the two processes. Bendrath (1974) uses completely different techniques (no probability generating function(al)s) in a generalization of Milne's (1970) result. I thank Dr R.K. Milne for pointing out the existence of this alternative generalization and Dr D.J. Daley for finding the reference.

6.2.2. IDENTIFIABILITY WITH GENERAL CLUSTER STRUCTURE

Our solution will be in the spirit of Milne (1970). We shall assume throughout that the subsidiaries are i.i.d., and employ the notation of Section 4.4.

To identify the cluster structure, it is sufficient to specify

$$P_2\{\mathrm{N} : N_1(A_1) = m_1, \ldots, N_1(A_k) = m_k\} \tag{2.2.1}$$

for $A_1, \ldots, A_k$ left-open, right-closed adjacent intervals with rational endpoints, and $k, m_1, \ldots, m_k$ each within Z_+ (Theorem 1.2.2). We need to define a mapping $(i, \eta_c) : N \times N^R \to N \times N_\infty$ (compare (4.1) of Chapter 4) by

$$(i, \eta_c)(N, \mathrm{N}) = \left(N, \int_R N_v(\cdot - v)dN(v)\right), \tag{2.2.2}$$

and define a measure P (with expectation denoted by E) on $(N \times N_\infty, \sigma(N) \times \sigma(N_\infty))$ by $P = (P_1 \times P_2)(i, \eta_c)^{-1}$. We will let N_c stand for a typical member of N_∞. Note that here we are assuming (i, η_c) is measurable with respect to $\sigma(N) \times \sigma(N)^R$ (which will follow from the measurability of η_c defined at (4.1) of Chapter 4). We will also assume that $(i, \eta_c) \in N^2$, $P_1 \times P_2$ - a.s. If $T_y : N \to N$ is defined as usual by $T_y N(\cdot) = N(\cdot + y)$, then it is easy to verify that P is stationary with respect to $T_y \times T_y$, and that the ergodicity of P will follow immediately from the ergodicity of $P_1 \times P_2$, which in turn follows from P_1 being ergodic, and P_2 weakly mixing (*cf.* Theorem 4.4.1).

Let $A \equiv (0, r]$ for a given rational $r > 0$, and $A_i \equiv (r_{i-1}, r_i]$, $r_0 \equiv 0$, r_i rational, $1 \le i \le k$. Setting $\hat{R} = \max(r, r_k)$, we define the sets $D_i \in \sigma(N)^2$ for $i \in Z_+$ by

$$D_i = \{(N, N_c) : N(A+i\hat{R}) = m,\ N_c(A_l+i\hat{R}) = m_l\ (1 \le l \le k)\} = (T_{\hat{R}} \times T_{\hat{R}})^i D_0 ,$$

where m, m_l $(1 \le l \le k)$ are each within Z_+. Clearly $1_{D_i} : N^2 \to \{0, 1\}$ satisfies $E|1_{D_i}| < \infty$, and hence by the pointwise ergodic theorem,

$$k^{-1} \sum_{i=0}^{k-1} 1_{D_i} \to E(1_{D_0}) = P(D_0) , \quad P \text{ - a.s.} \tag{2.2.3}$$

Hence for $z, z_1, \ldots, z_k$ rational, we have identified

$$M(z, z_1, \ldots, z_k) \equiv E\left[z^{N(A)} z_1^{N_c(A_1)} \cdots z_k^{N_c(A_k)}\right] . \tag{2.2.4}$$

By continuity, we have identified $M(z, z_1, \ldots, z_k)$ for all $(z, z_1, \ldots, z_k) \in [0, 1]^{k+1}$. In particular, we have also identified $\log M(z, z_1, \ldots, z_k)$.

Now, it is known for the Poisson cluster process that, for ξ, ξ_c within appropriate function classes (see Westcott (1971), Milne (1971),

p. 49), the probability generating functional of P is

$$G[\xi, \xi_c] \equiv E\left[\exp \int_{-\infty}^{+\infty} \{\log \xi(t)dN(t) + \log \xi_c(t)dN_c(t)\}\right]$$

$$= \exp\left(-\lambda \int_{-\infty}^{+\infty} \left[1-\xi(t)G_2\left(\xi_c|t\right)\right]dt\right) , \qquad (2.2.5)$$

where λ is the rate of the Poisson centre process, and $G_2(\xi_c|t)$ is the probability generating functional of a subsidiary process whose centre is at t , defined by

$$G_2(\xi_c|t) \equiv \int_{N^R} \left[\exp \int_{-\infty}^{+\infty} \log \xi_c(u+t)dN_1(u)\right]dP_2(\mathrm{N}) .$$

If, in particular, we set $\xi(t) = 1 - (1-z)1_A(t)$, and

$\xi_c(u) = 1 - \sum_{i=1}^{k} (1-z_i)1_{A_i}(u)$, we find that $\log M(z, z_1, \ldots, z_k)$ is a function linear in z , so differentiating we have identified

$$\lambda \int_{-\infty}^{+\infty} 1_A(t)E_2\left[z_1^{N_1(A_1-t)} \cdots z_k^{N_1(A_k-t)}\right]dt . \qquad (2.2.6)$$

Now λ can be identified as $\lim_{n\to\infty} N(0, n]/n$, P_1 - a.s. Since (2.2.6) is clearly an absolutely convergent power series in $z_1, \ldots, z_k$, we can equate coefficients to obtain

$$F(r) \equiv \int_0^r P_2\{N_1(A_1-t) = m_1, \ldots, N_1(A_k-t) = m_k\}dt . \qquad (2.2.7)$$

It is not difficult to show that the integrand in $F(r)$ is a continuous function of t . To prove this, let us write

$$P_2(\mathrm{A}, t, v) = P_2\{N_1[(A_i-t)\cap(A_i-v)] = m_i \quad (1 \le i \le k)\} .$$

Then for any $\varepsilon > 0$, there exists by consistency condition (v) of Theorem 1.2.1 a $\delta_1 > 0$ such that for $|v-t| < \delta_1$

$$|P_2(\mathrm{A}, t, t)-P_2(\mathrm{A}, t, v)| \le \varepsilon/2 ,$$

and, similarly, a $\delta_2 > 0$ such that for $|v-t| < \delta_2$,

$$|P_2(A,\ t,\ v)-P_2(A,\ v,\ v)| \le \varepsilon/2\ .$$

Hence, for $|v-t| < \min(\delta_1,\ \delta_2)$,

$$|P_2(A,\ t,\ t)-P_2(A,\ v,\ v)| \le \varepsilon\ ,$$

which we require. Thus, evaluating the limit

$$\left[F(r+\Delta_n)-F(r)\right]/\Delta_n$$

for a sequence of rationals $\Delta_n \downarrow 0$ $(n \to \infty)$, we obtain

$$P_2\{N_1(A_1-r) = m_1,\ \ldots,\ N_1(A_k-r) = m_k\}$$

and thus (2.2.1).

We state this result as

THEOREM 6.2.1. *The cluster structure of a stationary Poisson cluster process with i.i.d. subsidiaries is identifiable with probability one from a complete centre process-subsidiary process record.* □

6.2.3. SPECIAL CLUSTER PROCESSES

In both special processes below (see e.g. Lawrance (1972), Section 4, for definitions), we first wish to identify the distribution of $N_1(R)$. Note that $\{N_1(-k,\ k) \le l\} \downarrow \{N_1(R) \le l\}$ $(k \to \infty)$, $l \in Z_+$, and hence the distribution of $N_1(R)$ is identified.

(a) The Bartlett-Lewis process

By standard results

$$P_2\{N_1(0,\ t] = 0\} = 1 - P_2\{N_1(R) > 0\}F(t)\ ,$$

so F is identified.

(b) The Neyman-Scott process

Again, looking at the same event,

$$P_2\{N_1(A) = 0\} = \sum_{n=0}^{\infty} P_2\{N_1(R) = n\}[1-F(A)]^n$$

for any $A \in \mathcal{B}(R)$, so that F once again is uniquely specified.

6.2.4. CONCLUDING REMARKS

There are several unanswered questions associated with this problem. In particular, we point out that identifiability should be a robust property, and hence we should only need to assume P ergodic. Note that this means that P_1 must be weakly mixing and P_2 ergodic or vice versa. One way to tackle the general problem in the case of P_2 having independent increments would be to expand the p.g.fl. of P (*cf.* Brillinger (1974)). However, we consider that a more aesthetic technique should be available.

Secondly, Milne (1970) has shown that if we are given a realization which is in fact the superposition of a Poisson process and an i.i.d. translation of it (i.e., each point of the Poisson process is independently subjected to a translation with d.f. G say) then, provided G is either symmetric or concentrated on a half-line $[0, \infty)$ or $(-\infty, 0]$, G is still identifiable. We pose the corresponding problem for Poisson centres and right-hand clusters.

Finally, we remark that our formulation gives no idea of the robustness of identifiability of the cluster structure, i.e., given only finite realizations of our processes, can we identify the alternative possible cluster structures and do they differ substantially (in a way that will have to be defined)?

6.3. Problems concerning Chapter 2

P.6.3.1 (REMARKS FOLLOWING THEOREM 2.3.1). Define the service process instead as a probability measure P_2 on R_+^R. Then we can replace ϕ by $\phi' : N \times R_+^R \to \overline{Z}_+^R$ defined as

$$\phi'(N, x)(t) = \int 1_{[v,v+x_v)}(t)dN(v) .$$

With this new model we can weaken Theorem 2.3.1 *(b)*. But one pays a price - it is harder to prove measurability of ϕ', and (more significantly) the model only corresponds with the usual one (i.i.d. service times) if P_1 is a.s. orderly. We think lack of orderliness more important than the weakening, and so have retained the other model.

P.6.3.2 (THEOREM 2.4.1). Weaken the condition in Theorem 2.4.1 *(b)* to $E_2\{x_1\} < \infty$. Possibly the proof there is insufficient, in that the argument used is not via the ergodic theorem, unlike part *(a)*. However, if we disregard the contribution from $t_i < 0$, $E_2\{x_i\} < \infty$ is sufficient.

P.6.3.3 (LEMMA 2.5.4). Virtually unexplored is the question of replacing $\sqrt{n}$ and $\sqrt{\tau}$ in (5.2) and (5.3) of Chapter 2 by

$\mathrm{Var}_{P_2}\left\{\sum_{i=1}^{n} x_i\right\}$ and $\mathrm{Var}_{P_1} N(0, \tau]$. Initial probings reveal difficulties, which may, however, disappear under closer scrutiny.

P.6.3.4 (END OF SECTION 2.5). Suppose we have a general non-stationary arrival process P with first moment measure $M(A) \equiv EN(A)$. We would like the central limit theorem for $\int_0^\tau \phi(s)ds$ asking only (in addition to the components satisfying a central limit theorem and possessing suitable moments) that $M[0, x] = O(x)$ $(x \to \infty)$ (*cf.* Iglehart and Kennedy (1970)). Then, as in Lemma 2.5.2,

$$P_1 \times P_2\left\{\tau^{-\delta}\Phi_1(\tau) \geq \varepsilon\right\} \leq E_1 \times E_2\{\Phi_1(\tau)\}/\varepsilon\tau^\delta$$

$$= \int_\tau^\infty dF(l) \int_0^\tau (u+l-\tau)dM(u)/\varepsilon\tau^\delta$$

$$+ \int_0^\tau dF(l) \int_{\tau-l}^\tau (u+l-\tau)dM(u)/\varepsilon\tau^\delta .$$

Now the first term is $\leq M(0, \tau] \int_\tau^\infty l dF(l)/\varepsilon\tau^\delta$, and so is easily handled. The second term has an upper bound of

$$\int_0^\tau l M(\tau-l,\ \tau] dF(l)/\varepsilon\tau^\delta \ .$$

If P is stationary, then $M(\tau-l, \tau] = ml$, or if P corresponds to a Palm measure then we can use (4.5) of Chapter 2. However, if we have no information such as $M(I) = M(|I|)$, $|I| \to \infty$, i.e., no information on how nicely segments I (intervals) of our process are behaving (*cf.* functional convergence), then it seems difficult to handle this term.

The stationarity of the process of service times can also be removed if again one imposes a condition ensuring that no one service time contributes too much to the central limit theorem (*cf.* uniform asymptotic negligibility). For example, suppose that there exists a distribution function F such that

$$F_j(x) \equiv P_2\{x_j \le x\} \ge F(x) \ ,$$

so that, in a sense, the random variable corresponding to F is dominating the service times. Then, after some algebra,

$$\begin{aligned} E_1 \times E_2\{\Phi_2\} &= E_1 \sum_{j=-\infty}^{0} \left\{\tau - \int_0^\tau F_j(l-t_j)\,dl\right\} \\ &\le E_1 \sum_{j=-\infty}^{0} \left\{\tau - \int_0^\tau F(l-t_j)\,dl\right\} \\ &= \int_0^\infty dF(l) \int_{-l}^{0} \min(l+u,\ \tau)\,dM(u) \end{aligned}$$

and, similarly,

$$E_1 \times E_2\{\Phi_1\} \le \int_0^\tau dM(u) \int_0^\infty \min(l,\ \tau-u)\,dF(l)$$

so that the analysis may be carried out as before.

P.6.3.5 (COROLLARY 2.6.12). Our feeling is that this result is suboptimal. It would be interesting to compare this approach to one using characteristic functions and an inequality due to Esseen (Lemma 2, Chung (1968), p. 208) on the stationary $M/G/\infty$ queue (*cf.* Rao (1966)).

P.6.3.6 (BEGINNING OF SECTION 2.7). The results of Section 2.7 are therefore not really an extension of those of Section 2.6. An examination of other rates of functional convergence theorems reveals that they too are often expressible in terms of the metric of convergence in probability. This seems strange: one should attempt to obtain rates of functional convergence results asking only that the arrival process and service time process converge to the Wiener process at rates in terms of the Prohorov or dual bounded Lipschitz metrics, although for the $GI/G/\infty$ queue the stronger result is available. Dudley's (1972) comment that rates of functional convergence are better expressed via the Prohorov metric therefore seems open to dispute, although his comparison was with Lipschitz functional formulations. The clue to changing the metric from ρ to d in (7.1) of Chapter 2 is given in Billingsley (1968), p. 112, line 15.

P.6.3.7 (LEMMA 2.7.1). Rates of convergence results are prime examples for improvement by fine arguments in unexpected directions. It is by no means clear that the $(\log n)^{3/4}/n^{\frac{1}{4}}$ term is optimal; in particular, the logarithmic factor.

6.4. Problems concerning Chapter 3

P.6.4.1 (THEOREM 3.3.2). The tail sums could hopefully be dealt with as in Barbour (1974), although here we have triangular arrays of convergent series rather than a straight sequence.

P.6.4.2 (LEMMA 3.4.6 (b)). Note that Iglehart (1971a), Theorem 3.3 does not need (directly) the properties of K in his proof. In his notation, $\rho\left(S_n^j, \sigma_j\xi_n\right) \to 0$ a.s. has already been proved, where $\xi_n(t) = \xi(nt)/\phi(n)$ for ξ a standard Brownian motion. His result thus follows from inequality (5.13) of Chapter 1, the relative compactness of ξ_n (Strassen (1964)) and the Arzelà-Ascoli theorem.

P.6.4.3 (AFTER LEMMA 3.4.6). Of course, these remarks are vacuous if a process X_n is defined to obey a FLIL as in Problem 6.4.2 above, i.e., $\rho(X_n, \xi_n) \to 0$ a.s. for ξ_n defined there.

In (4.1.7) of Chapter 3 note that if

$$A_n \equiv \left\{\varepsilon^2\phi(n)^2 \le X_1^2 < \varepsilon^2\phi(n+1)^2\right\} ,$$

$$EX_1^2 = \sum_{n=1}^{\infty} \int_{A_n} X_1^2 dPr \ge 2\varepsilon^2 \sum_{n=1}^{\infty} nPr\{A_n\} = 2\varepsilon^2 \sum_{n=1}^{\infty} Pr\left\{X_n^2 \ge \varepsilon^2\phi(n)^2\right\}$$

provide the intermediate steps. The ideas succeeding (4.1.7) of Chapter 3 are probably well known - for an instance without details see Heyde and Scott (1973).

P.6.4.4. Removing the condition $m < 1$, which is not a logical requirement, but apparently a technical one, in SECTIONS 2.7, 3.3 AND 3.4 seems difficult.

P.6.4.5 (THEOREM 3.4.8). The RHS of (4.2.9) of Chapter 3 is a significant overestimate of the LHS. Using the current arguments, weakening the condition (4.2.4) of Chapter 3 means decreasing in some manageable way the RHS of (4.2.9) of Chapter 3, or estimating tail probabilities in the geometric subsequence argument by something other than Chebyshev's inequality.

P.6.4.6 (THEOREM 3.5.1). A direct proof of the relationship (5.1) of Chapter 3 is not known (to me).

P.6.4.7 (COROLLARY 3.5.5). We seem to require that the service time process obeys a FLIL; the problem is achieving a result parallel to Lemma 3.4.7 for the OLIL.

P.6.4.8 (SECTION 3.5.3). It is by no means clear whether infinite divisibility is a global property. Dr M. Westcott (private communication) has informed me that a student of Matthes is working on the problem: "a doubly stochastic Poisson process is infinitely divisible if and only if its stochastic intensity is infinitely divisible". We pose the question more generally in the light of Kingman's (1964) characterization (not his renewal characterization - see Section 3.5.3): "if a point process is subjected to a random change of time independent of the original process, then the resultant point process is infinitely divisible if and only if the original and time transformation processes are".

P.6.4.9. An open question is the generalization of Theorem 4 of Daley (1972) to a stationary process of subsidiaries, in the style of Lemma 2.5.4 for the $G/G/\infty$ queue. It is clear that this generalization goes through if we assume that the process of subsidiaries is say, strong mixing, with an appropriate condition on the rate of mixing. We would anticipate that we need only assume

$$\sum_{j=0}^{\infty} \left| E\left\{ \left(N_0(R)-\mu\right) \left(N_j(R)-\mu\right) \right\} \right| < \infty$$

where $\mu \equiv E\{N_0(R)\} < \infty$ (compare Lemma 3 of Billingsley (1968), p. 172). However, the problem then seems to involve justifying exchanging the order of some limits and integrals.

P.6.4.10 (LEMMA 4, Westcott (1972)). It would be interesting to know whether there exists a small class of functions ξ with nice properties such that the convergence of the p.g.fls.

$$G_n[\xi] \to G[\xi]$$

guarantees weak convergence for the corresponding point processes (*cf.* the dual bounded Lipschitz metric for weak convergence (Section 1.5)).

P.6.4.11 (WEAK CONTINUITY OF CLUSTER PROCESSES). Kennedy (1972a) and Whitt (1974a) have demonstrated weak continuity for queues. Such an idea should extend to cluster processes, but their proofs rely heavily on the queues beginning at $t = 0$, and the service times being non-negative. However, in the following, we will demonstrate that weak continuity for cluster processes is a valid concept. The proof is via the p.g.fl., and hence we have not striven for the finest possible conditions via this technique, which demands (unnecessarily, we surmise) that the centre process and process of subsidiaries be independent, and the subsidiaries themselves be i.i.d. We use the model (2.2) of Chapter 3, but now write

$P_2 = \bigtimes_{j=-\infty}^{+\infty} P_{2j}$, where P_{2j} is a probability measure on $(N, \sigma(N))$, and set $P_{2j} = P'$ for some generic point process P' for which $P'\{N(R) < \infty\} = 1$.

PROPOSITION 6.4.1. *If* P_1 *and the sequence* $\{P_{1n}\}$ *are stationary centre processes with finite intensities* m , $\{m_n\}$ *and satisfy*

$$P_{1n} \Rightarrow P_1 \quad \text{and} \quad m < \infty , \quad \limsup_{n\to\infty} m_n < \infty \tag{4.1}$$

and if P' *and the sequence* $\{P'_n\}$ *of subsidiaries have first moment measures* $M'(\cdot)$, $\{M'_n(\cdot)\}$ *satisfying*

$$P'_n \Rightarrow P' \quad \text{and} \quad M'(R) < \infty , \quad \limsup_{n\to\infty} M'_n(R) < \infty \tag{4.2}$$

then the sequence of corresponding cluster processes satisfy

$$P_{cn} \Rightarrow P_c .$$

Proof. By Lemma 4 of Westcott (1972) and equation (15) of Vere-Jones (1968), we must prove that the corresponding p.g.fls. satisfy

$$G_{1n}[G'_n(\xi|t)] \to G_1[G'(\xi|t)] \tag{4.3}$$

for $\xi \in V$ (recall that weak convergence and convergence of finite dimensional distributions correspond for a sequence of point processes converging to a point process).

We prove (4.3) in two parts. Firstly, by an inequality in the proof of Theorem 2 of Westcott (1972),

$$|G_{1n}[G'_n(\xi|t)]-G_{1n}[G'(\xi|t)]| \tag{4.4}$$

$$\leq m_n \int_{-\infty}^{+\infty} |G'_n(\xi|t)-G'(\xi|t)|\,dt \tag{4.5}$$

$$\leq m_n\left\{\int_{-\infty}^{+\infty} [1-G'_n(\xi|t)]dt + \int_{-\infty}^{+\infty} [1-G'(\xi|t)]dt\right\} . \tag{4.6}$$

By some easy inequalities, (4.6) has an upper bound of

$$m_n\{M_n'(R)+M'(R)\}|\text{supp.}[\log \xi]|$$

and hence by (4.1), (4.2) and the dominated convergence theorem, (4.5) $\to 0$ $(n \to \infty)$.

For the second half, let $\xi_m \in V$ be a sequence of functions $\downarrow G_2(\xi|t)$ (this is always possible since our functions take values in $[0, 1]$). Then

$$|G_{1n}[G'(\xi|t)]-G_1[G'(\xi|t)]| \tag{4.7}$$

$$\leq (m_n+m)\int_{-\infty}^{+\infty} |\xi_m-G'(\xi|t)|dt + |G_{1n}[\xi_m]-G_1[\xi_m]| \ .$$

Hence, by Lemma 4 of Westcott (1972),

$$\limsup_{n\to\infty} (4.7) \leq \kappa \int_{-\infty}^{+\infty} |\xi_m-G'(\xi|t)|dt \tag{4.8}$$

for some constant κ . But the RHS of (4.8) has a uniform upper bound of

$$2\kappa \int_{-\infty}^{+\infty} [1-G'(\xi|t)]dt$$

which we have already shown to be finite. Hence, by dominated convergence the RHS of (4.8) $\to 0$ $(m \to \infty)$, and thus we have proved (4.3). □

Hopefully, a more direct analysis will yield a stronger result. We remark, though, that continuity in itself is not really what we are looking for: instead (in an obvious phraseology) we require 'rate of continuity'.

6.5. Problems concerning Chapter 4

P.6.5.1 (THEOREM 4.4.1). It was pointed out in Section 4.5 that the condition $P_1 \times P_2(G_\tau) \to 1$ $(\tau \to \infty)$ is not a necessary condition for strong mixing of cluster processes to occur. In fact, we know of no example of a cluster process with strong mixing centre which is not strong mixing. But consider the Neyman-Scott process (Poisson centres) with two points per cluster. Then this process always exists (Westcott (1971)), but if F , the distribution of each point of a cluster from its centre, does not have finite mean, then (4.13) of Chapter 4 fails. We have no suggestions, however, as to the relevant events to consider for the potential counter-example.

P.6.5.2 (THEOREM 4.2.1) (D.J. Daley - solved problem). We point out that a continuous time Markov chain need not be ϕ-mixing. Let $Q(t)$ be the size of a classical immigration-death process; it is well-known that this is a Markov chain. Note that the classical immigration-death process corresponds to the $M/M/\infty$ queue, and $Q(t)$ to the number of servers. Again $Pr\{Q(\tau) = 0 \mid Q(0) = n\}$ can be made arbitrarily small by taking n large enough, so that counter-examples of the style of Section 4.2 follow (compare Bloomfield (1973)).

P.6.5.3. Given a point process which is ϕ-mixing, and $\phi_1 : [0, \infty) \to [0, 1]$ satisfying $\phi_1(\tau) \downarrow 0 \quad (\tau \to \infty)$ and $\phi_1(\tau) > \phi(\tau)$ (all $\tau \in R_+$), construct a process which is ϕ_1-mixing but not ϕ-mixing (P.A.P. Moran).

If we subject the original process to random deletions via an independent process which is $(\phi_1 - \phi)$-mixing, then by Corollary 4.4.2 *(v)*, the new process is ϕ_1-mixing. It is at least intuitively clear that the derived process will not be ϕ-mixing.

P.6.5.4 (THEOREMS 4.3.1, 4.3.3). We suspect that the following is true: if the centre process of a cluster process with i.i.d. clusters is complete, ϕ- or strong mixing, and its p.g.fl. G_1 satisfies (3.1), (3.2) or the p.g.fl. equivalent to (3.8) of Chapter 4 respectively, then so does the p.g.fl. G of the cluster process.

P.6.5.5 (D.J. Daley). ϕ-mixing point processes P may be represented as satisfying, for $C \in \sigma(N(-\infty, t])$ and $D \in \sigma(N(t+\tau, \infty))$,

$$|P(D|C) - P(D)| \le \phi(\tau) .$$

Taking $C = \{N(t-h, t] > 0\}$ and letting $h \downarrow 0$ suggests that ϕ-mixing processes may have nice properties for Palm-Khinchin measures.

P.6.5.6 (D.J. Daley). A more general question is the possibility of a p.g.fl. formulation for the Palm-Khinchin equations, i.e., to relate the p.g.fl. $G(\cdot)$ of a stationary point process to the "conditional p.g.fl. $G_0(\cdot)$, the p.g.fl. for the process given a point at zero".

P.6.5.7 (THEOREM 4.3.1) (P.A.P. Moran). Define and study the properties of the "conditional characteristic functional". This remark arose in relation to the inequality (3.2) of Chapter 4. A relevant reference is Bartlett (1938).

P.6.5.8 (COROLLARY 4.4.2). Let $(X, Y) \equiv \{(X_j, Y_j)\}$ be a double sequence of stationary random variables, where X and Y are independent. If X is strong mixing, Y is strong mixing, is (X, Y) strong mixing? Professor E.J. Hannan (private communication) conjectures that this is not so, but knows of no counter-example. Presuming he is correct, then, comparing (X, Y) to $P_1 \times P_2$ and thus P_c in Theorem 4.4.1, we conclude (tentatively) that Corollary 4.4.2 *(iv)* (and *(v)*) are optimal. If X is ergodic and Y is ergodic then (X, Y) is not necessarily: for a counter-example, see Hannan (1973), p. 163, or Breiman (1968), pp. 100 and 113.

P.6.5.9. Of the same ilk as Problem 8 is the question of superpositions. The techniques of Theorem 4.4.1 yield: independently superposing a completely mixing point process with a strong mixing point process results in a strong mixing point process. Characteristic functional arguments (Theorem 4.3.3) suggest the stronger conjecture: independently superposing a strong mixing $(\text{rate } \alpha_1)$ process with a strong mixing $(\text{rate } \alpha_2)$ process yields a strong mixing $(\text{rate } \alpha_1+\alpha_2)$ process. We do not believe the stronger statement.

P.6.5.10 (SECTION 4.4). Find conditions on η_c which guarantees its measurability. They will probably include the measurability of $v \mapsto N_v(A)$ from $R \to Z_+$ for any given $\{N_v\} \in N^R$ for which $N_v(R) < \infty$ (all $v \in R$)

and any given $A \in \mathcal{B}(R)$.

P.6.5.11. In LEMMA 4.4.3 we conjecture that if P_2 has independent increments, the necessity of (4.6) of Chapter 4 holds in general. The problem is to find a path proving

$$\sum_{t_i(N)\le t-\tau_1} P_2\{F(t_i)\} = \infty \qquad P_1 \text{ - a.s.}$$

$$\Rightarrow P_2\{F(t_i) \text{ i.o., } t_i \le t-\tau_1\} = 1 \qquad P_1 \text{ - a.s.}$$

by exploiting the independence of subsidiaries located at different points.

P.6.5.12. THEOREM 4.4.1 should generalize to multivariate point processes. However, a problem of definition is encountered in the generalization to multidimensional point processes: what is meant, for example, by strong mixing of a point process on the plane?

P.6.5.13 (SECTION 4.1). We have not attempted to indicate when a point process is complete, ϕ- or strong mixing. Many point processes are specified by their inter-epoch times. Therefore a relevant question is: if the inter-epoch times of a point process are complete, ϕ- or strong mixing, does the point process itself have the same property?

6.6. Problems concerning Chapter 5

P.6.6.1 (SECTION 5.1). Robustness of characterizations of queueing systems is very closely allied to the idea of continuity of queues (Kennedy (1972a), Whitt (1974a)). We discuss this idea further in Problem 6.4.11.

Dr D.J. Daley has suggested the following metric of robustness:

$$D_{\theta,z}(P, F) \equiv \int_0^\infty e^{-\theta u}\left|E^0\left(z^{N(0,u]}\right) - E_1\left(z^{N(0,u]}\right)\right| du$$

where $|z| < 1$, E^0 is the expectation with respect to the Palm measure of a given point process P , and E_1 the expectation of a renewal process P_1 with lifetime distribution F . This metric suffers from the defect (possibly) that it is zero when $z = 1$, but could easily be modified. The point is, though, that $D_{\theta,z}(P, F)$

(a) plays down the importance of what happens for large u ;

(b) is small when $N(0, u]$ is large,

i.e., it emphasizes local behaviour. (a) and (b) seem to be desirable properties of any 'robustness' metric.

P.6.6.2 (SECTION 5.1) (D.J. Daley). Characterize stationary $GI/M/1/N$ $(1 \le N < \infty)$ queueing systems with uncorrelated output. Daley (1968) has characterized $GI/M/1/\infty$ and $M/G/1/\infty$ systems with uncorrelated output, while Daley (1974) and Vlach (1971) have characterized $GI/M/1/0$ systems with such an output. King (1971) has answered the $M/G/1/1$ case, and has given an expression for the covariance in the $M/G/1/N$ situation. Preliminary investigations suggest that the techniques of these papers do not assist us here.

P.6.6.3 (SECTION 5.2.1) (D.J. Daley). Characterize stationary $GI/D/1/N$ $(0 \le N \le \infty)$ queueing systems with renewal output. It is known (Daley (1974)) that no restrictions on the inter-arrival distribution are required if $N = 0$, and that the output is renewal for $N = 1$ if the input is Poisson (King (1971)). We conjecture that no other non-degenerate situations (i.e., for which $Pr\{\text{arrival time} > \text{service time}\} < 1$) give rise to renewal output.

P.6.6.4 (SECTION 5.1) (D.J. Daley). For a pure loss $GI/G/1$ queue (i.e., $GI/G/1/0$) find

$$U(x) = E\{\text{number of departures in } (0, x] \mid \text{departure at } 0\} ,$$

and ask when is $U(x) = \lambda x$ for some constant λ (i.e., when do departures look like a Poisson process as far as second order properties go)? We could also investigate when $U(x)$ is a renewal function. This may not be a tractable problem, although Laplace transforms may be of assistance.

P.6.6.5 (SECTION 5.3) (D.J. Daley). Let $\eta = \eta_1 + \eta_2$ be a renewal process which is the superposition of independent alternating renewal processes. Prove that either η_1 or η_2 is Poisson. (If Daley's (1973a) technique is to be used, it may be necessary to assume that η_1 is derived from the jump epochs of a Markov chain, i.e., the lifetime distributions are exponential, with rates α and β ($\alpha \neq \beta$)).

P.6.6.6 (D.J. Daley). Let SM denote a point process derived from the jump epochs of a stationary irreducible semi-Markov process. Characterize those SM processes for which

$$P + SM = R \tag{6.1}$$

(P is Poisson, R renewal). In particular, it is of interest to know whether any two state semi-Markov process with

$p_{ij} \equiv Pr\{\text{transition } i \to j \mid \text{lifetime of type } i \text{ just ended}\}$, $i, j = 1, 2$

and with $p_{ii} \neq 0$ can satisfy (6.1). (This is to be compared with Daley's (1973a) counter-example.)

P.6.6.7 (D.J. Daley). Let MC denote a point process derived from the jump epochs of a stationary irreducible continuous time Markov chain on a countable state space. Daley (1973b) has obtained necessary and sufficient conditions on MC to guarantee

$$P + MC = R \, .$$

What is the distribution function F (or its Laplace Transform) of the renewal lifetimes in R ? Daley (private communication) anticipates that

$$1/\int_0^\infty e^{-\theta t} dF(t) = \left(\lambda + \theta + \int_{(0,\infty]} \left(1-e^{-\theta x}\right)\mu(dx)\right)/\lambda \, ,$$

for some measure μ , because the problem seems to be linked with continuous time regenerative phenomena. Note that this would mean that R is also a doubly stochastic Poisson process (e.g., Kingman, discussion to Bartlett (1963)).

P.6.6.8 (D.J. Daley). Let $\eta(\cdot)$ be a stationary renewal process, and let

$$\chi_n^h = 1 \quad \text{if} \quad \eta\big(nh,\ (n+1)h\big] > 0 \ ,$$
$$= 0 \quad \text{otherwise.}$$

CHARACTERIZATION CONJECTURE. $\{\chi_n^h\}$ *is a discrete time stationary renewal process for every* $h > 0$ *if and only if* $\eta(\cdot)$ *is Poisson. If so, can the conclusion hold if* h *is restricted to an interval.* □

P.6.6.9 (D.J. Daley). Let $\eta(\cdot)$ be a non-arithmetic renewal process with lifetime distribution F having finite mean λ^{-1}, and set

$H(x) = \sum_{n=1}^{\infty} F^{n*}(x)$, where $F^{j*}(x)$ denotes the j-fold convolution of F.

Can skeleton arguments (Kingman (1972), p. 34) be used to show

$$H(x+h) - H(x) \to \lambda h \quad (x \to \infty)$$

for every $h > 0$? (That is, can the Blackwell renewal theorem be deduced from the Erdös-Feller-Pollard theorem?)

P.6.6.10 (D.J. Daley). Let a stationary, orderly point process $\eta(\cdot)$ with finite rate λ be mixing in some appropriate sense. Does this imply a quasi-Blackwell renewal theorem? That is, does the expectation function $U(x) = E\big(\eta(0,\ x] \mid \eta(\{0\}) > 0\big)$ satisfy

$$U(x+h) - U(x) \to \lambda h \quad (x \to \infty) \ ?$$

BIBLIOGRAPHY

Ash, R.B. (1972). *Real Analysis and Probability*. New York and London: Academic Press.

Athreya, K.B. and Karlin, S. (1967). Limit theorems for the split times of branching processes. *J. Math. Mech.* **17**, 257-278.

Barbour, A.D. (1974). Tail sums of convergent series of independent random variables. *Proc. Camb. Phil. Soc.* **75**, 361-364.

Bartlett, M.S. (1938). The characteristic function of a conditional statistic. *J. Lond. Math. Soc.* **13**, 62-67.

Bartlett, M.S. (1954). Processus stochastiques ponctuels. *Ann. Inst. H. Poincaré* **14**, 34-60.

Bartlett, M.S. (1963). The spectral analysis of point processes. *J.R. Statist. Soc. B* **25**, 264-296.

Bendrath, R. (1974). Verallgemeinerung eines Satzes von R.K. Milne. *Math. Nachr.* **59**, 221-228.

Beneš, V.E. (1957). A sufficient set of statistics for a simple telephone exchange model. *Bell System Tech. J.* **36**, 939-964.

Beutler, F.J. and Leneman, O.A.Z. (1966). The theory of stationary point processes. *Acta Math.* **116**, 159-197.

Bhabha, H.J. (1950). On the stochastic theory of continuous parametric systems and its application to electron cascades. *Proc. Roy. Soc. Ser. A* **202**, 301-322.

Billingsley, P. (1956). The invariance principle for dependent random variables. *Trans. Amer. Math. Soc.* **83**, 250-268.

Billingsley, P. (1962). Limit theorems for randomly selected partial sums. *Ann. Math. Statist.* **33**, 85-92.

Billingsley, P. (1968). *Convergence of Probability Measures*. New York: Wiley.

Bloomfield, P. (1973). Note on a central limit theorem of I.A. Ibragimov, with an application to the sojourn times of the $M/G/\infty$ queue. *Technical Report* 41, Series 2. Princeton University: Dept. of Statistics.

Boes, D.C. (1969). Note on the output of a queueing system. *J. Appl. Prob.* 6, 459-461.

Breiman, L. (1968). *Probability.* Reading, Mass.: Addison-Wesley.

Brillinger, D.R. (1974). Cross-spectral analysis of processes with stationary increments including the stationary $G/G/\infty$ queue. *Ann. Probability* 2, 815-827.

Brown, M. (1970). An $M/G/\infty$ estimation problem. *Ann. Math. Statist.* 41, 651-654.

Brown, M. and Ross, S.M. (1969). Some results for infinite server Poisson queues. *J. Appl. Prob.* 6, 604-611.

Burke, P.J. (1956). The output of a queueing system. *Operat. Res.* 4, 699-704.

Chover, J. (1967). On Strassen's version of the log log law. *Z. Wahrscheinlichkeitstheorie verw. Geb.* 8, 83-90.

Chung, K.L. (1968). *A Course in Probability Theory.* 1st ed. New York: Harcourt Brace and World Inc.

Çinlar, E. (1972). Superposition of point processes. In *Stochastic Point Processes*, Ed. P.A.W. Lewis. New York: Wiley, 549-606.

Cox, D.R. and Lewis, P.A.W. (1966). *The Statistical Analysis of Series of Events.* London: Methuen and Barnes, New York: Noble.

Daley, D.J. (1968). The correlation structure of the output of some single server queueing systems. *Ann. Math. Statist.* 39, 1007-1019.

Daley, D.J. (1971). Weakly stationary point processes and random measures. *J. Roy. Statist. Soc. B* 33, 406-428.

Daley, D.J. (1972). Asymptotic properties of stationary point processes with generalized clusters. *Z. Wahrscheinlichkeitstheorie verw. Geb.* 21, 65-76.

Daley, D.J. (1973a). Poisson and alternating renewal processes with superposition a renewal process. *Math. Nachr.* **57**, 359-369.

Daley, D.J. (1973b). Markovian processes whose jump epochs constitute a renewal process. *Quart. J. Math. Oxford* (2), **24**, 97-105.

Daley, D.J. (1974). Characterizing pure loss $GI/G/1$ queues with renewal output. *Proc. Camb. Phil. Soc.* **75**, 103-107.

Daley, D.J. (1975). Queueing output processes. *Adv. Appl. Prob.* (to appear).

Daley, D.J. and Milne, R.K. (1973). The theory of point processes: a bibliography. *Internat. Statist. Rev.* **41**, 183-201.

Daley, D.J. and Shanbhag, D.N. (1975). Independent inter-departure times in $M/G/1/N$ queues. *J. Roy. Statist. Soc. B* (to appear).

Daley, D.J. and Vere-Jones, D. (1972). A summary of the theory of point processes. In *Stochastic Point Processes*, Ed. P.A.W. Lewis. New York: Wiley, 299-383.

Dharmadhikari, S.W., Fabian, V., and Jogdeo, K. (1968). Bounds on the moments of martingales. *Ann. Math. Statist.* **39**, 1719-1723.

Disney, R.L., Farrell, R.L. and Morais, P.R. de. (1973). A characterization of $M/G/1/N$ queues with renewal departure processes. *Management Sci.* **19**, 1222-1228.

Dudley, R.M. (1966). Convergence of Baire measures. *Studia Math.* **27**, 251-268.

Dudley, R.M. (1968). Distances of probability measures and random variables. *Ann. Math. Statist.* **39**, 1563-1572.

Dudley, R.M. (1972). Speeds of metric probability convergence. *Z. Wahrscheinlichkeitstheorie verw. Geb.* **22**, 323-332.

Feller, W. (1950). *An Introduction to Probability Theory and Its Applications.* Vol. 1, 1st ed. New York: Wiley.

Feller, W. (1966). *An Introduction to Probability Theory and its Applications.* Vol. 2. New York: Wiley.

Feller, W. (1968). *An Introduction to Probability Theory and Its Applications.* Vol. 1, 3rd ed. New York: Wiley.

Finch, P.D. (1959). The output process of the queueing system $M/G/1$. *J. Roy. Statist. Soc. Ser. B* 21, 375-380.

Fisher, L. (1972). A survey of the mathematical theory of multidimensional point processes. In *Stochastic Point Processes*, Ed. P.A.W. Lewis. New York: Wiley, 468-513.

Freedman, D. (1967). Some invariance principles for functionals of a Markov chain. *Ann. Math. Statist.* 38, 1-7.

Goldman, J.R. (1967). Infinitely divisible point processes in R^n . *J. Math. Anal. Appl.* 17, 133-146.

Grandell, J. (1971). On stochastic processes generated by a stochastic intensity function. *Skand. Aktuar. Tidskr.* 54, 204-240.

Greenwood, M. and Yule, G.U. (1920). An inquiry into the nature of frequency distributions representative of multiple happenings, with particular reference to the occurrence of multiple attacks of disease or of repeated accidents. *J.R. Statist. Soc.* 83, 255-279.

Gunning, R.C. and Rossi, H. (1965). *Analytic Functions of Several Complex Variables.* Englewood Cliffs, N.J.: Prentice Hall.

Halmos, P.R. (1950). *Measure Theory.* Princeton, N.J.: D. Van Nostrand Co. Inc.

Hannan, E.J. (1973). Central limit theorems for time series regression. *Z. Wahrscheinlichkeitstheorie verw. Geb.* 26, 157-170.

Harris, T.E. (1963). *The Theory of Branching Processes.* Berlin: Springer.

Hawkes, A.G. and Oakes, D. (1974). A cluster process representation of a self-exciting process. *J. Appl. Prob.* 11, 493-503.

Heyde, C.C. (1969). On extended rate of convergence results for the invariance principle. *Ann. Math. Statist.* 40, 2178-2179.

Heyde, C.C. (1974). On the central limit theorem for stationary processes. *Z. Wahrscheinlichkeitstheorie verw. Geb.* 30, 315-320.

Heyde, C.C. and Scott, D.J. (1973). Invariance principles for the law of the iterated logarithm for martingales and processes with stationary increments. *Ann. Probability* 1, 428-436.

Ibragimov, I.A. (1962). Some limit theorems for stationary processes. *Theor. Probability Appl.* 7, 349-382.

Iglehart, D.L. (1971a). Multiple channel queues in heavy traffic IV. Law of the iterated logarithm. *Z. Wahrscheinlichkeitstheorie verw Geb.* 17, 168-180.

Iglehart, D.L. (1971b). Functional limit theorems for the queue *GI/G/*1 in light traffic. *Adv. Appl. Prob.* 3, 269-281.

Iglehart, D.L. (1974). Weak convergence in applied probability. *Stoch. Procs. and their Applns.* 2, 211-241.

Iglehart, D.L. and Kennedy, D.P. (1970). Weak convergence of the average of flag processes. *J. Appl. Prob.* 7, 747-753.

Iglehart, D.L. and Whitt, W. (1971). The equivalence of functional central limit theorems for counting processes and associated partial sums. *Ann. Math. Statist.* 42, 1372-1378.

Jagers, P. (1973). On Palm probabilities. *Z. Wahrscheinlichkeitstheorie verw Geb.* 26, 17-32.

Jagers, P. (1974). Aspects of random measures and point processes. In *Advances in Probability* 3, Ed. P. Ney. New York: Marcel Dekker, Inc. (to appear).

Jagers, P. and Lindvall, T. (1974). Thinning and rare events in point processes. *Z. Wahrscheinlichkeitstheorie verw Geb.* 28, 89-98. (Erratum 29, 272).

Kallenberg, O. (1973). Characterization and convergence of random measures and point processes. *Z. Wahrscheinlichkeitstheorie verw Geb.* 27, 9-21.

Kallenberg, O. (1974). A limit theorem for thinning of point processes. *Institute of Statistics Mimeo Series No.* 98, University of North Carolina.

Kaplan, E.L. (1955). Transformations of stationary random sequences. *Math. Scand.* **3**, 127-149.

Kaplan, N. (1974). Limit theorems for a $GI/G/\infty$ queue. *Preprint*, University of California: Berkeley.

Keilson, J. (1966). The ergodic queue length distribution for queueing systems with finite capacity. *J. Roy. Statist. Soc. Ser. B* **28**, 190-201.

Kemp, C.D. and Kemp, A.W. (1965). Some properties of the 'Hermite' distribution. *Biometrika* **52**, 381-394.

Kennedy, D.P. (1972a). The continuity of the single server queue. *J. Appl. Prob.* **9**, 370-381.

Kennedy, D.P. (1972b). Rates of convergence for queues in heavy traffic. I. *Adv. Appl. Prob.* **4**, 357-381.

Kerstan, J. and Matthes, K. (1964). Stationäre zufällige Punktfolgen, II. *Jber. Deutsch. Math.-Verein.* **66**, 106-118.

Kerstan, J., Matthes, K. and Mecke, J. (1972). *Unbegrenzt teilbare Punktprozesse*. Berlin: Akademie Verlag.

Khinchin, A.Ya. (1955). *Mathematical Methods in the Theory of Queueing.* Trudy Mat. Inst. Steklov **49**. (English translation (1960). London: Griffin.)

King, R.A. (1971). The covariance structure of the departure process from $M/G/1$ queues with finite waiting lines. *J. Roy. Statist. Soc. Ser. B* **33**, 401-406.

Kingman, J.F.C. (1964). On doubly stochastic Poisson processes. *Proc. Camb. Phil. Soc.* **60**, 923-930.

Kingman, J.F.C. (1972). *Regenerative Phenomena.* New York: Wiley.

Krickeberg, K. (1974). Moments of point processes. In *Stochastic Geometry*, Eds. E.F. Harding and D.G. Kendall. New York: Wiley, 89-113.

Kurtz, T.G. (1974). Point processes and completely monotone set functions. *Z. Wahrscheinlichkeitstheorie verw. Geb.* **30**, 57-67.

Laslett, G.M. (1975). Characterising the finite capacity *GI/M/*1 queue with renewal output. *Management Sci.* (to appear).

Lawrance, A.J. (1972). Stationary series of univariate events. In *Stochastic Point Processes,* Ed. P.A.W. Lewis. New York: Wiley, 199-256.

Leadbetter, M.R. (1972a). On basic results of point process theory. *Proc. Sixth Berkeley Symp.* **3**, 449-461.

Leadbetter, M.R. (1972b). Point processes generated by level crossings. In *Stochastic Point Processes,* Ed. P.A.W. Lewis. New York: Wiley, 436-467.

Lee, P.M. (1967). Infinitely divisible stochastic processes. *Z. Wahrscheinlichkeitstheorie verw. Geb.* **7**, 147-160.

Lee, P.M. (1968). Some examples of infinitely divisible point processes. *Studia Sci. Math. Hung.* **3**, 219-224.

Lewis, P.A.W. (1964a). A branching Poisson process model for the analysis of computer failure patterns. *J.R. Statist. Soc. B* **26**, 398-456.

Lewis, P.A.W. (1964b). The implications of a failure model for the use and maintenance of computers. *J. Appl. Prob.* **1**, 347-368.

Lewis, P.A.W. (1969). Asymptotic properties and equilibrium conditions for branching Poisson processes. *J. Appl. Prob.* **6**, 355-371.

Lewis, P.A.W. (1972). *Stochastic Point Processes*. Ed. P.A.W. Lewis. New York: Wiley.

Lewis, T. (1961). The intervals between regular events displaced in time by independent random deviations of large dispersion. *J.R. Statist. Soc. B* **23**, 476-483.

Lindvall, T. (1973). Weak convergence of probability measures and random functions in the function space $D[0,\infty)$. *J. Appl. Prob.* **10**, 109-121.

Loève, M. (1960). *Probability Theory*, 2nd ed. Princeton, N.J.: D. Van Nostrand Co., Inc.

Lotka, A.J. (1939). A contribution to the theory of self-renewing aggregates, with especial reference to industrial replacement. *Ann. Math. Statist.* **10**, 1-25.

Loulou, R. (1973). Multi-channel queues in heavy traffic. *J. Appl. Prob.* **10**, 769-777.

Mackey, G.W. (1974). Ergodic theory and its significance for statistical mechanics and probability theory. *Advances in Mathematics* **12**, 178-268.

Matthes, K. (1963a). Stationäre zufällige Punktfolgen, I. *Jber. Deutsch. Math.-Verein.* **66**, 66-79.

Matthes, K. (1963b). Unbeschränkt teilbare Verteilungsgesetze stationärer zufalliger Punktfolgen. *Wiss. Z. Hochsch. Electro. Ilmenau* **9**, 235-238.

Mecke, J. (1967). Stationäre zufällige Masse auf localkompakten Abelschen Gruppen. *Z. Wahrscheinlichkeitstheorie verw. Geb.* **9**, 36-58.

Milne, R.K. (1970). Identifiability for random translations of Poisson processes. *Z. Wahrscheinlichkeitstheorie verw. Geb.* **15**, 195-201.

Milne, R.K. (1971). *Stochastic Analysis of Multivariate Point Processes.* Canberra: Australian National Univ. PhD thesis.

Moyal, J.E. (1962). The general theory of stochastic population processes. *Acta Math.* **108**, 1-31.

Narasimham, T. (1968). A note on the asymptotic distribution of the traffic-time average in a $GI/G/\infty$ with bulk arrivals. *J. Appl. Prob.* **5**, 476-480.

Nawrotzki, K. (1962). Eine Grenzwertsatz für homogene zufällige Punktfolgen. *Math. Nachr.* **24**, 201-217.

Nelsen, R.B. and Williams, T. (1968). Randomly delayed appointment streams. *Nature* **219**, 573-574.

Nelsen, R.B. and Williams, T. (1970). Random displacements of regularly spaced events. *J. Appl. Prob.* **7**, 183-195.

Neveu, J. (1968). Sur la structure des processus ponctuels stationnaires. *C.R. Acad. Sci. Paris Sér. A* **267**, A561-564.

Neyman, J. (1939). On a new class of contagious distributions applicable to entomology and bacteriology. *Ann. Math. Statist.* **10**, 35-57.

Neyman, J. and Scott, E.L. (1952). A theory of the spatial distribution of galaxies. *Astrophys. J.* **116**, 144-163.

Neyman, J. and Scott, E.L. (1958). Statistical approach to the problems of cosmology. *J. Roy. Statist. Soc. Ser. B* **20**, 1-43.

Neyman, J. and Scott, E.L. (1964). A stochastic model of epidemics. In *Stochastic Models in Medicine and Biology*, Ed. J. Gurland. Madison: The University of Wisconsin Press, 45-83.

Oodaira, H. and Yoshihara, K. (1971a). The law of the iterated logarithm for stationary processes satisfying mixing conditions. *Kōdai Math. Sem. Rep.* **23**, 311-334.

Oodaira, H. and Yoshihara, K. (1971b). Note on the law of the iterated logarithm for stationary processes satisfying mixing conditions. *Kōdai Math. Sem. Rep.* **23**, 335-342.

Ooodaira, H. and Yoshihara, K. (1972). Functional central limit theorems for strictly stationary processes satisfying mixing conditions. *Kōdai Math. Sem. Rep.* **24**, 259-269.

Palm, C. (1943). Intensitätsschwankungen im Fernsprechverkehr. *Ericsson Techniks* **44**, 1-189.

Papangelou, F. (1974). On the Palm probabilities of processes of points and processes of lines. In *Stochastic Geometry*, Eds. E.F. Harding and D.G. Kendall. New York: Wiley, 114-147.

Philipp, W. (1969). The central limit theorem for mixing sequences of random variables. *Z. Wahrscheinlichkeitstheorie verw. Geb.* **12**, 155-171.

Polya, G. (1931). Sur quelques points de la théorie des probabilités. *Ann. de l'Inst. Henri Poincaré* **1**, 117-162.

Prékopa, A. (1957a). On the compound Poisson distribution. *Acta Sci. Math. Szeged.* **18**, 23-28.

Prékopa, A. (1957b). On Poisson and compound Poisson stochastic functions. *Studia Math.* **16**, 142-155.

Prohorov, Yu.V. (1956). Convergence of random measures and limit theorems in probability theory. *Theor. Probability Appl.* **1**, 157-214.

Prohorov, Yu.V. and Rozanov, Yu.A. (1969). *Probability Theory.* Berlin: Springer-Verlag.

Ramakrishnan, A. (1950). Stochastic processes relating to particles distributed in a continuous infinity of states. *Proc. Camb. Phil. Soc.* **46**, 595-602.

Rao, J.S. (1966). An application of stationary point processes to queueing theory and textile research. *J. Appl. Prob.* **3**, 231-246.

Rényi, A. (1967). Remarks on the Poisson process. *Stud. Sci. Math. Hung.* **2**, 119-123.

Rosenblatt, M. (1956). A central limit theorem and strong mixing condition. *Proc. Nat. Acad. Sci. USA* **42**, 43-47.

Rosenblatt, M. (1962). *Random Processes.* New York: Oxford University Press.

Rosencrantz, W.A. (1968). On rates of convergence for the invariance principle. *Trans. Amer. Math. Soc.* **129**, 542-552.

Ross, S.M. (1970). Identifiability in *GI/G/k* queues. *J. Appl. Prob.* **7**, 776-780.

Rudemo, M. (1973). Point processes generated by transitions of Markov chains. *Adv. Appl. Prob.* **5**, 262-286.

Ryll-Nardzewski, C. (1961). Remarks on processes of calls. *Proc.* 4th *Berkeley Symp.* **2**, 455-465.

Saunders, D.I. (1975). Convergence of random measures. *Preprint*, Flinders University: School of Mathematical Sciences.

Skorokhod, A.V. (1956). Limit theorems for stochastic processes. *Theor. Probability Appl.* **1**, 261-290.

Slivnyak, I.M. (1962). Some properties of stationary flows of homogeneous random events. *Teor. Veroyatnost. i Primenen* **7**, 347-352. (English translation in *Theory Prob. Appl.* **7**, 336-341). Addendum **9**, 190 (English translation in *Theory Prob. Appl.* **9**, 168).

Slivnyak, I.M. (1966). Stationary streams of homogeneous random events. *Vest. Harkov. Gos. Univ., Ser. Mech.-Math.* **32**, 73-116. (English translation: Statistics Department, Australian National University.)

Smith, W.L. (1958). Renewal theory and its ramifications. *J. Roy. Statist. Soc. Ser. B* **20**, 243-302.

Störmer, H. (1969). Zur Überlagerung von Erneuerungsprozessen. *Z. Wahrscheinlichkeitstheorie verw. Geb.* **13**, 9-24.

Straf, M.L. (1972). Weak convergence of stochastic processes with several parameters. *Proc.* 6th *Berkeley Symp.* **2**, 187-221.

Strassen, V. (1964). An invariance principle for the law of the iterated logarithm. *Z. Wahrscheinlichkeitstheorie verw. Geb.* **3**, 211-226.

Thedeen, T. (1969). On road traffic with free overtaking. *J. Appl. Prob.* **6**, 524-549.

Thompson, H.R. (1955). Spatial point processes, with applications to ecology. *Biometrika* **42**, 102-115.

Tomko, J. (1972). The rate of convergence in central limit theorems for service systems with finite queue capacity. *J. Appl. Prob.* **9**, 87-102.

Vere-Jones, D. (1968). Some applications of probability generating functionals to the study of input/output streams. *J.R. Statist. Soc. B* **30**, 321-333.

Vere-Jones, D. (1970). Stochastic models for earthquake occurrence. *J.R. Statist. Soc. B* **32**, 1-62.

Vere-Jones, D. (1971). Appendix to Daley (1971).

Vere-Jones, D. (1973). *Lectures on Point Processes*. Unpublished notes, University of California, Berkeley: Department of Statistics.

Vere-Jones, D. (1974). An elementary approach to the spectral theory of stationary random measures. In *Stochastic Geometry*, Eds. E.F. Harding and D.G. Kendall. New York: Wiley, 307-321.

Verwaat, W. (1972). Functional central limit theorems for processes with positive drift and their inverses. *Z. Wahrscheinlichkeitstheorie verw. Geb.* **23**, 245-253.

Vlach, T.L. (1971). Further results concerning the simple queueing loss system. *Z. Operat. Res.* **15**, 55-57.

Von Bahr, B. and Esseen, C.G. (1965). Inequalities for the rth absolute moment of a sum of random variables, $1 \le r \le 2$. *Ann. Math. Statist.* **36**, 299-303.

Westcott, M. (1970). *Some applications of the probability generating functional to point processes.* Canberra: Australian National Univ. PhD Thesis.

Westcott, M. (1971). On existence and mixing results for cluster point processes. *J.R. Statist. Soc. B* **33**, 290-300.

Westcott, M. (1972). The probability generating functional. *J. Austral. Math. Soc.* **14**, 448-466.

Westcott, M. (1973). Results in the asymptotic and equilibrium theory of Poisson cluster processes. *J. Appl. Prob.* **10**, 807-823.

Whitt, W. (1970). Weak convergence of probability measures on the function space $C[0, \infty)$. *Ann. Math. Statist.* **41**, 939-944.

Whitt, W. (1971). Weak convergence of probability measures on the function space $D[0, \infty)$. *Technical report*, Yale University.

Whitt, W. (1973). On the quality of Poisson approximations. *Z. Wahrscheinlichkeitstheorie verw. Geb.* **28**, 23-36.

Whitt, W. (1974a). The continuity of queues. *Adv. Appl. Prob.* **6**, 175-183.

Whitt, W. (1974b). Preservation of rates of convergence under mappings. *Z. Wahrscheinlichkeitstheorie verw. Geb.* **29**, 39-44.

Whitt, W. (1975). Representation and convergence of point processes on the line. *Ann. Probability* (to appear).

Wichura, M.J. (1973). Some Strassen-type laws of the iterated logarithm for multiparameter stochastic processes with independent increments. *Ann. Probability* 1, 272-296.

Wold, H. (1949). Sur les processus stationairres ponctuels. *Le Calcul des Probabilités et ses Applications*. Publications du C.N.R.S. 13, 75-86.