SOME APPLICATIONS OF THE PROBABILITY GENERATING FUNCTIONAL TO POINT PROCESSES

by

M. Westcott

A thesis submitted for the degree of

Doctor of Philosophy

in the Australian National University.

February 1970

This thesis contains the original work of the author except where specific reference is made in the text. Chapter 5 is a modified version of a paper submitted for publication.

Mark Westcott

Mark Westcott

TABLE OF CONTENTS

Acknowledgements v

Preface vii

Notation xiii

1. POINT PROCESSES

1.1 Introduction 1
1.2 The Definition of a Point Process 2
1.3 Some Properties of Point Processes 4
1.4 A Theorem on Stationary Streams 8
1.5 Some Examples of Point Processes 11

2. THE PROBABILITY GENERATING FUNCTIONAL

2.1 Introduction 15
2.2 Definitions and Properties of the p.g.fl 15
2.3 The Connection with Moment Measures 24
2.4 Characterisations of Ergodicity and Mixing 29

3. CLUSTER PROCESSES

3.1 Introduction 36
3.2 Definitions, Properties and Examples of Cluster Processes 37
3.3 Existence Criteria for Cluster Processes 42
3.4 Mixing in Cluster Processes 50
3.5 Asymptotic Properties of Cluster Processes 55
3.6 Infinite-Server Queueing Systems 63

4. LIMIT THEOREMS FOR POINT PROCESSES

4.1 Introduction 67
4.2 Ergodic Theorems for Point Processes 69
4.3 Infinitely Divisible Point Processes 72
4.4 Deletion and Translation Theorems for Point Processes 78

5. IDENTIFIABILITY IN LINEAR STOCHASTIC POISSON PROCESSES

5.1 Introduction 84
5.2 Preliminary Results 84
5.3 Identifiability of ψ 86
5.4 Identifiability of f 91

References 98

ACKNOWLEDGEMENTS

The majority of this work was carried out at the Australian National University. The stimulating atmosphere of the two Departments of Statistics there contributed in no small measure to the development of this thesis, and I am most grateful to all members of the departments for assisting in this most important way.

My principal debt, of course, is to my supervisors, Professor P.A.P. Moran and Dr D. Vere-Jones, for their constant help and guidance. I thank Professor Moran especially for his encouragement during my early tentative efforts at research and his thoughtful questioning of all my work, and Dr Vere-Jones for his introduction to the probability generating functional and his lively interest in its development. I am also most grateful to Mr R.K. Milne for numerous enjoyable and stimulating conversations, for his patience at my interruptions and his careful reading of a draft of this thesis. Professor E.J. Hannan made several suggestions on the material of Chapter 5, and I thank him for his advice. The assistance of Dr P. Mandl in the translation of several German papers is deeply appreciated, and was of the greatest assistance.

I also spent a term at the Flinders University of South Australia working under Dr M. Jiřina. His help is particularly evident in the sections on ergodicity and mixing. I wish to express my thanks to him, as well as to Professor J.N. Darroch and the other members of the Mathematics Department at Flinders, for a most profitable and enjoyable stay.

I would like to thank Dr P.A.W. Lewis for making available, before publication, a copy of his paper [51], and Dr D.J. Daley for sending an advance copy of the abstract of his address [15] to the I.S.I. Conference. Both these works were most useful.

Finally, I am deeply grateful to Mrs B. Cranston for her splendid typing of the final draft of this thesis.

PREFACE

This thesis is concerned with a study of some aspects of the probability generating functional (p.g.fl) and its application to a variety of problems in the theory of point processes. Several of these problems are new, in particular the questions of existence and mixing for cluster processes. Others have been studied before, but the present unified approach may have some didactic advantages as well as giving elegant proofs of individual results.

The p.g.fl is one of a number of functionals which have been introduced into probability theory from time to time. The senior member of the family is the characteristic functional (c.fl), due initially to Kolmogorov [42] in 1935 for set functions on Banach space and rediscovered simultaneously in 1947 by Bochner [11] and Le Cam [44]; another form of definition is in Shiryaev [77]. A useful survey of work up to 1960 is contained in Prohorov [70] (see also Moran [60] §6.16 and Grenander [24] §1.4 and ch. 6). Recently the c.fl has become part of general investigations into Fourier transforms on abstract topological spaces, especially whether analogues of the theorems of Bochner and Lévy for characteristic functions are valid e.g. Badrikian [2] and Von Wandenfels [90].

The idea of a moment generating functional to describe the age-distribution of a population at a given time is due to D.G. Kendall [34] in 1949. He also recognised its connection with a set of cumulant functions or product densities. Further development of this approach may be found in Bartlett and Kendall [7] and Bartlett's book [6], where also the first mention of the p.g.fl occurs. However this is always associated with a population whose size is finite with probability one, a situation we wish to generalise in this thesis. For the almost surely finite case, a theory of the p.g.fl is contained in Moyal [61] and Harris [28] ch. 3.

The remaining member of the functional class is the Laplace functional, introduced by Jiřina [32] in 1962 for a study of finite non-negative random measures and subsequently used in [33] and Mecke [57]. One of his achievements is to derive a characterisation of weak convergence of these measures in terms of Laplace functionals, a concept which seems open to generalisation to point processes.

An obvious question now is, why choose the p.g.fl for study in preference to the other available functionals? The answer is twofold. Firstly, we shall see in Chapter 1 that there are several classes of moment measures which we can associate with a point process, each in turn associated with a particular functional. It turns out that the class of measures least affected by diagonal concentrations is the one associated with the p.g.fl. This aspect is developed in Chapter 2. Secondly, a principal attraction of the generating functional method is its potential for facilitating proofs of general limit theorems for point processes. Many such theorems assert convergence to a Poisson process, which is most naturally associated with the p.g.fl. This is further explained in the introduction to Chapter 4.

More difficult to answer is whether the p.g.fl merits further intensive study. The applications given in this thesis are certainly encouraging but in the area where most was expected from it, namely limit theorems, there are several unsolved problems barring progress towards a general theory. However it is not at all clear whether these difficulties are symptomatic of an overall intractability in the method or merely technical matters brought about by the infancy of the subject, and this is certainly worth consideration in the future. Also, even if the hopes held for the p.g.fl in the theory of point processes are overoptimistic Dr D. Vere-Jones has stated in his reply to the discussion of [88] that '... it seemed to me that its [the p.g.fl's] usefulness had been understated in the past, and that there was a need to redress the balance', a statement which the author supports and which is the motivation for this thesis.

We conclude with a brief summary of the contents of the thesis.

Chapter 1 is largely expository and draws together various facts about point processes needed in the future. The literature in this field is widely scattered, so that it is easy to miss useful contributions, but it is hoped that nothing of importance relevant to this thesis has been overlooked. In this chapter we give a formal definition of a point process followed by some relevant properties such as stationarity, ergodicity and mixing, convergence in distribution and infinite divisibility. The various classes of moment measures are also defined. A result on the interpretation of an arbitrary stationary point process as a marked point sequence in the sense of Matthes [55] is proved, and used to show the existence of a 'batch-size' distribution for multiple occurrences of events.

Chapter 2 introduces the p.g.fl of an arbitrary point process, following Vere-Jones [87], and we prove a number of properties analogous to those of the more common probability generating function. These include continuity, a theorem specifying which functionals are p.g.fls, and a characterisation of convergence in distribution by convergence of p.g.fls. Some examples of p.g.fls are given, in particular that of the doubly stochastic Poisson process. The connection with the factorial moment measures is then explained and put on a rigorous basis with the proof of a finite Taylor-type expansion for the p.g.fl and its logarithm. We give characterisations of ergodic and mixing point processes in terms of relations involving their p.g.fls, related to ideas of Leonov [47], and use them to establish mixing for several classes of point processes.

The important class of cluster processes is studied in Chapter 3. We first state the fundamental p.g.fl equation for a cluster process, due to Moyal, and give some examples of models introduced for various purposes. This includes results showing that several apparently distinct classes of point processes can be identical. Then comes a section on existence criteria for cluster processes, in the sense of having finite numbers of events in bounded sets with probability one. The general necessary and sufficient conditions obtained are applied to several of the examples mentioned above. The results of Chapter 2 are now invoked to prove a wide-ranging theorem on mixing in cluster processes, with applications to the class of infinitely divisible point processes.

Asymptotic results for the cumulants and probabilities of a generalised Bartlett-Lewis model ([4], [49]) follow, and a neat proof of the form of the equilibrium distribution is given in a particular case. Finally, we indicate applications of the p.g.fl to various aspects of infinite-server queueing systems.

Limit theorems for point processes are the subject of Chapter 4. The introduction explains some of the problems encountered in trying to set up a general limit theory using the p.g.fl. We establish the Weak and Strong Laws of Large Numbers for point processes, under suitable conditions, using a recent result of Daley [15]. The theory of Chapter 2 is then used to derive a canonical form for the p.g.fl of an infinitely divisible point process, based on previous work for the generating functions, and this in turn provides some characterisations of such processes including an equivalent of the 'accompanying laws' theorem in terms of p.g.fls. Then we establish Poisson convergence for a variety of operations on point processes, one of which illustrates the application of the Taylor expansion of Chapter 2. These operations are superposition, random deletion and random translation.

In Chapter 5 we use the p.g.fl to study the identifiability of the two parametric functions in a doubly stochastic Poisson process whose mean process is linear. This involves deciding when these functions are uniquely determined by the process, i.e. by its p.g.fl, although we never attempt to estimate them. The emphasis throughout is on the point process although in the author's paper [91], on which this chapter is based, additional results for the associated linear process are given in a more general setting. We prove that both functions are identifiable, under certain conditions, from quantities which may in practice provide reasonable estimators.

It is intended to submit the results of this thesis for publication in the near future.

NOTATION

We invariably use $\mathcal{P}$ to denote a probability measure, the space on which it is defined being clear from the context. P will always be the univariate or multivariate probability generating function (p.g.f.) of a discrete probability measure.

$E\{\cdot\}$ indicates expectation with respect to the appropriate $\mathcal{P}$. For the variance of a random variable we write $D(\cdot)$.

The letters a.s., p., d., above an arrow, e.g. $\xrightarrow{a.s}$, indicate convergence of a sequence of random variables almost surely, in probability or in distribution respectively (see especially 1.3). Above an equality sign, e.g. $\overset{d}{=}$, they indicate equality in that mode. In the text a.s. means almost surely, that is with probability one.

Theorems, lemmas and equations are numbered independently and consecutively in each chapter. Thus the ordered pair x.y is expression y in Chapter x.

$\chi_A(\cdot)$ is the indicator function of a set A, namely $\chi_A(t) = 1$ if $t \in A$, 0 otherwise.

L.H.S., R.H.S. mean the left-hand side and right-hand side of an equation.

All other notation is explained as it is introduced.

1. POINT PROCESSES

1.1. Introduction

A general theory of point processes is of comparatively recent origin, although particular cases have been extensively analysed. There are two areas of greatest interest, corresponding to processes with finite or infinite numbers of points. Population processes are the principal members of the first class, and here the theory is due initially to Bartlett [3], being substantially extended and completed by Moyal [61] (see also Harris [28] chapter 3). The more comprehensive case of infinitely many points was first discussed by Wold [92]. Then Khintchine [39] published a fundamental monograph on the general theory of streams of events which has led to a rapid development (cf. Ryll-Nardzewski [76], Moyal [61] §6, Matthes [55], Beutler and Leneman [10]). Contemporarily their concepts are being extended to random measures on topological spaces (Lee [45], Agnew [1], Mecke [56]).

This chapter is mostly expository, and aims to set out the basic definitions and properties of point processes in one dimension needed in this thesis. Drawing on many of the above accounts we give a formal definition and one useful generalisation to marked point sequences (Matthes [55]). Then follows a long list of known properties required later. A new proof for a basic relation is given to illustrate the potential of the marked point process notion. Finally some special processes are introduced for future reference.

1.2. The Definition of a Point Process

Probability theory is concerned with a space of events Ω, and a probability measure P defined on suitable subsets of the space. In the theory of stochastic processes, where the events are functions, the well-known Kolmogorov theorem (Kingman and Taylor [41] p.381) shows that P can be uniquely defined by extension of a consistent set of distributions over values of the functions at a finite set of points in their domain. It turns out that the same is true for stochastic point processes after some extra conditions are imposed.

Let Ω be the set of all countable sequences of real numbers $\{t_i\}$ without limit points and let $N(A)$ be the cardinality of the set $\{t_i \in A\}$ for all Borel sets A on the real line. Then $N(\cdot)$ is a counting measure (i.e. a non-negative integer-valued set function countably additive on the Borel sets). Since $\{t_i\}$ has no limit points $N(\cdot)$ is obviously finite on bounded sets. It is known (Moyal [61]) that there is a one-to-one correspondence between Ω and the set of all σ-finite counting measures $N(\cdot)$, which we can therefore also denote by Ω.

Consider now a set of functions $p(A_1,\ldots,A_k;\ r_1,\ldots,r_k)$ where $k,r_1,\ldots,r_k$ are non-negative integers and $A_1,\ldots,A_k$ are Borel sets. In order that they be the finite-dimensional distributions of a point process they must satisfy the following consistency conditions:

(1) $p(A_1,\ldots,A_k;r_1,\ldots,r_k) = p(A_{i_1},\ldots,A_{i_k};r_{i_1},\ldots,r_{i_k})$ for any permutation $(i_1,\ldots,i_k)$ of $(1,\ldots,k)$.

(2) $p(A_1,\ldots,A_k;r_1,\ldots,r_k) \geq 0$ and $\sum_{r=0}^{\infty} p(A_1,\ldots,A_k,A;r_1,\ldots,r_k,r)$
$= p(A_1,\ldots,A_k;r_1,\ldots,r_k)$.

(3) $p(A_1 \cup \ldots \cup A_k;r) = \sum_{r_1+\ldots+r_k=r} p(A_1,\ldots,A_k;r_1,\ldots,r_k)$ where the A_i are disjoint.

(4) If a sequence of bounded sets $A_k \downarrow \emptyset$, the null set, as k then $\lim_{k\to\infty} p(A_k;0) = 1$.

The fundamental result is

<u>Theorem 1.1.</u> (Moyal [61], Harris [28] p.55, Nawrotzki [63]). Corresponding to a set of functions $p(A_1,\ldots,A_k;r_1,\ldots,r_k)$ satisfying (1)-(4) there is a unique probability measure $\mathcal{P}$ defined on the σ-algebra $\mathcal{F}$ generated by the cylinder sets $\{N(\cdot);N(A_1) = r_1,\ldots,N(A_k) = r_k\}$ for which

$$\mathcal{P}\{N(A_1) = r_1,\ldots,N(A_k) = r_k\} = p(A_1,\ldots,A_k;r_1,\ldots,r_k).$$

A <u>point process</u> is specified by the triple $(\Omega, \mathcal{F}, \mathcal{P})$. Since the Borel sets on the real line can be generated by half-open intervals (Halmos [26] p.62) we may consider the $p(\cdot;\cdot)$ of Theorem 1.1. only for disjoint half-open intervals. This is sometimes convenient.

A word on conventions is in order. We use 'point process' and 'random stream' interchangeably, dropping the adjectives if no ambiguity arises. Notationally, we write $N(\cdot)$ for an arbitrary point process, because of the 1:1 correspondence between points and counts, though our outlook may vary between the two. $N(\cdot)$ is *always* assumed to be finite on bounded Borel sets. All our point processes evolve on the real line, though there is no difficulty in principle in extending the definitions at least to higher dimensions (Goldman [22]) and arbitrary spaces (Moyal [61]).

We mention one extension of this basic scheme, due to Matthes [55]. With each point t_i is associated a mark k_i from a fixed measurable space $[K, \mathcal{K}]$, so that the event space is $\Omega_K = (-\infty, \infty) \times K$ with counting measure $N(\cdot)$ taken over sets $I \times L$, where I is a Borel set and $L \subset \mathcal{K}$. A <u>marked point process</u> $(\Omega_K, \mathcal{F}_{\mathcal{K}}, P)$ can now be defined as before, and has a variety of applications in theory ([55] and 1.4) and practice (energies of earthquakes, velocities of cars or electrons).

1.3. Some Properties of Point Processes

A point process $N(\cdot)$ is <u>stationary</u> if all its finite-dimensional distributions are invariant under translation. This concept occurs frequently in our work. If we define the translation operator T by $T^t A = \{x : x+t \in A\}$, for some Borel set A, then stationarity means

$$p(T^t A_1, \ldots, T^t A_k; r_1, \ldots, r_k) = p(A_1, \ldots, A_k; r_1, \ldots, r_k).$$

We see that if $N(\cdot)$ is stationary, $N(-\infty, \infty) \overset{a.s.}{=} 0$ or ∞ and $P\{N(\{x\}) > 0\} = 0$ for all singletons $\{x\}$; cf. Ryll-Nardzewski [76].

A stationary stream is <u>orderly</u> if†

$$P\{N[0,t) \geq 2\} = o(t) \quad \text{as} \quad t \downarrow 0$$

(Khintchine [39] §1). It has <u>no multiple occurrences</u> if each point of the stream has multiplicity one. Clearly this is true for orderly streams.

A stationary point process is <u>ergodic</u> if all events invariant under translation have probability zero or one. A necessary and sufficient

† We write $N[a,b)$ instead of $N([a,b))$, for typographical convenience.

condition for ergodicity is (Rosenblatt [73] p.110)

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t P(A \cap T^{-\tau}B)d\tau = P(A).P(B) \qquad \text{for all} \quad A,B \in F. \qquad (1.1)$$

The process is <u>mixing</u> if

$$\lim_{t\to\infty} P(A \cap T^{-t}B) = P(A).P(B) \qquad \text{for all} \quad A,B \in F, \qquad (1.2)$$

and is <u>weakly mixing</u> if

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t |P(A \cap T^{-\tau}B) - P(A).P(B)|d\tau = 0 \quad \text{for all} \quad A,B \in F. \quad (1.3)$$

Obviously mixing $\Rightarrow$ weak mixing $\Rightarrow$ ergodic (Matthes [55]).

The <u>superposition</u> of n independent point processes $N_1(\cdot),\ldots,N_n(\cdot)$ is simply the aggregation of all their points, and is written $\Sigma_{i=1}^{n} N_i(\cdot)$. The finite-dimensional distributions are derived by convolution, so the probability law of the superposition follows from Theorem 1.1.

A sequence $\{N_n(\cdot)\}$ of point processes <u>converges in distribution</u> to a point process $N(\cdot)$, in symbols $N_n(\cdot) \xrightarrow{d} N(\cdot)$, if all the finite-dimensional distributions converge in the usual sense. For a stronger concept, see Jiřina [33].

We can define integrals with respect to a point process (counting measure) by†

$$\int f(t)\, dN(t) = \sum_i f(t_i) \qquad (1.4)$$

for such functions f as ensure the integral exists (Moyal [61]). Obviously the class of functions which vanish outside some bounded interval makes the integral finite, and other criteria are considered later.

---

† $\int$ will always mean $\int_{-\infty}^{\infty}$.

The moment structure of a point process is determined by a set of moment measures, namely expectations of product counting measures on $\Omega \times \ldots \times \Omega$. The first moment measure $M(\cdot) = E\{N(\cdot)\}$ (it is easily shown to actually be a measure) is always assumed to be a Borel measure, (Halmos [55] p.223). Thus $E\{N(A)\} < \infty$ for a bounded set A which implies $N(A) < \infty$ with probability one, an assumption already made. It is easy to see that for a stationary process $M(A) = m|A|$, where $|A|$ is the Lebesgue measure of A and $m = E\{N[0,1)\}$ is the intensity of the process.

The higher moment measures are defined like $M_{(2)}(A \times B) = E\{N(A)N(B)\}$, and it is obvious that they have concentrations on subsets of lower dimension. This is inconvenient, and it is more useful to work with the factorial moment measures $M_n(\cdot)$ defined by analogy with the usual factorial moments (Moyal [61], Vere-Jones [88])†. For example

$$M_2(A \times B) = E\{N(A)N(B)\} - E\{N(A \cap B)\},$$

which equals $M_{(2)}(\cdot)$ if $A \cap B = \emptyset$ and reduces to the usual factorial moment if $A = B$. Under reasonable conditions ($N(\cdot)$ stationary and orderly for example) the aberrant concentrations now vanish.

We can also define factorial cumulant measures $C_n(\cdot)$ by the usual moment-cumulant formulae. If these measures are absolutely continuous we may speak of factorial moment and cumulant densities

† Note that the notation of these papers is reversed.

$m_n(\cdot)$ and $c_n(\cdot)$, which are the product densities of Bartlett [6] pp.83, 122. These densities exist for stationary orderly streams under mild extra conditions. For a stationary stream the <u>covariance density</u> $c_2(u, v) = c_2(u-v)$, and similarly in higher dimensions.

Fubini's theorem now shows that for integrals (1.4)

$$E\{\int f(t)\, dN(t)\} = \int f(t)\, M(dt) \tag{1.5}$$

if $\int |f(t)| M(dt) < \infty$, in which case the integral exists. Similar considerations apply for higher moments; for instance

$$D\{\int f(t)\, dN(t)\} = \int f^2(t)\, M(dt) + \iint f(t)\, f(u)\, C_2(dt \times du) \tag{1.6}$$

if $C_2(\cdot)$ is a Borel measure on the plane and the R.H.S. integrals converge absolutely. By taking f as an indicator function we get the moments of $N[0, t)$ in terms of moment measures. These results are in [87], [88] and Cox and Lewis [14] chapter 4.

A point process is <u>weakly stationary</u> if its first and second order moments are invariant under translation. Using the identity

$$2\, \mathrm{Cov}\{N[0,t), N[t+v, t+v+u)\} = D(t+v+u) - D(t+v) - D(u+v) + D(v), \tag{1.7}$$

where $D(x) = D\{N[0,x)\}$, Daley [15] has proved

<u>Lemma 1.1</u>. A weakly stationary point process uniquely determines a σ-finite measure $\mu(\cdot)$ on $[0, \infty)$ for which $\theta^{-1}\mu([0,\theta]) \to \frac{\alpha}{\pi}$ as $\theta \to \infty$, $\alpha = \lim_{u \downarrow 0} u^{-1} D(u) < \infty$, and

$$D(u) = u^2 \mu(\{0\}) + \int_{(0,\infty)} \left(\frac{\sin \frac{1}{2}\theta u}{\frac{1}{2}\theta}\right)^2 \mu(d\theta) \tag{1.8}$$

This has close connections with the spectral analysis of point processes introduced in Bartlett [4] for orderly streams. If $\gamma(d\theta) = \mu(d\theta) - \frac{m d\theta}{\pi}$ is totally finite then $c_2(u) = \int_0^\infty \cos \theta u \, \gamma(d\theta)$, and if $c_2(\cdot)$ is integrable, as assumed by Bartlett, then $\gamma(\cdot)$ is absolutely continuous and defines a spectral density function $g(\theta)$.

Our survey of properties is necessarily brief, and reflects the requirements of the thesis. Further general results for point processes may be found in Khintchine [39], Ryll-Nardzewski [76], Slivnyak [78], Matthes et al. [55], [36], [20], Fieger [19], Cox and Lewis [14], and Beutler and Leneman [10].

1.4. A Theorem on Stationary Streams

A fundamental result for stationary point processes is that the parameter λ, defined by $\lambda = \lim_{t \downarrow 0} t^{-1} P\{N[0,t) \geq 1\}$, always exists (Khintchine [39] §7). Obviously $\lambda \leq m$, so that we always have $\lambda < \infty$. Korolyuk's theorem ([39] §11 and Zitek [93]) states that orderliness is necessary and sufficient for $\lambda = m$.

As we know, orderliness implies no multiple occurrences. Conversely, Dobrushin's lemma (Volkonskii [89]) says that if there are no multiple occurrences and $m < \infty$ then the stream is orderly.

Any extension of Korolyuk's theorem to non-orderly streams requires the notion of a 'batch-size' distribution for the number in a multiple occurrence, if such a quantity exists. Its existence may be deduced by the powerful analytical methods employed in Slivnyak [78], Beutler and Leneman [10] and Fieger [19] for other purposes. Recently Milne [58], using a technique of Leadbetter [43], has given a neat proof of more general extensions. We apply the same technique to prove Khintchine's result for a marked point process and show how some theorems for general stationary streams flow from it.

Lemma 1.2. For a stationary marked point sequence $(\Omega_K, \mathcal{F}_K, \mathcal{P})$

$$\lim_{t \downarrow o} \frac{\mathcal{P}\{N([0,t)\times L) = 1\}}{\mathcal{P}\{N([0,t)\times K) \geq 1\}} = \pi(L) \qquad L \subset K$$

exists and is a probability measure over K

Proof: Let $\chi_{in}(L) = 1$ if $N([\frac{i}{n}, \frac{i+1}{n}) \times L) \geq 1$

$= 0$ otherwise

$i = 0,1,\ldots,n-1$ and $L \subset K$. A direct imitation of Leadbetter's proof now gives

$$\lim_{t \downarrow o} t^{-1} \mathcal{P}\{N([0,t)\times L) \geq 1\} = E\{N([0,1)\times L)\} = \lambda(L), \tag{1.9}$$

$\lambda(K) = \lambda$, the parameter. An identical procedure with ≥ 2 establishes Dobrushin's lemma and hence that (1.9) is still true with $= 1$ on the L.H.S. Thus

$$\lim_{t \downarrow o} \frac{\mathcal{P}\{N([0,t)\times L) = 1\}}{\mathcal{P}\{N([0,t)\times K) \geq 1\}} = \frac{\lambda(L)}{\lambda} = \pi(L) \qquad L \subset K$$

Since $\lambda(\cdot)$ is obviously a finite measure over K the proof is complete. ▼

Now for a given stationary stream $N(\cdot)$ we define a new orderly stream $N^*(\cdot)$ by counting multiple events of $N(\cdot)$ as single events. Thus $N(\cdot)$ is a marked point process with $K = \{1,2,\ldots\}$, the mark of an $N^*(\cdot)$ event being its multiplicity. If $N^*(\cdot)$ is stationary, then taking $L = \{k\}$ shows that π_k, $k = 1,2,\ldots$ is the desired batch-size distribution, i.e. the probability of k events given that at least one has occurred. The mean batch size is $a = \Sigma_{k=1}^{\infty} k\pi_k$ and $\lambda a = \lambda\, \Sigma_{k=1}^{\infty} k\pi_k = \Sigma_{k=1}^{\infty} k\lambda_k = m < \infty$.

It remains only to show that $N^*(\cdot)$ is stationary when $N(\cdot)$ is. Suppose this were not so. If $P^*(z_1,\ldots,z_k;\ I_1,\ldots,I_k)$ are the p.g.fs of the joint distributions of $N^*(\cdot)$, $|z_i| \leq 1$ $i = 1,\ldots,k$, then for some integer $k > 0$, some intervals $I_1,\ldots,I_k$, and some t, we must have

$$P^*(z_1,\ldots,z_k;\ I_1,\ldots,I_k) \neq P^*(z_1,\ldots,z_k;\ T^tI_1,\ldots,T^tI_k) \quad (1.10)$$

in a neighbourhood of $(0,\ldots,0)$. Otherwise there is a sequence of points, having $(0,\ldots,0)$ as a limit point, for which equality holds in (1.10), so that it will hold for all $z_1,\ldots,z_k$ in the unit sphere, a contradiction. Hence taking $(z_1,\ldots,z_k) = (0,\ldots,0)$ in (1.10), as we may, we get

$$p^*(0,\ldots,0;\ I_1,\ldots,I_k) \neq p^*(0,\ldots,0;\ T^tI_1,\ldots,T^tI_k).$$

But obviously $p^*(0,\ldots,0;\ I_1,\ldots,I_k) = p(0,\ldots,0;\ I_1,\ldots,I_k)$, which is stationary by assumption. This contradiction proves that $N^*(\cdot)$ is stationary, and so *any* stationary stream may be regarded as a stationary marked point process.

<u>Theorem 1.2</u>. For a stationary point process the batch-size distribution π_k exists and has a mean equal to $\frac{m}{\lambda}$.

1.5. Some Examples of Point Processes

The most important point process is of course the Poisson process, characterised by being the only stationary orderly stream with independent increments i.e. $N(I)$, $N(J)$ independent if $I \cap J = \emptyset$. (Khintchine [39] §3). Its finite-dimensional distributions are thus products of the usual Poisson distributions. If its parameter (= intensity) is λ, we speak of a Poisson (λ) process.

As is well-known, the Poisson process is a special case of a renewal process, for which the intervals between events are independent (Khintchine [39] §13, Smith [79]).

Without orderliness we get the compound Poisson process ([39] §8), which may have multiple occurrences and is a special case of the situation in 1.4. Omitting only stationarity leads to a non-homogeneous Poisson process, where the parameter is now a function of time $\lambda(t)$ ([39] §5). When both hypotheses are dropped, the resulting process is characterised in Khintchine [38] and Fieger [19]. A more general non-homogeneous Poisson process assumes that the intensity is a measure $\Lambda(\cdot)$, which is obviously the expectation measure (cf. Moyal [61]).

Another generalisation is to take the time-dependent parameter $\lambda(t)$ as a realisation of a stochastic process $\Lambda(t)$. This defines the class of doubly stochastic Poisson processes, with mean process Λ, introduced by Cox [13] and studied further by Bartlett [4], Kingman [40] and Mecke [57]. Of particular interest is a *linear* mean process, defined by

$$\Lambda(t) = \int_{-\infty}^{\infty} f(t-u)dX(u) \tag{1.11}$$

where f is non-negative, integrable and square-integrable and $X(\cdot)$ is an additive homogeneous process with non-negative increments (Bartlett [6] p.161). If we assume $E\{X^2(t)\} < \infty$ then (1.11) exists as a mean square convergent integral (Doob [17]) §9.2). Such processes will be called linear stochastic Poisson processes. They are of considerable importance in the spectral analysis of point processes (Bartlett [4]), being a natural generalisation of the Poisson process for which a tractable theory can be formulated. Also, as noted by Moran [59], the statistical properties of the process can be expressed in terms of f, and should be easy to develop.

The properties of doubly stochastic processes are easily derived by conditional arguments (Cox and Lewis [14] §7.2). They are stationary if Λ is stationary. For linear stochastic Poisson processes we only note at present that the index of dispersion $I(t) = \frac{D(t)}{M(t)}$ is always convergent as $t \to \infty$, since f is integrable, and that they are never renewal processes as Kingman's characterisation [40] for these involves a mean process which is constant over random intervals, an obvious impossibility for (1.11).

A further class is the infinitely divisible point processes introduced by Matthes [54] and Lee [45] and studied in Kersten and Matthes [36], [37], Lee [46]. There are several possible definitions:

(A) If a sequence of independent *uniformly asymptotically negligible* (*u.a.n.*) point processes $\{N_{n,i}(\cdot)\}$ $i = 1,\dots,s_n$ $n = 1,2,\dots$ satisfies $\Sigma_{i=1}^{s_n} N_{n,i}(\cdot) \xrightarrow{d} N(\cdot)$ as $n \to \infty$ then $N(\cdot)$ is infinitely divisible. A sequence is u.a.n. if for each bounded interval I

$$\lim_{n\to\infty} \max_{1\le i\le s_n} P\{N_{n,i}(I)>0\} = 0.$$

(B) If there is a sequence $\{N_{n,i}(\cdot)\}$ of independent identically distributed point processes such that $N(\cdot) \overset{d}{=} \Sigma_{i=1}^{n} N_{n,i}(\cdot)$ for all $n = 1,2,\ldots$ then $N(\cdot)$ is infinitely divisible.

(C) If all the finite-dimensional distributions of $N(\cdot)$ are infinitely divisible then $N(\cdot)$ is infinitely divisible.

For the equivalence of (A)-(C) see e.g. Goldman [23]. From (C), the finite-dimensional distributions of an infinitely divisible $N(\cdot)$ are all compound Poisson with p.g.fs of the form $\exp(\Sigma_m a(m)z^m)$. In fact we have

Theorem 1.3. (Kersten and Matthes [36], Lee [45]). To each infinitely divisible point process (Ω, F, P) there corresponds exactly one measure $\tilde{P}(\cdot)$ on F with the properties

(1) $\tilde{P}\{N(I_1) = r_1,\ldots,N(I_k) = r_k\} = a(r_1,\ldots,r_k;\ I_1,\ldots,I_k)$ for all integers $k,r_1,\ldots,r_k$ and intervals $I_1,\ldots,I_k$.

(2) $\tilde{P}\{\emptyset\} = 0$.

(3) $\tilde{P}\{N(I) \neq 0\} < \infty$ for all bounded intervals I.

The KLM (Kersten-Lee-Matthes) measure $\tilde{P}$ is stationary if and only if P is.

We may now define regular and singular infinitely divisible point processes (Matthes [54]) corresponding to $\tilde{P}$ concentrated on members of Ω having finitely and infinitely many points respectively. Regular processes can be characterised in terms of cluster processes (see 3.2).

As an example of a singular process we introduce the <u>singular Poisson process</u> $E_{\lambda,N}$, defined as the n-fold superposition of an arbitrary point process $N(\cdot)$, where n is a Poisson random variable with parameter λ. If $N(\cdot)$ is stationary, $E_{\lambda,N}$ is obviously a singular infinitely divisible point process.

For examples of KLM measures see Lee [46].

2. THE PROBABILITY GENERATING FUNCTIONAL

2.1. Introduction

The probability generating functional (p.g.fl) of a point process $N(\cdot)$ is the principal tool used in this thesis. It originated in work on population processes (Kendall [34], Bartlett and Kendall [7], Moyal [61]), but has recently been extended from this essentially finite situation to cases where infinitely many points may occur (Vere-Jones [87]). This is more difficult, and it is with this case we are principally concerned.

The basic properties of the p.g.fl are derived, in close analogy with the p.g.f. it generalises, and some examples given. We discuss its relation to the factorial moment measures of $N(\cdot)$. Finally, we characterise mixing and ergodic properties of a point process by relations involving the p.g.fl, and give applications.

2.2. Definitions and Properties of the p.g.fl

We consider throughout a point process $N(\cdot)$ whose expectation measure is a Borel measure. It may or may not be a.s. finite.

Definition. The p.g.fl of $N(\cdot)$ is defined by

$$G[\xi] = E\{\exp \int \log \xi(t)\, dN(t)\} \qquad (2.1)$$

for a suitable class of functions ξ.

Suitable classes will be discussed shortly. First we give two alternative statements of (2.1) which are both useful in certain cases. Clearly equivalent is

$$G[\xi] = E\{\prod_i \xi(t_i)\} \tag{2.2}$$

where the $\{t_i\}$ are the times of occurrence of the points. If the point process is *a.s. finite* and

$$p_n = P\{N(-\infty,\infty) = n\}$$

$U_n(t_1,\ldots,t_n)$ = distribution function of the t_i given $N(-\infty,\infty) = n$,

then

$$G[\xi] = \sum_{n=0}^{\infty} p_n \int\ldots\int \xi(t_1)\ldots\xi(t_n)dU_n(t_1,\ldots,t_n) \tag{2.3}$$

For a.s. finite processes these definitions are due to Moyal [61]; in general, see Vere-Jones [87].

It is clear that heuristically the p.g.fl is an extension of the multivariate p.g.f. to the 'generating function' of an infinite set of 'random variables' $dN(t)$. We expect then that its properties will be similar to those of the p.g.f. and later we see this is generally true.

To ensure that the p.g.fl is non-trivial, the exponent in (2.1) must be finite with probability one. Motivated by the analogy with the p.g.f. we consider functions ξ such that for all real t

$$0 \leq \xi(t) \leq 1 \tag{2.4}$$

<u>Definition</u>. If ξ is measurable and satisfies (2.4) then

$\xi \in V$ if ξ vanishes outside a bounded interval

$\xi \in L(N)$ if $\int|\log\xi(t)|M(dt) < \infty$.

These are the classes introduced by Vere-Jones [87]. We must further decide what happens in (2.1) at zeros of $\xi(t)$. Isolated zeros are no problem, and if $\xi(t) \equiv 0$ over some set A the exponential in (2.1) is taken as zero, unless $N(A) = 0$ when it equals one. Now since $M(\cdot)$ is a Borel measure we have (cf. [87])

Lemma 2.1. $G[\xi]$ is non-trivial if

(i) $N(\cdot)$ is a.s. finite

or (ii) $1 - \xi \varepsilon V$

or (iii) $\xi \varepsilon L(N)$; in this case $\int[1 - \xi(t)]M(dt) < \infty$ also.

We make the convention that *if in future we use a p.g.fl without specific reference to ξ we are assuming that it belongs to either of the above classes.*

The next lemma shows the fundamental role that the p.g.fl plays in the theory of point processes.

Lemma 2.2. (Moyal [61], Vere-Jones [87]). The p.g.fl is uniquely determined by $N(\cdot)$ and, conversely, knowledge of the p.g.fl completely determines the probability structure of $N(\cdot)$.

An important consequence for us is that the p.g.fs of the joint distributions of $N(\cdot)$ are derived from the p.g.fl by setting $\xi(t)$ equal to a simple function (Halmos [26] p.84). Because of a previous convention, expressions such as $P\{N(A) = 0\}$ come from putting $\xi(t) \equiv 0$ for $t \varepsilon A$.

We now develop further properties of the p.g.fl. Obviously

(a) $0 \leq G[\xi] \leq 1$,

(b) G is monotonic i.e. $\xi_1 \leq \xi_2 \Rightarrow G[\xi_1] \leq G[\xi_2]$

(c) Continuity. We might hope that the p.g.fl is always continuous, in the sense that $\xi_n(t) \to \xi(t)$ pointwise as $n \to \infty$ implies $G[\xi_n] \to G[\xi]$. To see this is in general false, consider

$$\xi_n(t) = 1 - (1-z)\chi_{[n,n+1)}(t),$$

so that $G[\xi_n] = E\{z^{N[n,n+1)}\}$. Clearly $\xi_n(t) \to 1$ pointwise as $n \to \infty$ yet $G[\xi_n]$ need not tend to one (take $N(\cdot)$ stationary for example). However we do have

Theorem 2.1. The p.g.fl is continuous if one of the following holds

(i) $N(\cdot)$ is a.s. finite

(ii) the $1 - \xi_n \in V$ and have a common interval outside which they all vanish

(iii) $\xi_n(t) \geq \tilde{\xi}(t)$ for all n and $\tilde{\xi} \in L(N)$

(iv) $\int |\xi_n(t) - \xi(t)| M(dt) \to 0$ as $n \to \infty$

(v) $\xi \in L(N)$ and given $\varepsilon > 0$ there is $T(\varepsilon)$ such that

$$\int_{|t|>T} |\log \xi_n(t)| M(dt) < \varepsilon \quad \text{or} \quad \int_{|t|>T} [1-\xi_n(t)] M(dt) < \varepsilon \quad \text{for all } n.$$

Proof. In each case we prove that $\int \log \xi_n(t) dN(t) \to \int \log \xi(t) dN(t)$ in some sense, as then the bounded convergence theorem ensures continuity. (i) is obvious (see Harris [28] p.58). The assumptions in (ii) effectively reduce it to (i) and (iii) follows directly from dominated convergence. To prove (iv) we have the simple identity (Moyal [62])

$$\prod_{i=1}^{n} \zeta(t_i) - \prod_{i=1}^{n} \eta(t_i) = \sum_{i=1}^{n} [\xi(t_i) - \eta(t_i)] \eta(t_1)..\eta(t_{i-1}) \zeta(t_{i+1})..\zeta(t_n), \tag{2.5}$$

valid for $n = 1,2,\ldots$ and any functions ζ, η. With $\zeta = \xi_n$, $\eta = \xi$

we see from (2.2) that

$$|G[\xi_n] - G[\xi]| \leq E\{\int|\xi_n(t) - \xi(t)|dN(t)\}$$

$$= \int|\xi_n(t) - \xi(t)|M(dt) \to 0 \qquad \text{as} \quad n \to \infty.$$

Of course, to write down $G[\xi]$ implies that ξ is suitable in the sense of Lemma 2.1. Conditions (ii), (iii), (iv) all ensure that the limit is suitably integrable. If now we assume this explicitly then taking $\int_{|t|\leq T} + \int_{|t|>T}$ and recalling that $M(\cdot)$ is a Borel measure, (i) and (iv) can be used to establish (v). ▼

(d) Characterisation. An interesting question is, what functionals over V are p.g.fls? In the a.s. finite case, Harris [28] p.58 gives a result involving those functionals whose arguments are simple functions; a similar theorem for Laplace functionals may be found in Jiřina [32]. Moyal [61] has a characterisation in terms of restrictions to finite subsets of the population space. Our result is not essentially new, and draws on all these three theorems.

The basic technique is due to Harris ([28] p.53), namely that a set of functions $p_0(A_1,\ldots,A_k;\ r_1,\ldots,r_k)$ satisfying (1)-(4) for *disjoint* sets A_i (with slight changes in condition (3)) can be uniquely extended to functions $p(A_1,\ldots,A_k;\ r_1,\ldots,r_k)$ satisfying (1)-(4) and agreeing with the $p_0(\cdot;\cdot)$ whenever the A_i are disjoint.

Theorem 2.2. Suppose we have a functional $G[\xi]$ defined whenever $1 - \xi \in V$ and continuous for sequences ξ_n satisfying Theorem 2.1(ii). Further, if $1 - \xi$ is a simple function in V, i.e. $1 - \xi(t) = \Sigma_{i=1}^{k}(1-z_i)\chi_{A_i}(t)$ where the Borel sets A_i are disjoint, suppose

$G[\xi] = P(z_1,\ldots,z_k;\ A_1,\ldots,A_k)$ is the p.g.f. of an n-dimensional random variable. Then $G[\xi]$ is the p.g.fl of a point process.

Proof. Let the n-dimensional distributions associated with $P(z_1,\ldots,z_k;\ A_1,\ldots,A_k)$ be $p_o(A_1,\ldots,A_k;\ r_1,\ldots,r_k)$, for disjoint A_i. Now in the consistency conditions

(1) holds because $p_o(\cdot;\cdot)$ is certainly a probability distribution,

(2) holds by the obvious relation $P(z_1,\ldots,z_k,1;\ A_1,\ldots,A_k,A_{k+1}) = P(z_1,\ldots,z_k;\ A_1,\ldots,A_k)$,

(3) holds as $P(z,z;\ A_1,A_2) = P(z;\ A_1 \cup A_2)$ obviously, and this can be extended to the disjoint collections $\{A_{ij}\}$ making up the A_i (for explanation see Harris [28] pp.53-54),

(4) holds by the continuity of G for functions of the form $\xi_n(t) = 1 - (1-z)\chi_{A_n}(t)$, bounded $A_n \downarrow \emptyset$, for which $\xi_n \to 1$ pointwise.

So we may extend the $p_o(\cdot;\cdot)$ uniquely to a consistent set of functions, in the sense of 1.2, and by Theorem 1.1 there is a unique point process $N(\cdot)$ whose finite-dimensional distributions over disjoint Borel sets $A_1,\ldots,A_k$ are $p_o(A_1,\ldots,A_k;\ r_1,\ldots,r_k)$. $N(\cdot)$ has a p.g.fl $G^*[\xi]$, $1 - \xi \in V$, which must agree with $G[\xi]$ over simple functions. But arbitrary $1 - \xi \in V$ can be approximated uniformly by an increasing sequence of simple functions (Halmos [26] p.85), and G, G^* are continuous for such sequences by hypothesis and Theorem 2.1(ii) respectively. Therefore they agree for all ξ such that $1 - \xi \in V$. ▼

(e) Convergence. One of the most useful properties of the p.g.f. is that it provides necessary and sufficient conditions for the convergence of discrete probability measures. A similar result holds for convergence in distribution of point processes, namely

Lemma 2.3. (Vere-Jones [87]). A sequence of point processes $\{N_n(\cdot)\}$ converges in distribution to a point process $N(\cdot)$ if and only if the associated p.g.fls converge i.e. $G_n[\xi] \to G[\xi]$ for $1 - \xi \in V$.

The sufficiency is obvious, and the necessity is proved by approximating ξ above and below by simple functions and using the monotonicity and continuity (Theorem 2.1(ii) again) of the p.g.fl. We remark that a stronger assertion, corresponding to the weak convergence of measures, is given for Laplace functionals by Jiřina [33].

Lemma 2.3 can be generalised slightly if we assume only that the sequence $G_n[\xi]$ converges to some functional $G[\xi]$, $1 - \xi \in V$. Specifically, we have

Theorem 2.3. A sequence of point processes $\{N_n(\cdot)\}$ converges in distribution to a point process $N(\cdot)$ if and only if the p.g.fls $G_n[\xi]$ converge to a functional $G[\xi]$, $1 - \xi \in V$, which is continuous for sequences $\xi_m(t) \to 1$ pointwise. Then $G[\xi]$ is the p.g.fl of $N(\cdot)$.

Proof. The 'only if' part follows as before. For the sufficiency, take a simple function $\xi(t) = 1 - \Sigma_{i=1}^{k}(1-z_i)\chi_{A_i}(t)$. $G_n[\xi] = P_n(z_1,\ldots,z_k;\ A_1,\ldots,A_k)$ is therefore the p.g.f. of some joint distribution of $N_n(\cdot)$ which converges to a function $G[\xi] =$

$P(z_1,\ldots,z_k;\ A_1,\ldots,A_k)$ that is continuous as $(z_1,\ldots,z_k) \to (1,\ldots,1)$. So $G[\xi]$ is also a p.g.f., by a standard result, and therefore all the joint p.g.fs of $N_n(\cdot)$ converge to a set of p.g.fs. We must now prove that these limit p.g.fs are consistent, in the sense of satisfying (1)-(4). It is easy to see that

$$(1)' \quad P(z_1,z_2;A_1,A_2) = \lim_{n\to\infty} P_n(z_1,z_2;A_1,A_2) = \lim_{n\to\infty} P_n(z_2,z_1;A_2,A_1)$$
$$= P(z_2,z_1;A_2,A_1) \quad \text{etc,}$$

as the P_n satisfy (1).

$(2)'$ $P_n(z_1,z_2;A_1,A_2) \to P(z_1,z_2;A_1,A_2)$ by hypothesis

and $P_n(z_1,1;A_1,A_2) = P_n(z_1;A_1)$ by (2)

i.e. $P(z_1,1;A_1,A_2) = P(z_1;A_1)$ etc.

$(3)'$ $P_n(z_1,z_2,z_3;A_1,A_2,A_1 \cup A_2) \to P(z_1,z_2,z_3;A_1,A_2,A_1 \cup A_2), A_1 \cap A_2 = \emptyset$

and $P_n(z_1,z_2,z_3;A_1,A_2,A_1 \cup A_2) = P_n(z_1+z_3,z_2+z_3;A_1,A_2)$ by (3)

$\to P(z_1+z_3,z_2+z_3;A_1,A_2)$ etc.

$(4)'$ Take a sequence of bounded sets $A_m \downarrow \emptyset$ and define a simple function $\xi_m(t) = 1-(1-z)\chi_{A_m}(t)$. Then $\xi_m(t) \to 1$ pointwise as $m \to \infty$ and so $G[\xi_m] \to G[1] = 1$. But $G[\xi_m] = P(z;A_m)$.

So the limit distributions form a consistent set, and by Theorem 1.1 there is a unique point process $N(\cdot)$ having them as its finite-dimensional distributions. If $N(\cdot)$ has p.g.fl $G^*[\xi]$, $1 - \xi \in V$, then the 'only if' part gives $G_n[\xi] \to G^*[\xi]$ whereas $G_n[\xi] \to G[\xi]$ by hypothesis. Thus $G[\xi]$ is the p.g.fl of $N(\cdot)$. ▼

(f) Superposition. As we might expect from the p.g.f. analogy, superposition of independent streams is equivalent to multiplication of the p.g.fls. Thus if $N_1(\cdot),\ldots,N_k(\cdot)$ are independent point processes with p.g.fls $G_1[\xi],\ldots,G_k[\xi]$, the p.g.fl of $\Sigma_{i=1}^k N_i(\cdot)$ is

$$G[\xi] = \prod_{i=1}^{k} G_i[\xi] \tag{2.6}$$

This is immediately apparent from (2.1) or (2.2).

In view of Lemma 2.2 the p.g.fl is obviously a very powerful aid in the study of point processes, containing, as it does, information about all aspects of the process. Its disadvantage, however, is that it is rarely obtainable in closed form unless the point process involved is related to the Poisson process. We give some examples of this below. Nevertheless it is a valuable tool in a variety of theoretical problems, such as characterisations of ergodic and mixing properties (Section 2.4), the theory of cluster processes (Chapter 3) and limit theorems for point processes (Chapter 4).

The fundamental point process is, as we have seen, the Poisson process. The completely random property of this process makes the calculation of its p.g.fl particularly simple. We have, for a Poisson $(\Lambda(\cdot))$ process

$$G[\xi] = \exp\{-\int[1 - \xi(t)]\Lambda(dt)\} \tag{2.7}$$

cf. Ryll-Nardzewski [75], Moyal [61], Shiryaev [77]. In particular, for the stationary Poisson (λ) process

$$G[\xi] = \exp\{-\lambda\int[1-\xi(t)]dt\}. \tag{2.8}$$

These results are used repeatedly.

The p.g.fls of many point processes related to the Poisson process may be readily deduced from (2.7). We do this only for the doubly stochastic Poisson process with a stationary mean process $\Lambda(t)$. First we need a new concept.

Definition. The Laplace functional of a non-negative stochastic process $Y(t)$ is

$$L_Y[\xi] = E\{e^{-\int Y(t)\xi(dt)}\}, \tag{2.9}$$

where $\xi(\cdot)$ is a totally finite measure on the Borel sets of the line. Clearly, a sufficient condition for (2.9) to be non-trivial is that $Y(t)$ be stationary with finite mean. This definition is based on Shiryaev [77]; for a related idea see Jiřina [32], [33].

Now, conditional on a realisation $\lambda(t)$ of $\Lambda(t)$, the doubly stochastic Poisson process is a non-homogeneous Poisson process and so from (2.7) $G[\xi] = E_\Lambda\{e^{-\int\lambda(t)[1-\xi(t)]dt}\}$ i.e.

$$G[\xi] = L_\Lambda[\int_{-\infty}^{t}[1 - \xi(u)]du], \tag{2.10}$$

a result due to Bartlett (discussion to Cox [13]; see also [4] and Mecke [57]). This emphasises the close connection between the statistical properties of a doubly stochastic Poisson process and its mean process.

2.3. The Connection with Moment Measures

We saw in 1.3 that with any point process there is associated a set of factorial moment measures. We noted that these measures are more convenient than the usual moment measures, and it is because they are so intimately related to the p.g.fl that we prefer to work with this functional rather than the characteristic or moment generating functionals.

Formally, the relationship is very simple. If all relevant moment measures exist then on expanding the logarithm and exponential of (2.1) in their power series we have (Moyal [61], Vere-Jones [87])

$$G[1-\xi] = 1 + \sum_{k=1}^{\infty} \frac{(-1)^k}{k!} \int \ldots \int \xi(t_1)\ldots\xi(t_k) M_k(dt_1,\ldots,dt_k) \quad (2.11)$$

and

$$H[1-\xi] = \log G[1-\xi] = \sum_{k=1}^{\infty} \frac{(-1)^k}{k!} \int\ldots\int \xi(t_1)\ldots\xi(t_k)\; C_k(dt_1,\ldots,dt_k) \quad (2.12)$$

To put these relations on a more rigorous footing we first establish a Taylor-type expansion of the p.g.fl to a finite number of terms, in analogy with results for characteristic functions (Lukacs [53] p.31). This requires

<u>Lemma 2.4</u>. If $\xi_1,\ldots,\xi_N$ are real numbers, $0 < \xi_i < 1$ $i = 1,\ldots,N$, where N is an arbitrary positive integer, and we take

$$Q_N = \prod_{i=1}^{N} (1-\xi_i) = 1 - \sum_{i=1}^{N} \xi_i + \sum_{i_1=1}^{N} \sum_{i_2>i_1} \xi_{i_1}\xi_{i_2} - \ldots + (-1)^N \xi_1\ldots\xi_N$$

$$= 1 - q_1 + q_2 - \ldots + (-1)^N q_N \quad (2.13)$$

where $q_k = \sum_{1\le i_1<\ldots<i_k\le N} \xi_{i_1}\ldots\xi_{i_k}$, and

$$S_N^{(m)} = 1 - q_1 + q_2 - \ldots + (-1)^m q_m, \quad (2.14)$$

$m = 1,\ldots,N$ (so that $S_N^{(m)}$ is the m^{th} partial sum in (2.13)), then

$$S_N^{(2m-1)} \le Q_N \le S_N^{(2n)}$$

for all N and $m,n = 1,2,\ldots,[\frac{N}{2}]$ †

† $[x]$ is the greatest integer less than or equal to x.

Proof. A simple direct proof by induction is possible. However we note that this is a very special case of Bonferroni's inequalities on the probabilities of combined events (Moran [60] §1.18), if we interpret the ξ_i as the probabilities of a set of independent events A_i. Putting $m = 1$ in Theorem 1.5 of [60], $1 - P_1$ becomes equal to Q_N because of the assumed independence, and the lemma follows immediately. ▼

Corollary. (i) $0 \leq Q_N - S_N^{(2m-1)} \leq q_{2m}$ $\qquad 2m \leq N$

(ii) $0 \leq Q_N - S_N^{(2m-1)} \leq q_{2m-1}$ $\qquad 2m-1 \leq N$

(iii) $0 \leq S_N^{(2m)} - Q_N \leq q_{2m+1}$ $\qquad 2m+1 \leq N$

(iv) $0 \leq S_N^{(2m)} - Q_N \leq q_{2m}$ $\qquad 2m \leq N$

Theorem 2.4. For a point process $N(\cdot)$ with p.g.fl $G[\xi]$ whose m^{th} factorial moment measure is a Borel measure,

$$G[1-\varepsilon\xi] = 1-\varepsilon\int\xi(t)M(dt) + \frac{\varepsilon^2}{2!}\iint\xi(t_1)\xi(t_2)M_2(dt_1,dt_2)-\dots$$
$$+(-1)^m \frac{\varepsilon^m}{m!}\int\dots\int\xi(t_1)\dots\xi(t_m)M_m(dt_1,\dots,dt_m)+o(\varepsilon^m) \quad (2.15)$$

where $\xi \in V$ and $0 < \varepsilon < 1$.

Proof. Consider the function

$$\Gamma_m(\varepsilon,\xi) = \varepsilon^{-m}(-1)^{m+1}\{\prod_i[1-\varepsilon\xi(t_i)]-1+\varepsilon\int\xi(t)dN(t)-$$
$$-\frac{\varepsilon^2}{2!}\int_{t_1\neq t_2}\int \xi(t_1)\xi(t_2)dN(t_1)dN(t_2)$$
$$+\dots+(-1)^m\frac{\varepsilon^m}{m!}\int_{t_1\neq\dots\neq t_m}\dots\int \xi(t_1)\dots\xi(t_m)dN(t_1)\dots dN(t_m)\} .$$

For a realisation $\{t_i\}$ of $N(\cdot)$ we may rewrite the integrals as sums so that if $N = N(I)$, I the support of ξ,

$$\Gamma_m(\varepsilon,\xi) = \varepsilon^{-m}(-1)^{m+1}\{\prod_{i=1}^{N}(1-\varepsilon\xi_i)-1+\varepsilon q_1 - \frac{\varepsilon^2}{2!}q_2+..+(-1)^m\frac{\varepsilon^m}{m!}q_m\} \quad (2.16)$$

where $\xi_i = \xi(t_i)$ and the sums are zero for $m > N$.

Since $\xi \varepsilon V$, N is finite with probability one. Because the m^{th} factorial moment measure is a Borel measure we see from (2.16) and the Corollary to Lemma 2.4 that $\Gamma_m(\varepsilon,\xi)$ is positive, bounded by a random variable with finite expectation, namely $\frac{1}{m!}\int_{t_1\neq..\neq t_m}\!\!..\int \xi(t_1)..\xi(t_m)dN(t_1)..dN(t_m)$, and $\overset{a.s.}{\to} 0$ as $\varepsilon \to 0$. So by dominated convergence

$$E\{\Gamma_m(\varepsilon,\xi)\} \to 0 \quad \text{as} \quad \varepsilon \to 0$$

for $\xi \varepsilon V$, which proves the theorem. ▼

Corollary 1. If the $(m+1)$st moment measure is also a Borel measure, the error term $o(\varepsilon^m)$ is bounded by $\varepsilon^{m+1}\int..\int\xi(t_1)..\xi(t_{m+1})M_{m+1}(t_1,..,t_{m+1})$

This follows from (i) and (iii) of the Corollary to Lemma 2.4. It shows that there is a simple estimate for the remainder term in (2.15) if we assume the existence of higher-order moment measures.

Corollary 2. Under the conditions of the theorem

$$H[1-\varepsilon\xi] = -\{\varepsilon\int\xi(t)C_1(dt) + \frac{\varepsilon^2}{2!}\int\int\xi(t_1)\xi(t_2)C_2(dt_1,dt_2)+ .. + \frac{\varepsilon^m}{m!}\int..\int\xi(t_1)..\xi(t_m)C_m(dt_1,..,dt_m)\} + o(\varepsilon^m) .$$

Corollary 2 follows from the well-known expression $\log(1-x) = -\{x + \frac{x^2}{2} +..+ \frac{x^m}{m}\} + o(x^m)$. We remark that the remainder $o(\varepsilon^m)$ here is not estimable in the simple manner of Corollary 1. For consequences of this complication, see Chapter 4.

A result for characteristic functionals similar to Theorem 2.4 is given by Shiryaev [77], although he assumes $M_{m+1}(\cdot)$ to exist and has remainder $O(\varepsilon^{m+1})$. If we have $1-\xi \in L(N)$, rather than $\xi \in V$, the proof is still valid provided we also assume that $\int..\int \xi(t_1)..\xi(t_m)M_m(dt_1,..,dt_m) < \infty$, which holds automatically when $\xi \in V$ and $M_m(\cdot)$ is a Borel measure.

Equation (2.15) or a direct approach shows that the p.g.fl uniquely determines all existing factorial moment (and cumulant) measures. In fact there is a simple method of calculating them directly from the p.g.fl, given in the a.s. finite case by Moyal [61]. If we choose non-negative constants $x_1,..,x_k$ and functions $\xi_1,..,\xi_k$ suitably then a routine application of differentiation through an expectation operator gives

$$\frac{\partial^k}{\partial x_1..\partial x_k} G[1-x_1\xi_1-..-x_k\xi_k]\Big|_{(o,..,o)}$$

$$= E\Big\{ \sum_{i_1\neq..\neq i_k} .. \sum \xi_1(t_{i_1})..\xi_k(t_{i_k}) \prod_{j\neq i_1,..,i_k} [1-x_1\xi_1(t_j)-..-x_k\xi_k(t_j)]\Big\}\Big|_{(o,..o)}$$

$$= \int..\int \xi_1(t_1)..\xi_k(t_k)M_k(dt_1,..,dt_k),$$

and now put $\xi_i(t) = \chi_{A_i}(t)$ for Borel sets A_i. A suitable class of $\xi_1,..,\xi_k$ is that for which the integrals involved are all finite; as usual, this is true for $\xi_i \in V$ and $M_k(\cdot)$ a Borel measure.

The converse question is probably more interesting, namely when do the factorial moment measures uniquely determine the point process?

This is a random process analogue of the problem of moments (Feller [18] p.487, Leonov [47], [48]). In [48], Leonov has a comprehensive discussion of this topic. Our aims are more modest, and we state only the simplest of results. From (2.15) and the Corollary to Lemma 2.4.

$$G[1-x\xi] = 1-x\int\xi(t)M(dt)+..+(-1)^m \frac{x^m}{m!}\int..\int\xi(t_1)..\xi(t_m)M_m(dt_1,..,dt_m)$$
$$+\frac{x^m}{m!}R_m[\xi]$$

where $0 \leq R_m[\xi] \leq \int..\int\xi(t_1)..\xi(t_m)M_m(dt_1,..,dt_m)$, if all the integrals exist. This is certainly true for $\xi \in V$; then if the bounded interval I is the support of ξ

$R_m[\xi] \leq M_m(I \times .. \times I) = \mu_{[m]}(I)$, the m^{th} factorial moment of $N(I)$. So when all factorial moment measures exist and $\frac{1}{m!} x^m \mu_{[m]}(I) \to 0$ as $m \to \infty$, for some $x > 0$ and all I, the formal series (2.11) converges and G is uniquely determined by the coefficients $\int..\int\xi(t_1)..\xi(t_m)M_m(dt_1,..,dt_m)$ for all $\xi \in V$, hence by $\{M_m(\cdot)\}$.

Theorem 2.5. A point process $N(\cdot)$ is uniquely determined by its moment structure if for all bounded intervals I and some $x > 0$

$$\frac{x^m}{m!}\mu_{[m]}(I) \to 0 \quad \text{as} \quad m \to \infty.$$

A sufficient condition for this is $\limsup_{m\to\infty} \frac{1}{m}\mu_{[m]}^{1/m}(I) = \alpha(I) > 0$, all I (Feller [18]) p.487).

2.4. Characterisations of Ergodicity and Mixing

Leonov [47] has given, without proof, a series of theorems characterising ergodicity and mixing in a stationary process in terms of relations involving its characteristic functional. We will now derive analogous results for point processes and their p.g.fls, and give some examples. Similar ideas occur throughout general ergodic theory (e.g. Jacobs [31]).

The concepts of ergodicity and mixing for stationary point processes were introduced in 1.3, where we saw that both are equivalent to expressions involving limits of probabilities over measurable sets, i.e. members of F. We first consider ergodic point processes.

Theorem 2.6. A stationary point process $N(\cdot)$ with p.g.fl $G[\xi]$ is ergodic if and only if

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t G[\xi_1 S^\tau \xi_2]d\tau = G[\xi_1]G[\xi_2] \tag{2.17}$$

for $1-\xi_1$, $1-\xi_2 \varepsilon V$ or $\xi_1, \xi_2 \varepsilon L(N)$, where $S^\tau \xi(u) = \xi(u-\tau)$.

Proof. Suppose $N(\cdot)$ is ergodic. Then from (1.1)

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t P(A \cap T^{-\tau}B)d\tau = P(A)P(B) \tag{2.18}$$

for any $A, B \varepsilon F$ and so certainly for the cylinder sets $\{N : N(I_1) = n_1,..,N(I_k) = n_k\}$, $I_1,..,I_k$ any Borel sets and $k,n_1,..,n_k$ any integers. Thus

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t P\{N(I_1)=n_1,..,N(I_k)=n_k, N(J_1+\tau)=m_1,..,N(J_\ell+\tau)=m_\ell\}d\tau$$

$$= P\{N(I_1)=n_1,..,N(I_k)=n_k\}P\{N(J_1)=m_1,..,N(J_\ell)=m_\ell\}, \tag{2.19}$$

and it follows that the same relation holds for the corresponding p.g.fs, namely

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t P(I_1,\dots,I_k,J_1,\dots,J_\ell;\, y_1,\dots,y_k,z_1,\dots,z_\ell;\tau)d\tau$$

$$= P(I_1,\dots,I_k;y_1,\dots,y_k)P(J_1,\dots,J_\ell;z_1,\dots,z_\ell) \; . \tag{2.20}$$

Finally, since any measurable ξ can be uniformly approximated by an increasing sequence of simple functions, for which (2.17) holds by (2.20), we see that (2.17) is true for arbitrary ξ_1, ξ_2 in the appropriate classes.

Conversely, if (2.17) holds then by taking ξ_1, ξ_2 to be simple functions we deduce (2.20) and hence (2.19), since the L.H.S. of (2.20) is itself a p.g.f. (of the probability measure on the L.H.S. of (2.19)), and convergence of p.g.fs implies convergence of probabilities. So the ergodic relation (2.18) is established for the cylinder sets of F. To show that it holds for any measurable sets we use the following lemma, which is almost certainly known from general ergodic theory. As no specific statement of it has been found, we include a proof for completeness.

Lemma 2.5. Let R be a ring, with generated σ-ring $F(R)$ and an associated probability space (Ω, F, P). Let T be a measure-preserving transformation of Ω into itself. Then if

$$\lim_{t\to\infty} P(A \cap T^{-t}B) = P(A)P(B)$$

for any $A, B \in R$, it holds for any $A,B \in F(R)$.

Proof. From Halmos [26] p.56 we see that for any $\varepsilon > 0$ and each set E in $F(R)$ there is a set E_ε in R such that $P(E \,\Delta\, E_\varepsilon) < \varepsilon$.

Now given $\varepsilon > 0$ and arbitrary A, B in $F(R)$ with their 'approximating' sets A_ε, B_ε in R, consider

$$W = |P(A \cap T^{-t}B) - P(A)P(B)|$$

$$\leq |P(A \cap T^{-t}B) - P(A_\varepsilon \cap T^{-t}B_\varepsilon)| + |P(A_\varepsilon \cap T^{-t}B_\varepsilon) - P(A_\varepsilon)P(B_\varepsilon)|$$

$$+ |P(A_\varepsilon)P(B_\varepsilon) - P(A)P(B)|$$

$$= W_1(t,\varepsilon) + W_2(t,\varepsilon) + W_3(t,\varepsilon) \quad \text{say.}$$

And (i)

$$W_1(t,\varepsilon) \leq |P(A \cap T^{-t}B) - P(A_\varepsilon \cap T^{-t}B)| + |P(A_\varepsilon \cap T^{-t}B) - P(A_\varepsilon \cap T^{-t}B_\varepsilon)|,$$

while for arbitrary sets X, Y, Z

$$P(X \Delta Z) \geq P\{(X \Delta Z) \cap Y\} \geq P\{(X \cap Y) - (Z \cap Y)\}$$

$$\geq |P(X \cap Y) - P(Z \cap Y)|$$

so that

$$|P(A \cap T^{-t}B) - P(A_\varepsilon \cap T^{-t}B)| \leq P(A \Delta A_\varepsilon) < \varepsilon$$

$$|P(A_\varepsilon \cap T^{-t}B) - P(A_\varepsilon \cap T^{-t}B_\varepsilon) \leq P(T^{-t}B \Delta T^{-t}B_\varepsilon)$$

$$= P(B \Delta B_\varepsilon) < \varepsilon$$

because T is measure-preserving. Thus

$$W_1(t,\varepsilon) < 2\varepsilon \qquad \text{for all } t.$$

(ii) Since $A \cup A_\varepsilon = (A \Delta A_\varepsilon) \cup (A \cap A_\varepsilon)$,

$$P(A) \leq P(A \Delta A_\varepsilon) + P(A \cap A_\varepsilon)$$

$$\leq \varepsilon + P(A_\varepsilon)$$

and similarly for $P(B)$. Therefore

$$P(A)P(B) \leq \varepsilon^2 + 2\varepsilon + P(A_\varepsilon)P(B_\varepsilon)$$

and by symmetry $W_3(t,\varepsilon) < 3\varepsilon$ for all t.

(iii) $W_2(t,\varepsilon) < \varepsilon$ for t sufficiently large, by hypothesis. From (i), (ii), (iii) $W < 6\varepsilon$ if t sufficiently large. As this is true for arbitrary $\varepsilon > 0$ and any $A, B \in F(R)$ the lemma is proved. ▼

Now the cylinder sets in F form a semi-ring whose generated ring is the finite union of disjoint cylinders. (Halmos [26] p.26). Equation (2.18) obviously holds for the generated ring and, as an easy consequence of Lemma 2.5, it holds also for the σ-ring generated by the semi-ring, that is for all measurable sets. This proves the converse proposition and hence the theorem. ▼

An exactly similar argument proves

Theorem 2.7. A stationary point process $N(\cdot)$ with p.g.fl $G[\xi]$ is mixing if and only if

$$\lim_{t\to\infty} G[\xi_1 S^t \xi_2] = G[\xi_1]G[\xi_2] \tag{2.21}$$

for $1-\xi_1$, $1-\xi_2 \in V$ or $\xi_1, \xi_2 \in L(N)$.

Obviously, a like result is true for weakly mixing processes.

We now look for examples of ergodic and mixing point processes. Such processes are of interest in a variety of applications as the mixing condition (1.2), or (2.21), is a form of asymptotic independence of the numbers of points in widely separated intervals. Such a condition is often what seems required to prove limit theorems for point processes; similar considerations for mixing sequences of random variables have been intensively studied in recent years (e.g. Ibragimov [30]).

Since the stationary Poisson process has independent increments we expect that it is mixing (and hence ergodic cf. 1.3). This is easily proved on substituting (2.8) in (2.21). Likewise, processes related to the Poisson process should be mixing, and we prove this for doubly stochastic Poisson processes.

Theorem 2.8. A doubly stochastic Poisson process with stationary mean process $\Lambda(t)$ is mixing if Λ is mixing.

Proof. We need a slight modification of Leonov's [47] characterisation of mixing stochastic processes. Since $\Lambda(t)$ is non-negative and stationary its Laplace functional is well-defined, and from [47] Theorem 1 we easily deduce that Λ is mixing if and only if

$$\lim_{t\to\infty} L_\Lambda[\xi_1+S^t\xi_2] = L_\Lambda[\xi_1]\, L_\Lambda[\xi_2] \tag{2.22}$$

for all totally finite measures ξ_1, ξ_2.

From (2.10) we can transform the mixing condition (2.21) into an expression involving the Laplace functional of Λ. We write $\overline{L}[\xi]$ for $L[\int_{-\infty}^{t} \xi(u)du]$. Obviously

$$G[\xi_1 S^t\xi_2] = \overline{L}_\Lambda[1-\xi_1 S^t\xi_2]$$

$$G[\xi_1]G[\xi_2] = \overline{L}_\Lambda[1-\xi_1]\overline{L}_\Lambda[1-\xi_2]$$

for $1-\xi_1$, $1-\xi_2 \in V$, so to establish (2.21) we must show

$$\lim_{t\to\infty} \overline{L}_\Lambda[1-\xi_1 S^t\xi_2] = \lim_{t\to\infty} \overline{L}_\Lambda[1-\xi_1+1-S^t\xi_2] \tag{2.23}$$

Put $\eta_i = 1-\xi_i$, so that $\eta_i \in V$ $i = 1,2$. Then

$$1-\xi_1 S^t\xi_2 = 1-\xi_1+\xi_1(1-S^t\xi_2) = \eta_1+\xi_1 S^t\eta_2$$

and

$$
\begin{aligned}
0 &\le \bar{L}_\Lambda[1-\xi_1 S^t\xi_2] - \bar{L}_\Lambda[\eta_1+S^t\eta_2] \\
&= \bar{L}_\Lambda[\eta_1+\xi_1 S^t\eta_2] - \bar{L}_\Lambda[\eta_1+S^t\eta_2] \\
&= E\{\exp\{-\int\lambda(u)[\eta_1(u)+\xi_1(u)\eta_2(u-t)]du\}. \\
&\qquad\qquad (1-\exp\{-\int\lambda(u)\eta_1(u)\eta_2(u-t)du\})\} \\
&\le E\{1- \exp[-\int\lambda(u)\eta_1(u)\eta_2(u-t)du\} \\
&\le E\{\int\lambda(u)\eta_1(u)\eta_2(u-t)du\} \\
&= E\{\Lambda(u)\}\int\eta_1(u)\eta_2(u-t)du
\end{aligned}
$$

since Λ is stationary with finite mean. But the η_i are integrable, so the last expression tends to zero as $t \to \infty$. This proves (2.23), and the theorem follows from (2.22). ▼

A corresponding result can be proved if Λ is only ergodic. Because linear processes are all mixing (Rosenblatt [73] p.112) we have

<u>Corollary</u>. A linear stochastic Poisson process is mixing, and hence ergodic.

From the ergodic version of Theorem 2.8 it is easy to see that a mixed Poisson process is ergodic if and only if the mixing distribution is concentrated at one point, an observation which is clear independently. (The corollary does not apply because obviously a process with a.s. constant realisations is neither mixing nor ergodic).

Kersten and Matthes [37] give a number of results on mixing and ergodicity of infinitely divisible point processes. For related ideas, see section 3.4.

3. CLUSTER PROCESSES

3.1. Introduction

Cluster processes have developed coincidentally with point processes both because of their considerable practical application and because they are one of the few classes of point processes for which a reasonable theory can be derived. They have their origin in work on contagion mechanisms in ecology (Thompson [84]), but have since been applied to such diverse questions as the distribution of galaxies in space (Neyman and Scott [66], [67]), the flow of motor traffic (Bartlett [4]), computer failures (Lewis [49], [50]) and the occurrence of earthquakes (Vere-Jones [88]). So both one-dimensional and multi-dimensional cluster processes are of interest, though as usual we consider only the former. Important theoretical contributions are due to Moyal, who first pointed out the applicability of the p.g.fl to cluster processes in the discussion to [67] and subsequently developed a comprehensive theory in [61], [62].

In this chapter we define the general cluster process, indicate some simple properties and introduce useful special cases. We study the existence of such processes, in a reasonable sense, and also their mixing and asymptotic properties. Finally, we derive some results for two cluster models used in practice.

3.2. Definition, Properties and Examples of Cluster Processes

A cluster process $N(\cdot)$ has two components, the process of cluster centres $N_c(\cdot)$ and the process of cluster members $N_s(\cdot)$. Each point of the cluster centre process is assumed to initiate a cluster member process, generally called a *cluster*, independently for each point. The cluster process consists of the superposition of all the clusters, and possibly also the cluster centres. In general we will assume these to be *excluded*.

If the cluster centre process has p.g.fl $G_1[\xi]$ and the p.g.fl of a cluster, given its centre is at t, is $G_2[\xi|t]$, then the independent development of clusters shows that conditional on a realisation $\{t_i\}$ of $N_c(\cdot)$ the p.g.fl of the entire process is $\Pi_i\ G_2[\xi|t_i]$. Thus the p.g.fl of a cluster process is given by the fundamental relation

$$G[\xi] = G_1[G_2[\xi|t]], \tag{3.1}$$

a result due to Moyal (discussion to [67]). This compact formulation of a cluster process will be very useful in the study of a wide range of its properties.

Equation (3.1) has been further investigated by Moyal [61], [62], regarding it as the first 'generation' of a generalised branching process (see also Harris [28] chapter 3), and Vere-Jones [87], [88] specifically in our context. They derive, in particular, elegant relations for the factorial moment and cumulant measures of $N(\cdot)$ in terms of those for $N_c(\cdot)$ and $N_s(\cdot)$, though these do not concern us. In [4] there is work on the spectral analysis and interval properties of cluster processes.

We next give some examples of particular cluster models, introduced by various authors, and derive their p.g.fls from (3.1) where possible.

(a) Compound Poisson process (section 1.5). Here all the cluster members occur simultaneously at the cluster centre. If $\{p_n\}$ is the distribution of cluster size, with p.g.f. $P(z)$, then

$$G[\xi] = \exp\{-\int(1-P[\xi(t)])\Lambda(dt)\} \tag{3.2}$$

(b) $G/G/\infty$ queue. Each cluster centre (arrival) produces exactly one cluster member (departure) after a random time (service time). In other words, the cluster process is the output stream. If the service time distribution is $F(x)$, then Vere-Jones [87] shows

$$G[\xi] = G_1[\int_0^\infty \xi(t+x)\ dF(x)] \tag{3.3}$$

(c) Neyman-Scott model (Neyman and Scott [66], [67]). A Poisson (λ) process of cluster centres triggers clusters whose members are independently and identically distributed about the cluster centre with distribution $F(x)$. If $P(z)$ is the p.g.f. of the cluster size, assumed a.s. finite, then

$$G[\xi] = \exp\{-\lambda\int(1-P[\int\xi(t+x)\ dF(x)])\ dt\}, \tag{3.4}$$

as shown by Moyal (see [88]).

(d) Bartlett-Lewis model (Bartlett [4], Lewis [49]). Again $N_c(\cdot)$ is a Poisson (λ) process, but now the cluster members form a finite renewal process following the cluster centre. The p.g.fl cannot be written in closed form here, except as a series of multiple integrals cf. (2.3).

(e) Earthquake model (Vere-Jones [87]). $N_s(\cdot)$ is a non-homogeneous Poisson $(\mu(t))$ process, $t \geq 0$ and $\int_0^\infty \mu(t)dt < \infty$. Then

$$G_2[\xi|t] = \exp\{-\int_t^\infty \mu(x-t)[1-\xi(x)]dx\},$$

so that with a Poisson (λ) centre process

$$G[\xi] = \exp\{-\lambda\int(1 - \exp\{-\int_0^\infty \mu(x)[1-\xi(x+t)]dx\})dt\} \tag{3.5}$$

Models (a), (c), (d), (e), and (b) for $M/G/\infty$, are special cases of the general Poisson cluster process where the cluster centres form a Poisson process and the clusters are arbitrary. Such processes are studied in Matthes et. al. [36], [37], [54] and Goldman [23] in the context of infinitely divisible point processes, as it turns out that they are exactly equivalent to the regular infinitely divisible processes (section 1.5). Our methods enable us to derive and extend their results for these processes in a simple manner.

Among cluster processes we find several examples of distinct but equivalent formulations of point processes. One such case is noted in Bartlett [5] and Vere-Jones [88]; they show that with suitable choice of the parameters the following processes are identical (i.e. have the same p.g.fl):

Vere-Jones' earthquake model,

A doubly stochastic Poisson process with 'shot-noise' mean,

A Neyman-Scott cluster process with Poisson cluster size.

We extend these relations in the following theorems.

Theorem 3.1. A linear stochastic Poisson process is equivalent to a Neyman-Scott process.

Proof. From the linear stochastic Poisson process (1.5 and 5.2), define

$$\ell = \int f(x)dx < \infty$$

$$\lambda = -\psi(\ell);\ \lambda > 0 \quad \text{as clearly} \quad \ell > 0$$

$$\zeta_i = \delta_{1k} \frac{\gamma\ell}{\lambda} + \frac{\ell^k}{k!\lambda} \int_0^\infty x^{k-1} e^{-x\ell} dK(x) \qquad k = 1,2,\ldots$$

where δ_{ij} is the Kronecker delta. Clearly $\zeta_k \geq 0$ for all k. It is finite because K is proportional to a distribution function, so has a Laplace transform with all derivatives finite on $(0, \infty)$. And

$$\sum_{k=1}^{\infty} \zeta_k = \frac{\gamma\ell}{\lambda} + \int_0^\infty \frac{(e^{x\ell}-1)e^{-x\ell}}{\lambda x} dK(x) = 1$$

by definition of λ; we can interchange summation and integral because all terms are positive.

So $\{\zeta_k\}$, $k = 1,2,\ldots$ is a discrete probability distribution. Its p.g.f. is

$$Z(w) = \sum_{k=1}^{\infty} w^k \zeta_k \qquad w \varepsilon [0,1]$$

$$= \frac{\gamma\ell w}{\lambda} + \int_0^\infty \frac{(e^{x\ell w}-1)e^{-x\ell}}{\lambda x} dK(x)$$

as before.

Now for $1 - \xi \varepsilon V$, let $w' = \ell^{-1}\int\xi(v+t)f(v)dv$. For all t, $w' \varepsilon [0,1]$ and so we may substitute in $Z(w)$. After some simplification

$$Z(\int\xi(v+t)\ell^{-1}f(v)dv) = 1 + \lambda^{-1} \psi[\int[1-\xi(v+t)]f(v)dv]$$

Therefore

$$-\lambda\int[1-Z(\int\xi(v+t)\ell^{-1}f(v)dv]dt = \int\psi[\int[1-\xi(v+t)]f(v)dv]dt, \quad (3.6)$$

which is the log p.g.fl of a linear stochastic Poisson process (equation (5.4)). So the L.H.S. of (3.6) exists. But from example (c), the L.H.S. of (3.6) is the log p.g.fl of a Neyman-Scott process for which $F'(x) = \ell^{-1}f(x)$ and $Z(w)$ is the cluster size p.g.f. Hence the two processes are identical. ▼

Remarks. 1. In the case of a 'shot-noise' mean, i.e. $X(\cdot)$ a Poisson process in (1.11), this result is due to Bartlett [5].

2. If we define $\zeta_o = \alpha$, $0 \leq \alpha < 1$, and use $\frac{\lambda}{1-\alpha}$ in the previous definition of ζ_k, $k \geq 1$, the resultant process is unchanged. This reinforces the observation in [88] that the probability of no cluster members is unidentifiable.

Theorem 3.2. A doubly stochastic Poisson process with mean

$$Y(t) = \int f(t-u)dN(u), \quad (3.7)$$

where $f \geq 0$ is integrable and $N(\cdot)$ is a stationary point process, is equivalent to a cluster process with $N_c(\cdot) \stackrel{d}{=} N(\cdot)$ and an inhomogeneous Poisson $(f(t))$ process for the clusters.

Proof. From (3.7) $Y(t)$ is stationary, non-negative and has finite mean, so that $Y(t)$ exists with probability one. Its Laplace functional is

$$\begin{aligned} L_Y[\xi] &= E\{e^{-\int Y(t)\xi(dt)}\} \\ &= E\{e^{-\int\int f(t-u)\xi(dt)dN(u)}\} \\ &= G_1[e^{-\int f(t-u)\xi(dt)}], \end{aligned}$$

where $G_1[\xi]$ is the p.g.fl of $N(\cdot)$. Clearly

$$\int\left|\log e^{-\int f(t-u)\xi(dt)}\right|du = \int f(x)dx\int\xi(dt) < \infty$$

as f is integrable and $\xi(\cdot)$ totally finite, so by Lemma 2.1 the p.g.fl exists.

On taking $Y(t)$ for the mean process of a doubly stochastic Poisson process, (2.10) shows that its p.g.fl is given by

$$G[\xi] = G_1[e^{-\int f(t-u)[1-\xi(t)]dt}] \tag{3.8}$$

Comparison with example (e) shows that the argument of G_1 in (3.8) is the p.g.fl of an inhomogeneous Poisson $(f(t))$ process. The result now follows from (3.1). ▼

Remarks 1. When $N(\cdot)$ is a Poisson process, this is the result of Vere-Jones [88]. He also gives an illuminating heuristic derivation which carries over directly to this case.

2. Putting $\lambda = \int f(x)dx$, $g(x) = \lambda^{-1}f(x)$, (3.8) shows that the clusters are also of Neyman-Scott type with cluster size Poisson (λ) distributed and $F'(x) = g(x)$.

3.3. Existence Criteria for Cluster Processes

So far we have not considered whether our cluster processes exist, in the sense that with probability one there are only a finite number of points in a finite interval. In some cases, such as population processes, no problems arise as all component processes of the cluster process are taken to be a.s. finite so that the entire process must be a.s. finite. However the examples of 3.2 show that in general we can have an a.s. infinite set of cluster centres. This means that any interval may contain cluster members initiated by infinitely many centres, and it is apparent that some restriction on the cluster structure will be needed to ensure existence. Our results give such restrictions.

Matthes [54] (see also Goldman [23]) has stated a necessary and sufficient condition for the existence of Poisson cluster processes, and Neyman [65] has done work on the Neyman-Scott model. We derive a necessary and sufficient condition for existence of a wide class of cluster processes and deduce some useful sufficient conditions, giving also an extremely simple proof of Matthes' criterion.

The cluster process exists if for all bounded intervals I, $N(I) < \infty$ with probability one. In (3.1) put $\xi(u) = 1-(1-z)\chi_I(u)$, $0 \le z \le 1$. This reduces (3.1) to a p.g.f. so that existence is equivalent to

$$P_I(z) = G[G_2[1-(1-z)\chi_I|t]] \uparrow 1 \qquad \text{as} \quad z \uparrow 1 \tag{3.9}$$

for all bounded intervals I.

Write

$$\begin{aligned} Q_I(z;t) &= -\log G_2[1-(1-z)\chi_I|t] \\ &= -\log P_I(z;t), \end{aligned}$$

so that $Q_I(\cdot)$ is the logarithm of the p.g.f. of $N_s(I)$ if the cluster is centred at t, and

$$P_I(z) = E\{\exp[-\int Q_I(z;t)dN_c(t)]\} \tag{3.10}$$

Now $Q_I(z;t) \downarrow 0$ monotonically as $z \uparrow 1$, so (3.9) follows from the monotone convergence theorem if with probability one

$$\int Q_I(z;t)dN_c(t) \tag{3.11}$$

is finite for some z, $0 < z < 1$. If we assume that the clusters are a.s. finite, with cluster size p.g.f. $P(z)$, then

$$P_I(z;t) \geq P(z) \quad \text{for all } I,\ t,\ z \in [0,1) \tag{3.12}$$

For $\mathcal{P}\{N_s(I) > n/ \text{ cluster centre at } t\} \leq \mathcal{P}\{N_s(-\infty,\infty) > n\}$ and consequently

$$\frac{1-P_I(z;t)}{1-z} \leq \frac{1-P(z)}{1-z} ,$$

which implies (3.12). So $P_I(z;t)$ is bounded away from zero as t varies and (3.11) will converge or diverge with

$$\int[1 - P_I(z;t)]dN_c(t) . \tag{3.13}$$

Define $\rho_I(n;t) = \mathcal{P}\{N_s(I) > n/ \text{ cluster centre at } t\}$, with generating function

$$R_I(z;t) = \frac{1-P_I(z;t)}{1-z} .$$

So (3.13) becomes

$$(1-z)\int R_I(z;t)dN_c(t) = (1-z) \sum_{n=0}^{\infty} z^n \int \rho_I(n;t)dN_c(t)$$

by the positivity of the summands, which is

$$\leq \int \rho_I(0;t)dN_c(t)$$

because the $\{\rho_I(n;t)\}$ form a monotone decreasing sequence. Therefore $\int \rho_I(0;t)dN_c(t)$ finite with probability one is a necessary and sufficient condition for the finiteness of (3.11), and a sufficient condition for the existence of the cluster process.

We now prove it is also necessary. Suppose that $\int \rho_I(0;t)dN_c(t)$ is infinite on a set $\mathcal{D} \subset F_c$, $P(\mathcal{D}) = p > 0$. Then as $(1-z)\int R_I(z;t)dN_c(t)$ is both dominated by, and contains as a summand, $\int \rho_I(0;t)dN_c(t)$ it is infinite precisely for $N_c(\cdot)$ in $\mathcal{D}$ independently of $z \in [0,1)$. So both of (3.11), (3.13) are infinite precisely for $N_c(\cdot)$ in $\mathcal{D}$, and hence by (3.10)

$$
\begin{aligned}
P_I(z) &= \left(\int_{\mathcal{D}} + \int_{\Omega-\mathcal{D}}\right)\exp\{-\int Q_I(z;t)dN_c(t)\}P(dN_c) \\
&\leq P(\Omega - \mathcal{D}) = 1 - p \qquad\qquad z \in [0,1)
\end{aligned}
$$

If then $P_I(z) \uparrow 1$ as $z \uparrow 1$ we can choose $z' \in [0,1)$ such that $P_I(z') > 1-p$, a contradiction. So (3.9) cannot hold in this case, which proves the necessity and consequently

Theorem 3.3. A cluster process with a.s. finite clusters exists if and only if for every bounded interval I

$$\int \rho_I(0;t)dN_c(t) < \infty \text{ with probability one} \qquad (3.14)$$

Neyman [65] shows that some such theorem is required by constructing a Neyman-Scott process with an a.s. infinite number of points in a bounded interval.

It will be noticed that we have not specified the mode of convergence of the stochastic integral in (3.14). Since the integral is actually a sum of non-negative random variables it will always have an a.s. limit, and moreover under the condition (3.14) all modes of convergence are equivalent.

Unfortunately, (3.14) is still in a stochastic form and is therefore of limited value in applications. No non-stochastic if and only if results have as yet been proved in general, but there are several useful sufficient conditions deducible from (3.14).

Corollary 1. If $\int\rho_I(0;t)M_c(dt) < \infty$ for all bounded intervals I then (3.14) is satisfied. If $N_c(\cdot)$ is stationary this becomes $\int\rho_I(0;t)dt < \infty$ for *some* bounded interval I.

Corollary 2. If the cluster structure depends only on the position relative to the cluster centre and $N_c(\cdot)$ is stationary, (3.14) is satisfied when the mean cluster size is finite.

Remark. With the cluster structure of Corollary 2, and $N_c(\cdot)$ stationary, the cluster process itself is stationary (Vere-Jones [88]). In future, any stationary cluster process will be assumed to have these properties.

Corollary 2 follows from

$$\rho_I(0;t) = 1 - P_I(0;t) = P_I'(z_0;t) \qquad 0 < z_0 < 1$$

by the mean value theorem, whence $\rho_I(0;t) \leq P_I'(1;t)$, the mean number of events in I from a cluster centred at t. With our cluster structure this is $E\{\Sigma_1^\nu\chi_I(x_i+t)\}$, where ν is the a.s. finite random variable $N_s(-\infty,\infty)$ and the x_i are the times of occurrence of cluster members. So if $|I|$ is the Lebesgue measure of I

$$\int\rho_I(0,t)dt \leq E\{\nu|I|\} < \infty$$

and apply Corollary 1 to complete the proof.

Corollary 2 was derived by Neyman [65] in the special case of a Neyman-Scott process, and in general by Vere-Jones (unpublished). One can produce examples to show that $E(\nu) < \infty$ is not a necessary condition (see [65]), and also that it is not implied by $\int \rho_I(0;t)dt < \infty$.

For a stationary Poisson cluster process we may obtain more satisfying results. Equation (3.9) becomes

$$P_I(z) = \exp\{-\lambda \int [1-P_I(z;t)]dt\}$$

and so will $\uparrow 1$ with z if and only if $\int [1-P_I(z;t)]dt$ is convergent for all $z \in [0,1)$. Arguing as in the proof of Theorem 3.3 we see this process exists if and only if

$$\int \rho_I(0;t)dt < \infty \quad \text{for some bounded interval } I, \tag{3.15}$$

which is a very simple proof of the Matthes-Goldman theorem (cf. [54], [23], where the Borel-Cantelli lemmas are used repeatedly).

We can apply these results to some of the special processes discussed earlier. In each case, assume I is bounded.

In the compound process,

$$\begin{aligned} \rho_I(0;t) &= 1 \quad \text{if the cluster centre falls in } I, \\ &= 0 \quad \text{otherwise.} \end{aligned}$$

So (3.14) equals $N_c(I)$, which is always a.s. finite, and we conclude that such processes always exist.

For a $G/G/\infty$ queue we see that $\rho_I(0;t) = F\{I-t\}$ so that from Corollary 1 a sufficient condition for existence is $\int F\{I-t\}M_c(dt) < \infty$. But this is just the first moment measure of the output stream (Vere-Jones [87]), so the result is obvious a priori.

A discussion of the Bartlett-Lewis model is deferred until 3.5.

Since the Neyman-Scott model is a special Poisson cluster process we have (3.15) as a necessary and sufficient condition for its existence.

From (3.4)

$$\rho_I(0;t) = 1 - P[1-F\{I-t\}].$$

The process being stationary, take $I = [0,x)$ without loss of generality. Then (3.15) becomes, for any $x > 0$,

$$\int\{1-P[1-F(x+t)+F(t)]\}dt < \infty \tag{3.16}$$

We now remark

(i) by Corollary 2 a finite mean cluster size is sufficient for (3.16), without any restrictions on F,

(ii) if F has finite range then the integrand in (3.16) vanishes for large t and the integral must be finite (Neyman [65]).

These are both sufficient conditions. As yet no simple necessary and sufficient conditions have come from (3.16); however we do generalise (i) by assuming something about F. Write

$$\Theta(t) = -\log[1-F(x+t)+F(t)]$$

$$P(z) = \int_0^\infty z^u dB(u) = \sum_{n=0}^\infty b_n z^n \qquad 0 \le z \le 1,$$

so that $b_n = P\{\nu = n\}$ and $B(\cdot)$ is its distribution function. Then (3.16) becomes

$$\int_0^\infty \int_{-\infty}^\infty \{1-e^{-u\Theta(t)}\}dt\ dB(u) < \infty, \tag{3.17}$$

where $\Theta(t) \ge 0$, $\Theta(t) \to 0$ as $t \to \pm\infty$ and $\Theta(t)$ integrable over $(-\infty,\infty)$.

(a) Suppose F is the exponential distribution with parameter one. Then (3.17) becomes

$$\sum_{n=1}^\infty b_n \int_0^\infty \{1-(1-ce^{-t})^n\}dt < \infty$$

where $c = 1-e^{-x} > 0$. Changing the variable by $y = 1-ce^{-t}$ we see that (3.17) is true if and only if

$$\sum_{n=1}^{\infty} b_n \int_0^1 \left(\frac{1-y^n}{1-y}\right) dy = \sum_{n=1}^{\infty} b_n \sum_{j=1}^{n} \frac{1}{j} < \infty,$$

and the well-known result $\Sigma_{j=1}^{n} \frac{1}{j} \sim \log n$ as $n \to \infty$ shows us that (3.17) holds if and only if $E(\log \nu) < \infty$. Clearly this is unaffected by F having exponential tails with an arbitrary parameter.

(b) Suppose that F has regularly varying tails with exponent $-\alpha$, $\alpha > 0$ (see Feller [18] p.268) i.e. as $s \to \infty$

$$1 - F(s) \sim s^{-\alpha} L(s),$$

where $L(s) \geq 0$ is of slow variation. Then as $t \to \infty$

$$\begin{aligned} F(x+t) - F(t) &\sim t^{-\alpha} L(t)\left\{1 - \frac{L(x+t)}{L(t)}\left(1 + \frac{x}{t}\right)^{-\alpha}\right\} \\ &\sim t^{-\alpha} L(t)\left\{1 - \frac{L(x+t)}{L(t)} + \frac{L(x+t)}{L(t)} \cdot \frac{x\alpha}{t} + o(t^{-1})\right\} \\ &\sim ct^{-(\alpha+1)} L(t) \qquad\qquad \alpha > 0 \end{aligned}$$

where c is a constant. A similar result holds for the lower tail. Since the values of u, t around the origin do not affect convergence in (3.17), we see that (3.17) converges if and only if

$$\int_C^{\infty} \int_{C'}^{\infty} (1 - \exp\{-ut^{-(\alpha+1)} L(t)\}) dt \, dB(u) < \infty, \qquad \alpha > 0,$$

and similarly for negative t.

Change the variable by $y = ut^{-(\alpha+1)}$. As $L(s^{\beta}) \sim L(s)$ we require the convergence of

$$\int_C^{\infty} \int_o^{u/C''} \frac{1-\exp\{-yL(\frac{u}{y})\}}{y^{\frac{1}{\alpha+1}+1}} u^{\frac{1}{\alpha+1}} dy \, dB(u).$$

Write $\int_0^{u/C''}$ as $\int_0^{\delta} + \int_{\delta}^{u/C''}$ for some $\delta > 0$. Clearly the second integral will converge if $\int_1^{\infty} u^{1/(\alpha+1)}\, dB(u) < \infty$. For the first integral, take

$$\int_C^{\infty} u^{1/(\alpha+1)} L(u)\{\int_0^{\delta} \frac{L(\frac{u}{y})}{L(u)}\, y^{-\frac{1}{\alpha+1}} dy\} dB(u), \tag{3.18}$$

which dominates it. The canonical form for a slowly varying function (Feller [18] p.274) shows that for any $\gamma > 0$ $\frac{L(u/y)}{L(u)} < y^{-\gamma}$ for u sufficiently large. Choosing C so that $\gamma + \frac{1}{\alpha+1} < 1$, the inner integral in (3.18) converges for all $u > C$ and (3.18) converges if $\int_1^{\infty} u^{1/(\alpha+1)}\, L(u)dB(u) < \infty$.

We collect these results as

Theorem 3.4. Suppose a Neyman-Scott cluster process has an F with regularly varying tails with exponent $-\alpha$, $\alpha > 0$. Then
(1) a sufficient condition for its existence is $E\{\max[\nu^{\frac{1}{1+\alpha}}, \nu^{\frac{1}{1+\alpha}} L(\nu)]\}$ finite,
(2) if $L(s)$ is bounded, $E(\nu^{\frac{1}{1+\alpha}}) < \infty$ is sufficient for existence, and if $L(s) \to$ constant as $s \to \infty$ the condition is necessary and sufficient.

If the tails of F are exponential, the process exists if and only if $E(\log \nu) < \infty$.

3.4. Mixing in Cluster Processes

The general mixing characterisation given in Theorem 2.7, when combined with (3.1), enables us to prove statements about the mixing (and ergodic) properties of cluster processes. We consider only mixing here, though as before all the theorems carry over to weak mixing and ergodicity with obvious changes.

Suppose we have a stationary cluster process with p.g.fl (3.1). If it exists, Theorem 3.3 implies that for $1 - \xi \in V$

$$\int(1 - G_2[\xi|t])dN_c(t) < \infty \quad \text{a.s.} \tag{3.19}$$

The principal result is

Theorem 3.5. A stationary cluster process is mixing if it exists and the cluster centre process is mixing.

Proof. Since $N_c(\cdot)$ is mixing, application of Theorem 2.7 for $\xi \in L(N)$ gives

$$\lim_{\tau\to\infty} G_1[G_2[\xi_1|t]S^\tau G_2[\xi_2|t]] = G_1[G_2[\xi_1|t]].G_1[G_2[\xi_2|t]] \tag{3.20}$$

for $1 - \xi_i \in V$.

The R.H.S. of (3.20) equals $G[\xi_1]G[\xi_2]$. To establish mixing we must prove that $\lim_{\tau\to\infty} G[\xi_1 S^\tau \xi_2] = G[\xi_1]G[\xi_2]$, so the problem becomes to prove that

$$\lim_{\tau\to\infty} G_1[G_2[\xi_1 S^\tau \xi_2|t]] = \lim_{\tau\to\infty} G_1[G_2[\xi_1|t]S^\tau G_2[\xi_2|t]] \tag{3.21}$$

for $1 - \xi_i \in V$, where the R.H.S. limit exists from (3.20).

Let $\Delta_\tau(t) = G_2[\xi_1 S^\tau \xi_2|t] - G_2[\xi_1|t]S^\tau G_2[\xi_2|t]$. We now prove

Lemma 3.1. $\int \Delta_\tau(t)dN_c(t) \overset{a.s.}{\to} 0$ as $\tau \to \infty$

Proof. Define random variables

$$X(t) = \exp \int \log \xi_1(u-t)dN_s(u), \quad Y_\tau(t) = \exp \int \log \xi_2(u-t-\tau)dN_s(u).$$

Then clearly

$$\Delta_\tau(t) = E\{X(t)Y_\tau(t)\} - E\{X(t)\}E\{Y_\tau(t)\}$$
$$= E\{X(t) - E[X(t)]\}\{Y_\tau(t) - E[Y_\tau(t)]\}.$$

Let I, J be the supports of the functions $1-\xi_1$, $1-\xi_2 \in V$, so they are both bounded intervals. Since the clusters $N_s(\cdot)$ are finite with probability 1 we can choose τ so large that, given $\varepsilon > 0$, the probability of one or more events at a distance $> \frac{1}{2}\tau$ from the cluster centre is at most $\frac{1}{4}\varepsilon$. But at least one of $I+t$, $J+t+\tau$ is at a distance $> \frac{1}{2}\tau$ from the cluster centre for all t, so that

$$\min[P\{N_s(I+t) \neq 0\},\ P\{N_s(J+t+\tau) \neq 0\}] < \tfrac{1}{4}\varepsilon \tag{3.22}$$

Then because $|a-b| \leq 1$ if $0 \leq a,b \leq 1$ we have

$$|\Delta_\tau(t)| \leq E\{|X(t) - E[X(t)]|.|Y_\tau(t) - E[Y_\tau(t)]\}$$
$$\leq \{1 - E[X(t)]\}.\{1 - E[Y_\tau(t)]\}P\{N_s(I+t) = 0, N_s(J+t+\tau) = 0\}$$
$$+ \{1 - E[X(t)]\}\ P\{N_s(I+t) = 0,\ N_s(J+t+\tau) \neq 0\}$$
$$+ \{1 - E[Y_\tau(t)]\}\ P\{N_s(I+t) \neq 0,\ N_s(J+t+\tau) = 0\}$$
$$+ P\{N_s(I+t) \neq 0, N_s(J+t+\tau) \neq 0\}\ .$$

But $1 - E[X(t)] = E[1-X(t)] \leq P\{N_s(I+t) \neq 0\}$, and similarly for $1 - E[Y_\tau(t)]$, so that from (3.22)

$$|\Delta_\tau(t)| \leq 4 \min\ [P\{N_s(I+t) \neq 0\},\ P\{N_s(J+t+\tau) \neq 0\}] < \varepsilon \tag{3.23}$$

uniformly in t, for large τ.

Also $|\Delta_\tau(t)| \leq E|X(t) - E[X(t)]|$

$$\leq E\{1 - X(t)\} + E\{1 - E[X(t)]\}$$

$$= 2\{1 - E[X(t)]\}$$

$$\text{i.e.} \quad |\Delta_\tau(t)| \le 2(1 - G_2[\xi_1|t]) \qquad 1 - \xi_1 \in V. \tag{3.24}$$

The lemma is now a consequence of (3.19), (3.23), (3.24) and the dominated convergence theorem. ▼

To complete the proof of the theorem, we use the identity

$$\prod_{i=1}^{n} \zeta_1(t_i) - \prod_{i=1}^{n} \zeta_2(t_i) = \sum_{i=1}^{n} [\zeta_1(t_i)-\zeta_2(t_i)]\zeta_2(t_1)\ldots\zeta_2(t_{i-1})\zeta_1(t_{i+1})\ldots \ldots \zeta_1(t_n), \tag{3.25}$$

valid for $n = 1,2,\ldots$ and any functions ζ_1, ζ_2.
Put $\zeta_1^{(\tau)}(t) = G_2[\xi_1 S^\tau \xi_2|t]$, $\zeta_2^{(\tau)}(t) = G_2[\xi_1|t]S^\tau G_2[\xi_2|t]$,
so that $\Delta_\tau(t) = \zeta_1^{(\tau)}(t) - \zeta_2^{(\tau)}(t)$. The modulus of the L.H.S. of (3.25) is bounded by 1, while that of the R.H.S. is bounded by $\int|\Delta_\tau(t)|dN_c(t)$, if we take $\{t_i\}$ as a realisation of the cluster centre process. Taking expectations over $\{t_i\}$ in (3.25), Lemma 3.1 and dominated convergence prove (3.21) and hence the theorem. ▼

From 2.4 we have the immediate

Corollary. The Poisson cluster process, and therefore the regular infinitely divisible point processes, are all mixing.

This was originally proved by Kersten and Matthes [37]. They deduced it from a necessary and sufficient condition for mixing in stationary infinitely divisible point processes, which is easily established by p.g.fl methods as follows.

In 4.3 we will show that the log p.g.fl of an infinitely divisible point process $N(\cdot)$ is

$$H[\xi] = \int_{\Omega-\{\emptyset\}} \{e^{\int \log\xi(t)dN(t)} - 1\} \tilde{P}(dN).$$

Taking logarithms in (2.21) we see that mixing means

$$H[\xi_1 S^\tau \xi_2] - H[\xi_1] - H[\xi_2] \to 0 \quad \text{as} \quad \tau \to \infty, \quad 1 - \xi_i \in V.$$

The L.H.S. becomes, in the stationary case,

$$\int_{\Omega-\{\emptyset\}} \{1-\exp \int \log \xi_1(t)dN(t) - \exp \int \log \xi_2(t-\tau)dN(t)$$

$$+ \exp \int \log \xi_1(t)\xi_2(t-\tau)dN(t)\} \tilde{P}(dN)$$

$$= \int_{\Omega-\{\emptyset\}} \{1 - \exp \int \log \xi_1(t)dN(t)\}\{1 - \exp \int \log \xi_2(t-\tau)dN(t)\}\tilde{P}(dN) \tag{3.26}$$

Suppose that I, J are the supports of $1-\xi_1$, $1-\xi_2$ respectively. Then the integrand is zero for the events $\{N(I) = 0\}$ and $\{N(J+\tau) = 0\}$, so they may be removed from the range of integration. Otherwise, (3.26) is non-negative and

$$\leq \tilde{P} \{\Omega-(\{N(I) = 0\} \cup \{N(J+\tau) = 0\})\}$$

$$= \tilde{P}\{N(I) \neq 0, N(J+\tau) \neq 0\}$$

Therefore if this converges to zero as $\tau \to \infty$ for any bounded I, J, so does (3.26) and the process is mixing. Conversely, if the process is mixing then by choosing ξ_1, ξ_2 to be zero over I, J respectively we see from (3.26) that $\tilde{P}\{N(I) \neq 0, N(J+\tau) \neq 0\} \to 0$ as $\tau \to \infty$.

This proves

Theorem 3.6 (Kersten and Matthes [37]). A stationary infinitely divisible point process $N(\cdot)$ is mixing if and only if for all bounded intervals I, J

$$\tilde{P}\{N(I) \neq 0, N(J+\tau) \neq 0\} \to 0 \quad \text{as} \quad \tau \to \infty.$$

Several of their other results for stationary infinitely divisible point processes (e.g. that weak mixing is equivalent to ergodicity) can be proved just as simply in similar fashion.

3.5 Asymptotic Properties of Cluster Processes

We now study the asymptotic properties of a generalised form of the Bartlett-Lewis model (example (d) of 3.2), in which we assume only that the cluster members all follow the cluster centre with a distribution depending solely on the distance from it. These results are of two types; the behaviour as $T \to \infty$ of $N[0,T)$ when the process starts at the origin, and work on the stationary equilibrium distribution. The latter includes an existence criterion for the stationary Bartlett-Lewis process, deferred from 3.3, and a direct proof of the form of its equilibrium p.g.f.

In part, this generalises Lewis [51]. On occasions we take the cluster structure of the Bartlett-Lewis model for more exact results (this is particularly true in the equilibrium situation), but a surprising amount can be achieved in the more general framework.

From (3.1) the p.g.fl of our process is

$$G[\xi] = \exp\{-\lambda \int (1-G_2[\xi|t])dt\} \tag{3.27}$$

Set $\xi(t) = 1-(1-z)\chi_{[0,T)}(t)$, so that as usual the p.g.fl becomes the p.g.f. $P_T(z)$. Assume now that the cluster centre process began at $t = 0$, so that it is no longer stationary. If $P(z;u)$ is the p.g.f. of $N_s[0,u)$, with cluster centre at zero, then (3.27) becomes

$$T^{-1} \log P_T(z) = -\lambda(1 - T^{-1} \int_0^T P(z;u)du); \qquad (3.28)$$

another derivation including the cluster centres is in Lewis [49]. We first study the behaviour of the R.H.S. as $T \to \infty$.

Let $p_i = P(\nu=i)$, so that we may write $P(z;\infty) = \Sigma_{i=0}^{\infty} p_i z^i$. Define the quantities $p_i(u)$ by $P(z;u) = \Sigma_{i=0}^{\infty} p_i(u) z^i$ and the dist^n fn of the time to the j^{th} event in the cluster as $F_j(u)$, with $\overline{F}_j(u) = 1 - F_j(u)$ and $R_i = \Sigma_{j=i}^{\infty} p_j$. Then

$$P\{N_s[0,u) \geq i, \text{ cluster centre at zero}\} = R_i \, F_i(u),$$

so by the usual p.g.f. formulae

$$\frac{1-P(z;\infty)}{1-z} - \frac{1-P(z;u)}{1-z} = \sum_{i=1}^{\infty} z^{i-1}\{R_i - R_i F_i(u)\}$$

and

$$P(z;u) - P(z;\infty) = (1-z) \sum_{i=1}^{\infty} z^{i-1} R_i \overline{F}_i(u) \qquad (3.29)$$

In view of (3.28), consider $\lim_{T\to\infty} T^{-1} \int_0^T \{P(z;u)-P(z;\infty)\}du$. From (3.29) this equals

$$\lim_{T\to\infty} (1-z) \sum_{i=1}^{\infty} z^{i-1} R_i \, T^{-1} \int_0^T \overline{F}_i(u)du,$$

the positivity of the summands validating the interchange. Because $\overline{F}_i(u)$ is monotone decreasing in u for all i, $T^{-1} \int_0^T \overline{F}_i(u) \to 0$ as $T \to \infty$, and the summands are dominated by $(1-z)z^{i-1}$ which form a convergent series. Thus the limit in question is zero, and from (3.28)

$$T^{-1} \log P_T(z) \to -\lambda\{1-P(z;\infty)\} \qquad \text{as} \quad T \to \infty \tag{3.30}$$

This means that the distribution of $N[0,T)$ is asymptotically of a compound Poisson form, as one might expect heuristically from the nature of the process (see Vere-Jones [88]).

We obtain more exact statements as follows. From (3.29) we conclude, as before, that

$$\int_0^T \{P(z;u) - P(z;\infty)\}du = (1-z) \sum_{i=1}^{\infty} z^{i-1} R_i \int_0^T \overline{F}_i(u)du \tag{3.31}$$

and

$$\lim_{T\to\infty} \int_0^T \{P(z;u) - P(z;\infty)\}du = (1-z) \sum_{i=1}^{\infty} z^{i-1} R_i \, \mu_i , \tag{3.32}$$

where μ_i is the expected time to the i^{th} cluster member, though both sides of (3.32) may be infinite. When they are finite, this generalises various moment relations of Lewis [49], [51] and Smith (discussion to [49]), as follows on differentiating (3.32) at $z = 1$.

<u>Theorem 3.7</u> If $\Sigma_{i=1}^{\infty} R_i \, \mu_i < \infty$ then as $T \to \infty$

$$\log P_T(z) = -\lambda T\{1-P(z;\infty)\} + \lambda\int_0^{\infty} \{P(z;u)-P(z;\infty)\}du + o(1) \tag{3.33}$$

<u>Corollary</u> For a Bartlett-Lewis process $\mu_i = i\mu$ and (3.33) holds if $E(\nu^2) < \infty$, $\mu < \infty$.

The theorem comes directly from (3.32), the corollary from $\Sigma_{i=1}^{\infty} R_i \mu_i = \mu \, \Sigma_{i=1}^{\infty} iR_i = \frac{1}{2}\mu\{E(\nu^2) + E(\nu)\}$. We note that (3.30) shows the cumulants of a Poisson cluster process are asymptotically linear, and, as remarked above, we may obtain further terms in an asymptotic expansion from (3.33), under suitable conditions.

As a further application, consider the asymptotic distribution of the random variable

$$\frac{N[0,T) - E\{N[0,T)\}}{\sqrt{\lambda TE(\nu^2)}} .$$

Its characteristic function has logarithm $\psi_T(\theta)$ given by

$$\psi_T(\theta) = -\lambda T\{1-P(\theta\sigma^{-1};\infty)\} + \lambda\int_0^T \{P(\theta\sigma^{-1};u)-P(\theta\sigma^{-1};\infty)\}du$$
$$- i\theta\sigma^{-1} E\{N[0,T)\},$$

where $\sigma = \sqrt{\lambda TE(\nu^2)}$ and $P(\theta;\cdot) \equiv P(e^{i\theta};\cdot)$. From (3.28)

$$E\{N[0,T)\} = \lambda \sum_{j=1}^{\infty} R_j \int_0^T F_j(u)du$$
$$= \lambda TE(\nu) - \lambda \sum_{j=1}^{\infty} R_j \int_0^T \overline{F}_j(u)du,$$

so we have, by (3.31),

$$\psi_T(\theta) = \lambda T\{P(\theta\sigma^{-1};\infty) - 1 - i\theta\sigma^{-1}E(\nu)\}$$
$$+ \lambda\{(1-e^{i\theta/\sigma}) \sum_{j=1}^{\infty} e^{i(j-1)\theta/\sigma}R_j\int_0^T \overline{F}_j(u)du+i\theta\sigma^{-1} \sum_{j=1}^{\infty} R_j\int_0^T \overline{F}_j(u)du\}$$
$$= \Gamma_T(\theta) + \Delta_T(\theta) \qquad \text{say} \tag{3.34}$$

Now $\Gamma_T(\theta)$ is the second characteristic function of a standardised compound Poisson distribution and so if $E(\nu^2) < \infty$

$$\Gamma_T(\theta) \to -\frac{1}{2}\theta^2 \quad \text{as} \quad T \to \infty \tag{3.35}$$

And

$$\Delta_T(\theta) = \lambda(1-e^{i\theta/\sigma}+i\theta\sigma^{-1}) \sum_{j=1}^{\infty} R_j \int_0^T \overline{F}_j(u)du$$
$$+ \lambda(1-e^{i\theta/\sigma}) \sum_{j=1}^{\infty} (e^{i(j-1)\theta/\sigma}-1)R_j \int_0^T \overline{F}_j(u)du,$$

so that elementary inequalities (Feller [18] p.485) give

$$|\Delta_T(\theta)| \le \tfrac{1}{2}\lambda\theta^2\sigma^{-2} \sum_{j=1}^{\infty} R_j \int_0^T \overline{F}_j(u)\,du$$

$$+ \lambda\theta\sigma^{-1} \sum_{j=1}^{\infty} (j-1)\theta\sigma^{-1} R_j \int_0^T \overline{F}_j(u)\,du$$

$$\le \frac{\theta^2}{TE(\nu^2)} \sum_{j=1}^{\infty} jR_j \int_0^T \overline{F}_j(u)\,du$$

if $E(\nu^2) < \infty$. But in that case $\Sigma_{j=1}^{\infty} jR_j < \infty$ (cf. the Corollary to Theorem 3.7), and $T^{-1} \int_0^T \overline{F}_j(u)du \to 0$ as $T \to \infty$, so by dominated convergence

$$|\Delta_T(\theta)| \to 0 \quad \text{as} \quad T \to \infty \tag{3.36}$$

Then from (3.34), (3.35), (3.36), if $E(\nu^2) < \infty$,

$$\psi_T(\theta) \to -\tfrac{1}{2}\theta^2 \quad \text{as} \quad T \to \infty$$

which proves

Theorem 3.8 If $E(\nu^2) < \infty$, the standardised random variable

$$\frac{N[0,T) - E\{N[0,T)\}}{\sqrt{\lambda TE(\nu^2)}}$$

is asymptotically normally distributed.

This is an improvement of a theorem of Lewis [51], who requires extra assumptions. For the Bartlett-Lewis model, his theorems are a direct consequence of our result.

We now consider the process which begins at a time x after the initiation of the cluster process. If this is well-defined in the limit as $x \to \infty$ it is called the equilibrium process. From (3.27) it is obvious that an equivalent situation is the cluster process beginning at $-x$ while we consider only events after time zero, so that the equilibrium process can be taken to start in the remote past.

With this interpretation we can readily find existence conditions for the process and also the form of the equilibrium p.g.f. in the particular case of a Bartlett-Lewis model. Lewis [49] derived the transform of this p.g.f. by a limiting argument but could not prove it directly; such a proof was given by Franken and Richter [21]. It relied heavily on the infinite divisibility of the process and the attendant structure developed in an earlier paper [20]. We will give a direct proof without such machinery.

From (3.27) it is clear that the equilibrium process is a stationary Poisson cluster process and so the necessary and sufficient condition for its existence is given by (3.15). To calculate the equilibrium p.g.f. $P_T(z)$ for $N[0,T)$ we let $P(z;a,b)$ be the p.g.f. of $N_s[a,b)$ with its centre at zero; if $a = 0$ we write $P(z;b)$ as before. Then from (3.27)

$$\log P_T(z) = -\lambda\{\int_0^T [1-P(z;t)]dt + \int_0^\infty [1-P(z;t,t+T)]dt\} \qquad (3.37)$$

Note that this implies that the existence criterion (3.15) can be slightly modified to $\int_0^\infty \rho_{[0,T)}(0;t)dt < \infty$, as the first integral in (3.37) does not affect matters.

Now with our general cluster structure

$$\int_0^T [1-P(z;t)]dt = (1-z)\sum_{k=1}^\infty z^{k-1}R_k \int_0^T F_k(y)dy, \qquad (3.38)$$

by identical reasoning to (3.31). However for the second term in (3.37) we will assume the Bartlett-Lewis cluster structure, so that F_k is the k-fold convolution of the interval distribution F and the intervals between cluster members are independent. Then

$$\int_0^\infty [1-P(z;t,t+T)]dt = (1-z)\sum_{k=1}^\infty z^{k-1}\int_0^\infty P\{N_s[t,t+T) \geq k,\text{ centre at zero}\}dt \tag{3.39}$$

and using the independence

$$P\{N_s(t,t+T) \geq k,\text{ centre at zero}\} =$$
$$= \sum_{i=o}^\infty R_{i+k}\int_o^t \int_{t-u}^{T+t-u} F_{k-1}(T+t-v-u)dF(v)dF_i(u) \tag{3.40}$$

It is easily shown that

$$\int_o^\infty \int_o^t \int_{t-u}^{T+t-u} F_{k-1}(T+t-v-u)dF(v)dF_i(u)dt = \int_o^\infty \int_x^{T+x} F_{k-1}(T+x-v)dF(v)dx$$
$$= \int_o^\infty \{F_k(T+x)-F_k(x)\}dx + \int_o^\infty \{F_{k-1}(y)-F_{k-1}(T+y)\}dy \tag{3.41}$$

But if $B(\cdot)$ is an arbitrary distribution function on $[0,\infty)$,

$$\int_o^\infty \{B(x+a)-B(x)\} = \int_o^a \{1-B(y)\}dy \qquad a > 0, \tag{3.42}$$

so finally, from equations (3.39)-(3.42),

$$\int_o^\infty [1-P(z;t,t+T)]dt = (1-z)\sum_{k=1}^\infty \sum_{i=k}^\infty R_i z^{k-1}\int_o^T \{F_{k-1}(y)-F_k(y)\}dy \tag{3.43}$$

Together with (3.38) this gives an explicit result for $P_T(z)$. After some simplification, if $E(\nu) < \infty$,

$$\log P_T(z) = -\lambda(1-z)\{T.E(\nu)+(z-1)\sum_{k=2}^\infty \sum_{i=k}^\infty R_i z^{k-2}\int_o^T F_{k-1}(y)dy\} \tag{3.44}$$

This differs slightly from the Franken-Richter result [21], because we do not include the cluster centres in a cluster process. If they are included, (3.27) becomes

$$G[\xi] = \exp\{-\lambda\int(1-\xi(t)G_2[\xi|t])dt\} \tag{3.27a}$$

and so

$$\log P_T(z) = -\lambda\{\int_0^T [1-zP(z;t)]dt + \int_0^\infty [1-P(z;t,t+T)]dt\} \qquad (3.37a)$$

Equation (3.43) is still relevant and (3.38) becomes

$$\int_0^T [1-zP(z;t)]dt = (1-z)\{T+\sum_{k=1}^\infty R_k z^k \int_0^T F_k(t)dt\}. \qquad (3.38a)$$

Then (3.37a), (3.38a) and (3.43) give, after simplification,

$$\log P_T(z) = -\lambda(1-z)\{T[1+E(\nu)]+(z-1)\sum_{k=1}^\infty \sum_{i=k}^\infty R_i z^{k-1} \int_0^T F_k(y)dy\} \qquad (3.44a)$$

which is equation (1) of [21].

We collect these results in

<u>Theorem 3.9</u> The equilibrium Poisson cluster process exists if and only if $\int_0^\infty \rho_{[0,T)}(0;t)dt < \infty$. For a Bartlett-Lewis process this is equivalent to $E(\nu) < \infty$, in which case the equilibrium p.g.f. is given by (3.44) or (3.44a) and the limiting forward recurrence time W has distribution

$$\log P(W \geq T) = \log P_T(0) = -\lambda\{T+E(\nu)\int_0^T [1-F(y)]dy\}$$

<u>Proof</u>. The only unverified statement is the existence criterion for a Bartlett-Lewis process. From (3.43) we see that

$$\int_0^\infty \rho_{[0,T)}(0;t)dt = \int_0^\infty [1-P(0;t,t+T)]dt$$

$$= \sum_{i=1}^\infty R_i \int_0^T [1-F(y)]dy,$$

which is finite if and only if $\Sigma_{i=1}^\infty R_i = E(\nu) < \infty$. ▼

This theorem shows that the condition $E(\nu) < \infty$ is incorporated into the equilibrium Bartlett-Lewis cluster process and so may be dispensed with in statements of results. Also, the equations (3.44), (3.44a) are very useful in deducing moment relations for the equilibrium process and may be used to prove a theorem, analogous to Theorem 3.8, giving asymptotic normality for this process.

3.6 Infinite-server Queueing Systems†

Infinite-server queues, with their alternative interpretation as a randomly delayed stream of events, have been widely studied in different contexts. The classical immigration-death process (Moran [60] p.176) is an $M/M/\infty$ queue. Smith [79] has used the $GI/G/\infty$ system to illustrate infinite products occurring in renewal theory, and as we see from (2.2) these are closely tied to p.g.fls. In the discussion to [79], Skellam mentions the $M/G/\infty$ queue as a tractable system; see also Lewis [49] for applications to the number of operative clusters at a given time. Lewis [52] and Nelsen and Williams [64] consider random delays to a deterministic schedule ($D/G/\infty$). Recently, Rao [72] has worked with infinite-server queueing situations arising in textile research.

An important quantity in such systems is $n(t)$, the number of servers busy (or the number of customers in service) at time t. We find the Laplace functional of this process $n(\cdot)$ for a $G/G/\infty$ queue, which enables us to generalise many earlier results.

The p.g.fl of the output is given by (3.3). To calculate the Laplace functional $L_n[\xi]$ of $n(t)$, consider an input $\{t_i\}$, with associated service times $\{x_i\}$ where the x_i are independently and

† These are not strictly queues, since no customer has to wait, but it is a convenient terminology.

identically distributed as F, independent of the t_i. Then

$$n(t) = \sum_{t_i \le t} \delta_i \qquad \delta_i = \begin{cases} 1 & t_i + x_i > t \\ 0 & t_i + x_i \le t \end{cases}$$

and

$$\begin{aligned} L_n[\xi] &= E\{e^{-\int n(t)\xi(dt)}\}, \quad \xi(\cdot) \text{ a totally finite measure} \\ &= E_{\{t_i\}}\{E_{\{x_i\}}(E\{e^{-\int n(t)\xi(dt)} \mid t_i, x_i\})\} \\ &= E\{E(\exp[-\Sigma\xi(t_i, t_i + x_i)] \mid t_i)\} \\ &= E\{\prod_i \int_0^\infty \exp[-\xi(t_i, t_i + x)]dF(x)\} \end{aligned}$$

by the independence of the x_i. This proves

<u>Theorem 3.10</u> The Laplace functional of $n(\cdot)$ for a $G/G/\infty$ queue is given by

$$L_n[\xi] = G_1[\int_0^\infty \exp\{-\xi(t, t+x)\}dF(x)] \tag{3.45}$$

If the input is stationary then so is $n(\cdot)$.

The stationarity assertion generalises Rao [72], Theorem 3. The comprehensive nature of the Laplace functional now permits us to deduce many properties of $n(\cdot)$.

<u>Corollary 1.</u> $E\{z^{n(a)}\} = G_1[1-(1-z)(1-F(a-t))] \qquad 0 \le z \le 1.$

Thus $E\{n(a)\} = \int_{-\infty}^{a} [1-F(a-t)]M_1(dt)$, where $M_1(\cdot)$ is the expectation measure of the input. For a stationary input

$$E\{n(a)\} = m \int_0^\infty [1-F(t)]dt = mE(S),$$

where S is the service time.

For the $M/G/\infty$ queue starting at time zero

$$E\{z^{n(a)}\} = \exp\{-\lambda(1-z)\int_0^a [1-F(t)]dt\},$$

a non-homogeneous Poisson distribution cf. Skellam in [79].

The second-order properties of $n(\cdot)$ also follow readily from (3.45). For instance

$$E\{e^{-\theta_1 n(a)} e^{-\theta_2 n(a+b)}\} = G_1[\int_0^\infty \exp\{-\phi(t,x,a,b,\theta_1,\theta_2)\}dF(x)] \quad \theta_i > 0$$

where

$$\begin{aligned}\phi(t,x,a,b,\theta_1,\theta_2) &= 0 \quad \text{if} \quad a \geq t+x \quad \text{or} \quad a \leq t \leq t+x \leq a+b \\ &= \theta_1 \quad \text{if} \quad t \leq a \leq t+x \leq a+b \\ &= \theta_2 \quad \text{if} \quad a \leq t \leq a+b \leq t+x \\ &= \theta_1+\theta_2 \text{ if } t \leq a \leq a+b \leq t+x\end{aligned}$$

and so

$$E\{z_1^{n(a)} \; z_2^{n(a+b)}\} = G_1[\psi(z_1,z_2,a,b;t)] \tag{3.46}$$

where

$$\begin{aligned}\psi(z_1,z_2,a,b;t) &= F(a-t)+z_1\{F(a+b-t)-F(a-t)\}+z_1z_2\{1-F(a+b-t)\} \quad t \leq a \\ &= F(a+b-t)+z_2\{1-F(a+b-t)\} \quad t>a\end{aligned} \tag{3.47}$$

<u>Corollary 2</u>. The joint distribution of $n(a)$, $n(a+b)$ is given by (3.46), (3.47). For a stationary input

$$\begin{aligned}\psi(z_1,z_2,b;t) &= (1-z_1)F(t-b)+z_1F(t)+z_1z_2\{1-F(t)\} \qquad t \geq b \\ &= F(t) + z_2\{1-F(t)\} \qquad 0 < t < b\end{aligned}$$

For similar results with a GI stationary input, see Rao [72], Theorem 6. From Corollary 2, $\mathrm{Cov}\{n(a),n(a+b)\}$ can be calculated. The special case of $M/G/\infty$ leads to more elegant expressions which we do not state explicitly - for details see [72].

Corollary 3. The Laplace transform of the traffic time average $\frac{1}{T}\int_0^T n(a)\,da$ comes from setting

$$\xi(da) = T^{-1}\theta\,da \qquad 0 \le a \le T$$
$$= 0 \qquad \text{otherwise} \quad .$$

in (3.45). The final result is not stated, though it is easily calculated just as before.

We remark that this approach can also lead to joint generating functionals for such quantities as $n(t)$ and the output, thus combining the work of this section and Vere-Jones [87].

4. LIMIT THEOREMS FOR POINT PROCESSES

4.1. Introduction

In recent years the sustained interest in limit theorems for sequences of independent variates has extended to dependent situations, particularly random processes. The field of point processes provides a unique variety of limiting operations which basically involve the addition, deletion or translation of points, with appropriate changes of scale. In all cases we consider, the limit process, in the sense of 1.3, is a Poisson process, which explains in part its ubiquitous occurrence as a model.

The first limiting result was for the superposition of a large number of, in some sense, negligible streams, a point process version of the Central Limit Problem (Khintchine [39] §16). Many rediscoveries and generalisations have followed e.g. Ososkov [68], Grigelionis [25], Goldman [22]. The limit for the converse notion of randomly deleting events of an arbitrary stream was given, under general conditions, by Belyaev [9] (for higher dimensions see [22]). Our third operation, randomly translating the points of some initial stream, was first studied by Dobrushin [16], with subsequent extensions from Goldman [22] and Stone [80]. This idea also arises in work on road traffic flows, for which see Breiman [12] and Thedéen [82], [83].

Because the characteristic function is so useful in proving limit theorems for random variables it was hoped, in view of Lemma 2.3, that the p.g.fl would be of equal value in the study of limit theorems for point processes. The difficulties which arise in trying to implement this idea are worth discussing. Basically, we wish to use the asymptotic expansion of the p.g.fl in Theorem 2.4, with $m = 1$, imposing conditions sufficient to make the error negligible (cf. the classical Central Limit Theorem, Feller [18] p. 488), so we require reasonable bounds or approximations to the error term. The problem is that this error, written $o(\varepsilon^m)$ in (2.15), is not uniform in ξ. For superposition theorems, when we can use $G[\xi]$ directly, this is not so serious since Corollary 1 to Theorem 2.4 provides error bounds if we assume the existence of higher order moment measures. However the other limiting operations are more naturally associated with $H[\xi]$, and here no workable bounds have been obtained. Consequently, a unified approach to point process limit theorems via the p.g.fl has not eventuated as yet, though we do give some results established by a technique of Vere-Jones [87].

It is because most limit theorems assert convergence to a Poisson process that we prefer the p.g.fl to other potential functionals in attempting to develop a general convergence theory. As the Poisson process has $C_n(\cdot) \equiv 0$ if $n \geq 2$, it seems natural to use the logarithm of the p.g.fl with which the $C_n(\cdot)$ are associated, introducing such conditions as ensure they are negligible in the limit. Asymptotic independence of counts in distant intervals, i.e. a mixing condition, appears to be appropriate, although little is known about this in the present context.

In this chapter we first look at some special ergodic results which are related to the usual Laws of Large Numbers. We next establish the canonical form of the p.g.fl of an infinitely divisible point process and prove some related limit theorems on superpositions of random streams. Finally we investigate the limit theorems arising from deletion and translation of point processes.

4.2 Ergodic Theorems for Point Processes

Since there is a wide literature on ergodic theory for stationary stochastic processes we try to relate ergodic theorems for stationary point processes $N(\cdot)$ to this work. There are two such relations, corresponding to the discrete and continuous parameter ergodic theory; we will only consider the latter.

If $X_b(t) = N[t-b,t)$ then for $2b < t$

$$\int_b^t X_b(u)du = \int_b^t \int_{u-b}^u dN(v)du = \int_o^b v dN(v) + b\int_b^{t-b} dN(v) + \int_{t-b}^t (t-v)dN(v)$$

$$\leq bN[0,t)$$

and similarly $\int_o^{t+b} X_b(u)du \geq bN[0,t)$,

i.e.

$$\frac{1}{bt}\int_o^t X_b(u)du \leq \frac{N[0,t)}{t} \leq \frac{1}{bt}\int_o^{t+b} X_b(u)du \tag{4.1}$$

So the behaviour as $t \to \infty$ of $t^{-1}N[0,t)$ is identical with that of $t^{-1}\int_o^t X_b(u)du$, where $X_b(\cdot)$ is a stationary stochastic process. This idea is used by Goldman [22] and Beutler and Leneman [10] in similar circumstances.

We deduce from (4.1) and the ergodic theorem for stationary process

(Doob [17] p.515) that as $t \to \infty$ $t^{-1}N[0,t)$ converges a.s. to a random variable invariant under translation (cf. Beutler and Leneman [10]). When this limit is a.s. constant we say that the Strong Law of Large Numbers holds for $N(\cdot)$, and that the Law of Large Numbers holds when there is mean square convergence to a constant.

In future we consider only weakly stationary $N(\cdot)$, when $X_b(t)$ is also weakly stationary. Let $C_b(t) = \text{Cov}\{X_b(\tau), X_b(t+\tau)\}$. Then from the identity (1.7)

$$C_b(t) = \frac{1}{2}\{D(t+b) + D(t-b) - 2D(t)\}, \tag{4.2}$$

and on substituting the spectral representation (1.8) for $D(t)$

$$C_b(t) = b^2\mu(\{0\}) + 2\int_{(0,\infty)} \cos\theta t\left(\frac{1-\cos\theta b}{\theta^2}\right)\mu(d\theta) \tag{4.3}$$

Since $C_b(t)$ is the covariance function of a real stationary process it must be the cosine transform of a spectral measure $F_b(\cdot)$ ([17], p.519). Here we can actually express $F_b(\cdot)$ in terms of the spectral measure for $D(t)$ by

$$F_b(d\theta) = \frac{1-\cos\theta b}{\theta^2}\mu(d\theta) \tag{4.4}$$

We are interested in $C_b(t)$ because Doob [17] p.530 and Verbitskaya [86] have shown how intimately its behaviour is connected with the validity of the two Laws of Large Numbers. Firstly, we look for mean square convergence of $t^{-1}N[0,t)$.

From (4.1), $t^{-1}N[0,t) \to m = E\{N[0,1)\}$ in mean square if and only if this occurs for $(bt)^{-1}\int_0^t X_b(u)du$. But by Doob [17] p.530 this is equivalent to $t^{-1}\int_0^t C_b(\tau)d\tau \to 0$ as $t \to \infty$ i.e. $F_b(\cdot)$ has no atom at the origin. Then from (4.4) we deduce

<u>Theorem 4.1</u> The Law of Large Numbers holds for a weakly stationary point process $N(\cdot)$ if and only if $\mu(\{0\}) = 0$. A sufficient condition for this is that $N(\cdot)$ have an integrable covariance density.

Verbitskaya [86] has stated sufficient conditions for the Strong Law of Large Numbers to hold for $X_b(t)$ in terms of its covariance function, related to its behaviour as $t \to \infty$. We shall express this in terms of properties of $\mu(\cdot)$ in one simple case.

Suppose $C_b(t) = O(t^{-\alpha})$, $\alpha > 0$, as $t \to \infty$. Then clearly $\overline{C}_b(t) = t^{-1}\int_0^t C_b(u)du = O(t^{-\alpha})$ as $t \to \infty$ which from [86] implies the Strong Law for $X_b(t)$, and hence $N(\cdot)$. As $C_b(t)$ is the cosine transform of a totally finite measure $F_b(\cdot)$ if $\mu(\{0\}) = 0$, it is well known (Lukacs [53] p.27) that its behaviour as $t \to \infty$ may be irregular unless $F_b(\cdot)$ is absolutely continuous with respect to Lebesgue measure. Assume $\mu(\cdot)$ has no singular component; then from (4.4)

<u>Lemma 4.1</u> $F_b(\cdot)$ is absolutely continuous with respect to Lebesgue measure if and only if the atoms of $\mu(\cdot)$ are concentrated on $\{2\pi n b^{-1};\ n = 1,2,\ldots\}$, i.e. on a lattice with span $2\pi b^{-1}$.

Under these conditions, if $h(\theta)$ is the density of the absolutely continuous part of $\mu(\cdot)$ (4.3) becomes

$$C_b(t) = 2\int_0^\infty \cos\,\theta t\left(\frac{1-\cos\,\theta b}{\theta^2}\right) h(\theta)d\theta \tag{4.5}$$

We can now prove

<u>Theorem 4.2</u> Let $N(\cdot)$ be a weakly stationary point process whose spectral measure has no singular component, atoms concentrated on

a lattice, and whose density with respect to Lebesgue measure is of bounded variation in (δ, ∞) for each $\delta > 0$ and satisfies $h(\theta) \sim c\theta^{-\gamma}$, $0 < \gamma < 1$, as $\theta \to 0$. Then the Strong Law of Large Numbers holds for $N(\cdot)$.

Proof. We know the result is true if $C_b(t) = O(t^{-\alpha})$, $\alpha > 0$, as $t \to \infty$. Since $h(\theta)$ is of bounded variation in (δ, ∞) so is $f_b(\theta) = \frac{1-\cos\theta b}{\theta^2} \cdot h(\theta)$. But also $f_b(\theta) = \theta^{-\gamma}\{\frac{1-\cos\theta b}{\theta^2} \cdot \theta^{\gamma}h(\theta)\}$, where the term in brackets is bounded, of bounded variation on $(0, \infty)$, and tends to a non-zero finite constant as $\theta \to 0$. So the conditions of Titchmarsh [85], Theorem 126, are satisfied and consequently

$$C_b(t) = O(t^{\gamma-1}) \qquad \text{as} \quad t \to \infty, \quad 0 < \gamma < 1. \qquad \blacktriangledown$$

4.3 Infinitely Divisible Point Processes

In 1.5 we introduced the notion of an infinitely divisible point process. In our further investigations we first derive its p.g.fl in terms of the KLM measure $\tilde{P}$, as a minor extension of Kersten and Matthes [36] and Lee [45].

Consider the functional

$$G[\xi] = \exp\{\int_{\Omega-\{\emptyset\}} [e^{\int \log \xi(t)dN(t)} - 1]\tilde{P}(dN)\} \tag{4.6}$$

for $1 - \xi \in V$. For simple functions ξ we see from Theorem 1.3 that $G[\xi]$ becomes the p.g.f. of a compound Poisson distribution, which is obviously a proper p.g.f. Now suppose we have a sequence of functions $1 - \xi_n \in V$ all vanishing outside a common interval I and that $\xi_n \to \xi$ pointwise as $n \to \infty$. By (3) of 1.5, $\tilde{P}\{N(I) \neq 0\} < \infty$ and as always $N(I)$ is a.s. finite. Since the integrand in (4.6) vanishes for all $N(\cdot)$ in Ω with $N(I) = 0$ it is now a simple consequence of dominated convergence that $G[\xi_n] \to G[\xi]$ as $n \to \infty$. So by Theorem 2.2 $G[\xi]$ is a p.g.fl, and obviously it is the p.g.fl of the infinitely divisible point process. This proves

Theorem 4.3 The p.g.fl of an infinitely divisible point process with KLM measure $\tilde{P}$ is given by (4.6). Conversely, if $\tilde{P}$ has the properties of a KLM measure, (4.6) is the p.g.fl of an infinitely divisible point process.

This canonical form for an infinitely divisible p.g.fl has a variety of applications. In cases where the p.g.fl is expressible in closed form, (4.6) gives the associated KLM measure directly (e.g. Lee [46]). It also helps us deduce some characterisation results analogous to those of Lukacs [53] p.83.

Corollary 1. A point process is infinitely divisible if and only if its p.g.fl can be expressed as

$$G[\xi] = \lim_{n\to\infty} \exp\{\lambda_n(G_n[\xi]-1)\} \tag{4.7}$$

where the $G_n[\xi]$ are all p.g.fls and the λ_n are positive real numbers.

This means that a point process is infinitely divisible if and only if it is the limit in distribution of a sequence of singular Poisson processes E_{λ_n, N_n}.

Corollary 2. An infinitely divisible point process has the form $E_{\lambda,N}$ if and only if $\tilde{P}$ is totally finite. Then $\lambda = \tilde{P}(\Omega)$ and $N(\cdot)$ has probability measure $\lambda^{-1}\tilde{P}$.

Both these results are implicitly stated in [36]. We note that the 'curiousity' of Lee [46] §4 has a totally finite KLM measure and so must be a singular Poisson process. It is of course a special case of the doubly stochastic Poisson process, and (2.10) shows that doubly stochastic processes are infinitely divisible if their mean processes have this property. In particular, linear stochastic Poisson processes are all infinitely divisible (section 5.3).

By definition (A) of 1.5, the limit in distribution of the superposition of independent u.a.n. point processes $\{N_{n,i}(\cdot)\}$ $i = 1,\dots s_n$ $n = 1,2,\dots$ is an infinitely divisible point process. To exploit this we express the u.a.n. condition in terms of p.g.fls.

<u>Lemma 4.2</u> The sequence $\{N_{n,i}(\cdot)\}$ is u.a.n. if and only if for given $\varepsilon > 0$ and $1 - \xi \in V$ there is $n_o(\varepsilon)$ such that if $n > n_o$,

$$1 - G_{n,i}[\xi] < \varepsilon \quad \text{uniformly in } i.$$

<u>Proof</u>. Suppose the $N_{n,i}(\cdot)$ are u.a.n. with p.g.fls $G_{n,i}[\xi]$. Given $\varepsilon > 0$ and arbitrary $1 - \xi \in V$ vanishing outside a bounded interval I,

$$\begin{aligned} 1 - G_{n,i}[\xi] &= 1 - E\{\exp \int_I \log\xi(t)\,dN_{n,i}(t)\} \\ &= 1 - \{ \int_{\{N(I)=0\}} + \int_{\{N(I)>0\}} \}\exp[\int_I \log \xi(t)\,dN_{n,i}(t)]P_{n,i}(dN) \\ &\leq 1 - P_{n,i}\{N(I) = 0\} \\ &< \varepsilon \quad \text{uniformly in } i \quad \text{if} \quad n > n_o \end{aligned}$$

The converse implication follows on putting $\xi \equiv 0$ over suitable bounded intervals. ▼

Lemma 4.3 For a sequence of independent u.a.n. point processes $\{N_{n,i}(\cdot)\}$, $-\Sigma_{i=1}^{s_n} \log G_{n,i}[\xi]$ and $\Sigma_{i=1}^{s_n}(1-G_{n,i}[\xi])$ converge or diverge together, $1-\xi \in V$.

Proof. By u.a.n. and Lemma 4.2 $\log G_{n,i}[\xi]$ exists if n is sufficiently large. Then the expansion $-\log x = (1-x) + O(1-x)^2$ $0 < \delta \leq x \leq 1$ shows that

$$-\sum_{i=1}^{s_n} \log G_{n,i}[\xi] = \sum_{i=1}^{s_n} (1-G_{n,i}[\xi]) + \sum_{i=1}^{s_n} O(1-G_{n,i}[\xi])^2 \tag{4.8}$$

The remainder is positive and $\leq K \sum_{i=1}^{s_n} (1-G_{n,i}[\xi])^2$

$$\leq K \max_{1\leq i\leq s_n} (1-G_{n,i}[\xi]) \sum_{i=1}^{s_n} (1-G_{n,i}[\xi])$$

$$< \varepsilon \sum_{i=1}^{s_n} (1-G_{n,i}[\xi]) \qquad 1-\xi \in V$$

for large n, by Lemma 4.2.

If $\Sigma_{i=1}^{s_n}(1-G_{n,i}[\xi])$ converges then it is bounded and the remainder will be $o(1)$, so that the L.H.S. of (4.8) converges. If the L.H.S. of (4.8) converges then as both terms on the R.H.S. are positive they must be bounded and again the remainder is $o(1)$. ▼

This leads to another useful characterisation theorem (Kersten and Matthes [36])

Theorem 4.4 For a sequence of independent u.a.n. point processes $\{N_{n,i}(\cdot)\}$, $\Sigma_{i=1}^{s_n} N_{n,i}(\cdot) \xrightarrow{d} N(\cdot)$ if and only if $\Sigma_{i=1}^{s_n} E_{1,N_{n,i}}(\cdot) \xrightarrow{d} N(\cdot)$ as $n \to \infty$ (by definition $N(\cdot)$ is infinitely divisible).

Remark. This is a more exact statement of (4.7).

Proof. The logarithm of the p.g.fl of $\Sigma_{i=1}^{s_n} N_{n,i}(\cdot)$ is $\Sigma_{i=1}^{s_n} \log G_{n,i}[\xi]$, while for $\Sigma_{i=1}^{s_n} E_{1,N_{n,i}}(\cdot)$ it is $-\Sigma_{i=1}^{s_n}(1-G_{n,i}[\xi])$, so the theorem follows from Lemmas 2.3 and 4.3. ▼

We now illustrate the remarks of 4.1 on the application of the p.g.fl to limit theorems by proving the convergence of a superposition to the Poisson process in two different ways.

(α) Suppose the $N_{n,i}(\cdot)$ are identically distributed, for fixed n, as a stationary orderly point process $N(\cdot)$ with intensity λ, a Borel second factorial moment measure $M_2'(\cdot)$, and its time scale dilated by a factor n (this is the format of Vere-Jones [87]). Because $E\{N_{n,i}(I)\} = n^{-1}\lambda|I|$ the array is u.a.n., and because of orderliness we have, for all I,

$$n\, M_2\left(\frac{I}{n} \times \frac{I}{n}\right) \to 0 \quad \text{as} \quad n \to \infty \qquad \text{(Milne [58])}$$

We now apply (2.15) to get

$$1 - G_{n,i}[1-\xi] = n^{-1}\lambda \int \xi(t)dt + R_{n,i}[\xi], \qquad \xi \in V$$

where by Corollary 1 to Theorem 2.4

$$0 \leq R_{n,i}[\xi] \leq \frac{1}{2} M_2\left(\frac{I}{n} \times \frac{I}{n}\right) = o(n^{-1}) \quad \text{from above.}$$

Therefore as $n \to \infty$

$$\sum_{i=1}^{n} (1 - G_{n,i}[1-\xi]) = \lambda \int \xi(t)dt + o(1)$$

and by Lemma 4.3 $\Sigma_{i=1}^{n} N_{n,i}(\cdot)$ converges in distribution to a Poisson (λ) process.

This proof is elementary but the conditions imposed are stringent. In particular we cannot dispense with $M_2(\cdot)$ a Borel measure, which emphasises the limitations of the expansion (2.15).

(β) For more delicate results we imitate Vere-Jones [87]. By definition

$$G[1-\xi] = p(0;I) + E\{1-\xi(t)|N(I)=1\}p(1;I) + O[P\{N(I)\geq 2\}]$$

where $\xi \in V$ vanishes outside the bounded interval I. Then

$$\sum_{i=1}^{s_n}(1-G_{n,i}[1-\xi]) = \sum_{i=1}^{s_n} E\{\xi(t)|N_{n,i}(I)=1\}p_{n,i}(1;I)+O[\sum_{i=1}^{s_n} P\{N_{n,i}(I)\geq 2\}] \tag{4.9}$$

Theorem 4.5 (Grigelionis [25]). For a sequence of u.a.n. point processes $\{N_{n,i}(\cdot)\}$, $\Sigma_{i=1}^{s_n} N_{n,i}(\cdot) \xrightarrow{d}$ Poisson $(\Lambda(\cdot))$ if and only if for each bounded interval I

$$\lim_{n\to\infty} \sum_{i=1}^{s_n} p_{n,i}(1;I) = \Lambda(I) \tag{4.10}$$

$$\lim_{n\to\infty} \sum_{i=1}^{s_n} [1-p_{n,i}(0;I)-p_{n,i}(1;I)] = 0 . \tag{4.11}$$

Proof. The necessity follows as in [25] and Lemma 4.3 proves the sufficiency if the L.H.S. of (4.9) converges to $\int\xi(t)\Lambda(dt)$ as $n \to \infty$. By (4.11) the remainder on the R.H.S. of (4.9) is $o(1)$. Choose for $\xi \in V$ a simple function $\Sigma_{j=1}^{k} z_j \chi_{I_j}$, $\bigcup_{j=1}^{k} I_j = I$, whence

$$E\{\xi(t)|N_{n,i}(I)=1\}p_{n,i}(1;I) = \sum_{j=1}^{k} z_j P\{N_{n,i}(I)=1, N_{n,i}(I_j)=1\}$$

$$\sim \sum_{j=1}^{k} z_j p_{n,i}(1;I_j)$$

by (4.11) so that from (4.10)

$$\lim_{n\to\infty} \sum_{i=1}^{s_n} E\{\xi(t)|N_{n,i}(I)=1\}p_{n,i}(1;I) = \sum_{j=1}^{k} z_j\Lambda(I_j) = \int\xi(t)\Lambda(dt).$$

By taking a sequence of simple functions converging monotonely to ξ this last result is true for arbitrary $\xi \in V$. ▼

4.4 Deletion and Translation Theorems for Point Processes

Numerous writers have considered limit theorems arising from the operations of deletion and translation and have characterised the Poisson process as the only stream invariant under these operations. We illustrate applications of the p.g.fl in this field, using the following technique (Vere-Jones [87]); if a sequence of p.g.fls $G_n[\xi]$ can be represented as $G[\xi_n]$, for some p.g.fl G, then clearly $\int \log \xi_n(t)dN(t)$ converging in any mode to a random variable Y implies $G_n[\xi] = G[\xi_n] \to E(e^{-Y})$.

Formally, the deletion operation cancels the points of an arbitrary point process $N(\cdot)$ with probability $1-q$ and retains them with probability q, independently for each point, thus generating a new stream $N_q(\cdot)$. If G is the p.g.fl of $N(\cdot)$ then as $N_q(\cdot)$ is a very special cluster process, with clusters of size 0 or 1 occurring at the cluster centre, (3.1) shows that

$$G_q[\xi] = G[1-q+q\xi] \tag{4.12}$$

So our general technique is applicable and we have

<u>Theorem 4.6</u> (Belyaev [9]). Suppose

$$\lim_{|I|\to\infty} P\left\{\left|\frac{N(I)}{|I|} - \lambda\right| > \varepsilon\right\} = 0 \tag{4.13}$$

uniformly over all intervals I of finite length. Then if we contract the time scale by a factor q, $N_q(\cdot) \overset{d}{\to}$ Poisson (λ) process as $q \to 0$.

Proof. We shall prove that for $\xi \in V$

$$Z_q = \int \log[1-q\xi(qt)]dN(t) \overset{p}{\to} -\lambda\int\xi(t)dt \quad \text{as} \quad q \to 0 \tag{4.14}$$

Expanding the logarithm, as $q,\xi(t) < 1$, we have

$$\log\,[1-q\xi(qt)] = -q\xi(qt) - B(q,\xi)q^2\xi(qt)$$

where $B(\cdot)$ is uniformly bounded for $q < 1$, $\xi \in V$. Then from (4.14)

$$Z_q = -q\int\xi(qt)dN(t) - q^2\int B(q,\xi)\xi(qt)dN(t),$$

so that if $q\int\xi(qt)dN(t) \overset{p}{\to} \lambda\int\xi(dt)dt$ as $q \to 0$ so will Z_q as the remainder will be $o(1)$ in probability. (4.12) and Lemma 2.3 then complete the proof.

When $\xi(t)$ is a simple function $\Sigma_{i=1}^{k} z_i\chi_{I_i}(t)$, elementary considerations give

$$P\{|q\int\xi(qt)dN(t)-\lambda\int\xi(t)dt|>\varepsilon\} \leq \sum_{i=1}^{k} P\{\left|qN(\tfrac{I_i}{q})-\lambda|I_i|\right| > \tfrac{\varepsilon}{k}\} \tag{4.15}$$

$$< \varepsilon \tag{4.15}$$

as $q \to 0$ by (4.13). So our result is true for simple functions ξ.

For arbitrary $\xi \in V$ choose a monotone sequence of simple functions $\{\xi_n\}$ with $\xi_n(t) \uparrow \xi(t)$ uniformly as $n \to \infty$, so that

$$\lambda\int\xi_n(t)dt \to \lambda\int\xi(t)dt \qquad \text{as} \quad n \to \infty \tag{4.16}$$

And given $\varepsilon' > 0$ we can choose n so large that for all q

$$0 \leq q\int[\xi(qt) - \xi_n(qt)]dN(t) < \varepsilon' qN(\tfrac{J}{q}),$$

where J is the bounded interval outside which ξ vanishes. Since $qN(\frac{J}{q})$ is convergent in probability,

$$P\{q\int[\xi(qt) - \xi_n(qt)]dN(t) > \varepsilon\} < \varepsilon \tag{4.17}$$

as $n \to \infty$, for fixed q near zero.

Thus for $\xi \in V$ and $\varepsilon > 0$,

$$P\{|q\int\xi(qt)dN(t)-\lambda\int\xi(t)dt|>\varepsilon\} \leq P\{|q\int\xi_n(qt)dN(t)-\lambda\int\xi_n(t)dt|>\tfrac{\varepsilon}{3}\}$$

$$+ P\{q\int[\xi(qt)-\xi_n(qt)]dN(t)>\tfrac{\varepsilon}{3}\}+P\{\lambda\int[\xi(t)-\xi_n(t)]dt>\tfrac{\varepsilon}{3}\}$$

$$< \varepsilon$$

by (4.15)-(4.17), which proves the theorem. ▼

A simple corollary is that if λ in (4.13) is a random variable, a conditional argument proves convergence to a mixed Poisson process. Theorem 4.6 generalises Goldman [22], as he requires a.s. convergence in (4.13).

The translation operation adds to each point t_i of a point process $N(\cdot)$ a random variable Y_i, where the Y_i are independently and identically distributed and independent of the t_i. This is again a cluster process, with cluster size one. If the Y_i are non-negative we have a $G/G/\infty$ queue (cf. 3.6).

Let $N(\cdot)$ have p.g.fl $G[\xi]$ and the translations $Y(x)$ have distribution function $F_x(y)$. Then from (3.3) the p.g.fl of the translated stream $N_x(\cdot)$ is

$$G_x[\xi] = G[\int\xi(t+y)dF_x(y)] \tag{4.18}$$

$$= G[\xi_x] \qquad \text{say, as we require.}$$

<u>Theorem 4.7</u> If for all bounded intervals I

(A) $\sup_y F_x(I-y) \to 0$ as $x \to \infty$

(B) $\int F_x(I-t)dN(t) \overset{p}{\to} \lambda|I|$ as $x \to \infty$

then $N_x(\cdot)$ converges in distribution to a Poisson (λ) process.

<u>Proof</u>. Consider again the variable $Z_x = \int \log[1-\xi_x(t)]dN(t)$, where $\xi \in V$. The usual expansion gives

$$Z_x = -\int \xi_x(t)dN(t) - \int \xi_x^2(t)R_x[\xi,t]dN(t),$$

where

$$0 \leq R_x[\xi,t] \leq \tfrac{1}{2}[1-\xi_x(t)]^{-1}.$$

Let I be the support of ξ. Then $\xi_x(t) \leq F_x(I-t)$ and as before if $\int \xi_x(t)dN(t) \overset{p}{\to} \lambda \int \xi(t)dt$ as $x \to \infty$ so does Z_x, because the remainder will converge to zero in probability by (A). Then (4.18) and Lemma 2.3 complete the proof.

When $\xi(t)$ is a simple function $\Sigma_{i=1}^k z_i \chi_{I_i}(t)$ then by (B)

$$\int \xi_x(t)dN(t) = \sum_{i=1}^k z_i \int F_x(I_i-t)dN(t)$$

$$\overset{p}{\to} \sum_{i=1}^k z_i \lambda|I_i| = \lambda \int \xi(t)dt \tag{4.19}$$

as $x \to \infty$. So as in Theorem 4.6, taking an increasing sequence of simple functions converging uniformly to $\xi \in V$, we can show that (4.19) holds for arbitrary ξ in V. This proves the theorem. ▼

Theorem 4.7. can obviously be extended to the case of λ a random variable, just as before. It generalises similar results of Goldman [22] and Thedéen [83] by assuming only convergence in probability and not with probability one in (B) and not requiring that $N(\cdot)$ be well-distributed. However it does not provide a necessary condition for convergence.

Corollary 1. If $F_n(y)$ is the n-fold convolution of a distribution function $F(y)$ and $N(\cdot)$ is stationary with intensity λ and an integrable covariance density, then (A) and (B) are satisfied as $n \to \infty$.

This is the output through a sequence of $G/G/\infty$ queues (Vere-Jones [87]).

Corollary 2. For the $D/M/\infty$ queue, with service distribution $F_x(y) = 1 - \exp\{-\frac{\alpha y}{x}\}$, (A) and (B) are satisfied with $\lambda = 1$ as $x \to \infty$.

This result was asserted by Nelsen and Williams [64]. (A) is obvious, and for (B) let $I = [a,b)$. Then

$$\int F_x(I-y)\,dN(y) = \int_{-\infty}^{a} \left(\exp\{-\frac{\alpha(a-y)}{x}\} - \exp\{-\frac{\alpha(b-y)}{x}\}\right) dN(y) + \int_{a}^{b} F_x(b-y)\,dN(y) \; .$$

Obviously the last term goes to zero, and the first term is

$$(1-\exp\{-\frac{\alpha}{x}(b-a)\}) \cdot \int_{-\infty}^{0} \exp\{\frac{\alpha z}{x}\} dN(z+a) \sim \frac{1-\exp\{-\frac{\alpha}{x}(b-a)\}}{1-\exp\{-\frac{\alpha}{x}\}}$$

$$\to (b-a) = |I| \quad \text{as} \quad x \to \infty.$$

Finally, a word on characterisation theorems. Dobrushin [16], Goldman [22] and Thedéen [83] prove that the mixed Poisson process is characterised, among various classes of point processes, by invariance under translation and Mecke [57] (see also Nawrotzki [63]) establishes this invariance for doubly stochastic Poisson processes and deletion.

A related question is considered by Szász [81], namely what cluster mechanisms ensure that a Poisson cluster process is again a Poisson process? We can easily prove his result from the basic p.g.fl formula; it shows that only deletion and translation are allowed.

Theorem 4.8 A Poisson cluster process is a stationary Poisson process if and only if $p_n = P\{N_s(-\infty,\infty)=n\} = 0, \quad n \geq 2.$

Proof. The sufficiency is obvious, and the Poisson cluster process is then a Poisson $(\lambda[1-p_o])$ process. Conversely, if this is true then taking logarithms in (3.27) we find

$$\lambda\int(1-G_2[\xi|t])dt = \lambda(1-p_o)\int[1-\xi(t)]dt \ .$$

Now $N_s(\cdot)$ is a.s. finite so $G_2[\xi|t]$ has a representation (2.3). If $\pi_n = (1-p_o)^{-1}p_n$ then

$$\int\{1-\sum_{n=1}^{\infty}\pi_n\int\cdot\cdot\int\xi(x_1+t)\cdot\cdot\xi(x_n+t)dU_n(x_1,\ldots,x_n)\}dt = \int[1-\xi(t)]dt \ .$$

Since $1 - \xi \in V$, choose $\xi(x) = 1-(1-z)\chi_I(x)$ for some interval I. On equating coefficients of z we obtain equations like

$$\sum_{n=2}^{\infty}\pi_n\int U_n(I-t,I-t,\overline{I-t},\ldots,\overline{I-t})dt \equiv 0,$$

where $\overline{A}$ is the complement of the set A. Since I is arbitrary, this implies $\pi_n = 0$ for $n = 2,3,\ldots$ ▼

5. IDENTIFIABILITY IN LINEAR STOCHASTIC POISSON PROCESSES

5.1. Introduction

In 1.5 it was pointed out that the linear stochastic Poisson process has a great variety of uses in the theory and applications of point processes. It will be shown that this process is determined by two parametric functions, hereafter called parameters for brevity, and so it is important to be able to determine them uniquely from the process. This is the identifiability problem studied in the present chapter. We do not attempt to estimate the parameters although we do prove that, under certain conditions, the parameters are identifiable from quantities which should provide reasonable estimates in practice.

The results established below for linear processes are of interest in their own right and may be extended to a wider class of linear processes than those we consider. In particular, we can drop certain non-negativity assumptions. For details see the author's paper [91], on which this chapter is based.

5.2 Preliminary Results

The linear stochastic Poisson process was defined in 1.5, where we assumed f was non-negative, integrable and square-integrable. We now write $f \varepsilon B_2$, where we say that $f \varepsilon B_n$ if f is non-negative and f^p is integrable for $p = 1,\ldots,n$. A simple criterion for this is given by

Lemma 5.1 If f is bounded and integrable then $f \varepsilon B_n$ for all n.

The cumulants of $X(t)-X(t-1)$ are denoted by κ_n, with $\kappa_2 = \sigma^2 < \infty$ by assumption.

Since the moments of Λ involve integrals of powers of f, Lemma 5.1 is a convenient way of ensuring that all these integrals are finite. In many cases the boundedness assumption will be reasonable, e.g. the 'shot-noise' process defined in Bartlett [6] p.161, although we do not use this assumption for general results.

The fundamental relation between the p.g.fl of a doubly stochastic Poisson process and the Laplace functional of its mean process is given by (2.10). When the mean is a linear process, as considered here, we can evaluate the functionals explicitly. Since $X(\cdot)$ is homogeneous additive with non-negative increments its Laplace transform is

$$E\{\exp(-\theta[X(t+\tau)-X(t)])\} = \exp\{-\tau\psi(\theta)\} \tag{5.1}$$

where

$$\psi(\theta) = \gamma\theta + \int_0^\infty \frac{1-e^{-\theta x}}{x}\, dK(x), \tag{5.2}$$

γ is a non-negative constant and K a non-decreasing function on $(0, \infty)$ with $K(0) = 0$ and $\int_1^\infty x^{-1} dK(x) < \infty$ (Baxter and Shapiro [8], Zolotarev [94]; this form is from Feller [18] p.426). Because $E\{X^2(t)\} < \infty$, in fact K is bounded and $\int_0^\infty x dK(x) < \infty$. It turns out that this kernel is mathematically more convenient in our work than the usual $1-e^{-\theta x}$.

It is now easily shown that

$$L_\Lambda[\xi] = \exp\{-\int\psi[\int f(t-u)\xi(dt)]du\} \tag{5.3}$$

(Bartlett [6], p.161, Shiryaev [77]), and consequently

$$G[\xi] = \exp\{-\int\psi(\int f(t)[1-\xi(t+u)]dt)du\}. \tag{5.4}$$

From this result we can of course derive any desired property of the process.

Equations (5.3), (5.4) show that the parameters f, ψ determine the process (linear or linear stochastic). In this chapter we investigate the converse question of when they are uniquely specified by the process i.e. when they are identifiable. Clearly the results will be the same whichever process we choose to discuss, because of (2.10); in fact we give a detailed account of the identifiability of ψ for linear processes and of f for the linear stochastic Poisson process.

It is plausible that because of the comprehensive nature of the two functionals identifiability will always hold, in the sense that each process is associated with an f and ψ unique up to constant multipliers. However the difficulty of obtaining a sample estimate of the p.g.fl, say, means that in practice estimators of the parameters will be quantities derivable from the p.g.fl but not containing all its information about the process. We are principally concerned with this problem.

An identifiability theorem for another special class of doubly stochastic Poisson processes, namely those which are renewal processes, is given in Kingman [40]. As remarked before, they cannot be linear stochastic Poisson processes.

5.3 Identifiability of ψ

Since $X(t)$ is homogeneous and additive it can be expressed as

$X(t) \overset{d}{=} \Sigma_{i=1}^{n} X_{i,n}(t)$, where the $X_{i,n}(t)$ are independent and identically distributed homogeneous additive processes. Consequently

$$\Lambda(t) \overset{d}{=} \sum_{i=1}^{n} \int f(t-u)dX_{i,n}(u)$$

for any n, so that $\Lambda(t)$ is infinitely divisible (Lee [45]). In view of (2.10) and 4.3, linear stochastic Poisson processes must also be infinitely divisible. We now calculate the canonical form of its Laplace transform in

Lemma 5.2 Let $\chi(f,\theta)$ be the logarithm of the Laplace transform of a linear process (1.11) and let

$$h(y) = m\{x : f(x) > y\}, \quad y > 0,$$

where $m\{\cdot\}$ is Lebesgue measure. If $f \in B_2$ and h has a derivative h' existing and non-zero at each point, then $\chi(f,\theta)$ has the canonical form given by (5.6), (5.7) below.

Proof. From (5.3)

$$\chi(f,\theta) = -\int\psi[\theta f(u)]du. \tag{5.5}$$

Since f is non-negative and integrable we may rearrange it as a decreasing function g on $[0, \infty)$ whose inverse exists and equals h(the technique of rearrangement is described in Hardy, Littlewood and Pólya [27] §10.12). Our hypotheses show that g is uniquely defined everywhere with finite derivative g', and $g'.h' = 1$.

Now integrals both of a function and measurable functions of the function are invariant under rearrangement. Since ψ is measurable,

$$-\chi(f,\theta) \equiv -\chi(g,\theta) = \gamma\theta\int_0^\infty g(u)du + \int_0^\infty\int_0^\infty \frac{1-e^{-\theta x g(u)}}{x} dK(x)du.$$

Change the variable by $v = xg(u)$. Then

$$-\chi(f,\theta) = \hat{\gamma}\theta + \int_0^\infty \int_0^{xg(o)} \frac{1-e^{-\theta v}}{v} \cdot \frac{v}{x^2}|h'(\frac{v}{x})|dv\, dK(x)$$

since clearly $g(u) \to 0$ as $u \to \infty$. Set $b = g(0) = \sup_u f(u) \leq \infty$; if $b = \infty$ then $b^{-1} = 0$. So

$$-\chi(f,\theta) = \hat{\gamma}\theta + \int_0^\infty \int_{vb^{-1}}^\infty \frac{1-e^{-\theta v}}{v} \cdot \frac{v}{x^2}|h'(\frac{v}{x})|dK(x)dv$$

$$= \hat{\gamma}\theta + \int_0^\infty \frac{1-e^{-\theta v}}{v}\, d\hat{K}(v) \tag{5.6}$$

where

$$\hat{\gamma} = \gamma \int_0^\infty g(u)du = \gamma \int_{-\infty}^\infty f(u)du$$

$$\frac{d}{dv}\hat{K}(v) = \hat{k}(v) = \int_{vb^{-1}}^\infty \frac{v}{x^2}|h'(\frac{v}{x})|dK(x) \tag{5.7}$$

cf. (5.2). Since we know Λ is infinitely divisible there is no need to check the suitability of (5.7). ▼

Lemma 5.3 (Kendall and Lewis [35]). If $\phi(\theta)$ is the characteristic function of a non-negative random variable, θ real, then the set of zeros of ϕ has Lebesgue measure zero.

We can now prove an identifiability result for the transform χ.

Theorem 5.1 Let $\chi(f,\theta)$, f, and h be defined as in Lemma 5.2, and in addition let f be bounded. If f is known, then $\chi(f,\theta)$ uniquely determines ψ and conversely.

Proof. The uniqueness of the canonical form (5.6) ensures that $\hat{\gamma}$, $\hat{K}$ are unique to χ. Clearly γ is uniquely determined by $\hat{\gamma}$, since f is assumed known, and we show the same is true for K.

As f is bounded we may assume without loss of generality that $b = 1$. Then

$$\hat{K}(\infty) - \hat{K}(v) = \int_v^\infty \int_y^\infty \frac{y}{x^2} \left|h'\left(\frac{y}{x}\right)\right| dK(x)dy$$
$$= \int_v^\infty \int_{vx^{-1}}^1 z|h'(z)|\,dz\,dK(x) \ . \qquad (5.8)$$

Now $z|h'(z)|$ is integrable over $[0, 1]$. For on integration by parts we see that this assertion is equivalent to $\lim_{z\to 0+} zh(z)$ finite. But

$$\int f(x)dx \geq \delta \sum_{n=1}^{\infty} h(n\delta)$$

for any $\delta > 0$, and as f is integrable the result follows. Together with K bounded this shows that (5.8) is proportional to the upper tail of the distribution function of a product of two independent random variables, say W_1, W_2, on $[0, \infty)$ and $[0, 1]$ respectively. If $U = W_1.W_2$ the problem becomes, does knowledge of the distributions of U and W_2 determine the distribution of W_1? Take logarithms to get a sum of independent random variables. Then as $-\log W_2$ has range $[0, \infty)$ Lemma 5.3 shows that its Fourier transform is non-zero almost everywhere, and we immediately conclude that K is uniquely determined by $\hat{K}$ and hence χ. As the converse is trivial, the theorem is proved. ▼

<u>Remarks</u> 1. The restrictions on h in Lemma 5.2 exclude functions f with intervals of constancy and certain kinds of discontinuities, in particular simple functions. It is possible to prove a similar result for simple functions by a different method.

2. If f is unbounded the theorem is still valid if the logarithm of a random variable with density function proportional to $z|h'(z)|$ has a Fourier transform non-zero almost everywhere. Boundedness ensures in a simple manner that this is so.

3. For a general linear process, with signed f and signed increments for $X(t)$, Lemma 5.2 is still true if we replace Laplace transforms by characteristic functions and use the canonical form for an infinitely divisible distribution with finite variance (Lukacs [53] p.90). However the uniqueness result of Theorem 5.1 cannot be proved, at least by the same method, unless f is non-negative (for details see [91]).

All this leads directly to

Theorem 5.2 A linear process (1.11) uniquely determines the Laplace transform of $X(t)$, and conversely, for known $f \in B_2$.

Proof. From (5.3), the Laplace functional of a linear process has the form (5.5) with a modified f. Since $\xi(\cdot)$ is an arbitrary measure we can choose $\xi(\cdot)$ so that $\int f(t-u)\xi(du)$ satisfies the conditions of Theorem 5.1. But $L_\Lambda[\xi]$ contains full information about the process. ▼

We note that if f is bounded then ψ is identifiable from the Laplace transform of $\Lambda(t)$ alone, that is from the distribution at a point.

Since (5.4) shows that the p.g.fl of the linear stochastic Poisson process is also of the form (5.5), and $1-\xi$ is arbitrary other than being in V, say, we deduce immediately

<u>Theorem 5.3</u> A linear stochastic Poisson process uniquely determines the Laplace transform of $X(t)$, and conversely, for known $f \varepsilon B_2$.

5.4 <u>Identifiability of f</u>

In the previous section we established the identifiability of ψ for known f. Here we investigate the converse question and assume throughout that ψ is known.

As we know, the factorial cumulant measures of a point process are uniquely determined by the process and may be calculated, for the linear stochastic Poisson process, by the use of Theorem 2.4 and (5.4). If $\kappa_n < \infty$ and $f \varepsilon B_n$, we expand $\frac{1-e^{-\theta x}}{x}$ as a Taylor series to n terms to get

$$H[1-\xi] = \int\left\{\sum_{k=1}^{n} \frac{(-1)^k}{k!} \kappa_k[r(w)]^k + R_n(w)\right\}dw \tag{5.9}$$

where $\xi \varepsilon V$ and

$$r(w) = \int f(v-w)\xi(v)dv$$

$$R_n(w) = \frac{(-1)^n}{n!} \int_0^\infty x^{n-1}[r(w)]^n e^{-\delta x r(w)} dK(x) \qquad 0 < \delta < 1.$$

Since $[r(w)]^n$ is integrable and $|R_n(w)| \leq \frac{\kappa_n}{n!}[r(w)]^n$, (5.9) and the Corollary to Theorem 2.4 show that the linear stochastic Poisson process has factorial cumulant *densities* $c_k(\cdot)$ existing, with

$$c_k(t_1,\ldots,t_k) = \kappa_k \int f(t_1-w)\ldots f(t_k-w)dw \tag{5.10}$$

$k=1,\ldots,n$. $c_2(\cdot)$ is of course the covariance density.

The uniquely determined functions $c_k(\cdot)$ are possible estimators of f, so it is of interest to see whether they provide identifiability. We first establish this in the important case when f is restricted to a half line i.e. $f(x) \equiv 0$ for $x < 0$. This means that the linear process is expressed in terms of past values of $X(t)$ only, which does not seem unreasonable.

Since the point process is stationary $c_2(t_1,t_2) = c_2(t_1-t_2)$, and $c_3(t_1,t_2,t_3) = c_3(t_2-t_1,t_3-t_1)$. Now take Fourier transforms, denoted by an asterisk, in (5.10). We have

$$c_2^*(\omega) = \sigma^2\phi(\omega)\overline{\phi(\omega)} \tag{5.11}$$

$$c_3^*(\omega_1,\omega_2) = \kappa_3\ \phi(\omega_1)\phi(\omega_2)\ \overline{\phi(\omega_1+\omega_2)}, \tag{5.12}$$

where the bar denotes complex conjugate and ϕ is the Fourier transform of f which exists because f is integrable. As ψ is known so are σ^2, κ_3, so that (5.11),(5.12) are exactly equivalent to equations (5), (6) of Kendall and Lewis [35], with ϕ replacing their Ψ, because f vanishes on a half line. From their paper we conclude that f is identifiable up to a location factor; clearly from the form of (5.10) we can expect no better than this. Hence

<u>Theorem 5.4</u> A linear stochastic Poisson process with ψ known, $\kappa_3 < \infty$, and $f \in B_3$ vanishing identically in $(-\infty,0)$ determines f uniquely up to a location factor.

Theorem 5.4 of course answers the question completely. However it involves the knowledge of both second and third factorial cumulant densities and in estimation we would like to use just one. This leads to the investigation of identifiability of f from a single $c_k(\cdot)$ of low order.

If we know only $c_2(\cdot)$ it is easy to construct instances where two different functions f give the same $c_2(\cdot)$. For example, $c_2^*(\omega) = (1+\omega^2)^{-2}$ arises from a gamma density with two degrees of freedom and also a bilateral exponential density. However if we again restrict attention to functions vanishing on $(-\infty, 0)$ the problem is equivalent to a much-studied question in the theory of linear prediction of a stationary process from its past history (Doob [17], chapter 12, Bartlett [6], chapter 7). The crucial point is to determine a Fourier transform knowing only its modulus, a topic discussed briefly below. This is essentially Doob's Theorem 12.5.2.

As f^2 is integrable and vanishes on $(-\infty, 0)$ a theorem of Paley and Wiener (Hoffman [29] p.131) shows that its complex Laplace transform $f^{**}(s) = \int_0^\infty e^{-sx} f(x)dx$, $\text{Re } s \geq 0$, is a member of the Hardy class H^2 in the right half plane. It is known ([29] p.132) that H^2 functions are uniquely factorisable into a product of three functions - one involving only the zeros of f^{**} in $\text{Re } s > 0$, one involving a singular measure which vanishes for f^{**} continuous on the imaginary axis, and one involving only $|f^{**}(i\omega)| = |\phi(\omega)|$. So if $f^{**}(s) \neq 0$ for $\text{Re } s > 0$ then $f^{**}(s)$, and hence $\phi(\omega)$, is uniquely determined by $|\phi(\omega)|$ up to a factor $e^{ic\omega}$, which proves

Theorem 5.5 A linear stochastic Poisson process with ψ known, $f \in B_2$ vanishing identically on $(-\infty, 0)$ and $f^{**}(s) \neq 0$ for $\text{Re } s > 0$ determines f uniquely from $c_2(\cdot)$ up to a location factor.

We note here that another theorem of Paley and Wiener [69] shows $c_2^*(\omega)$ is the squared modulus of the Fourier transform of a function vanishing on a half line if and only if

$$\int \frac{\left|\log\left|c_2^*(\omega)\right|\right|}{1+\omega^2}\, d\omega < \infty \tag{5.13}$$

Unfortunately there seems to be no known characterisation of the class of functions for which $f^{**}(s) \neq 0$, $\operatorname{Re} s > 0$, though it contains most of the common weighting functions.

This covers the case of knowing $c_2(\cdot)$ alone. Initially there seems little interest in working with $c_3(\cdot)$ alone, as $c_2(\cdot)$ will generally be easier to estimate and we already have Theorem 5.4. However we shall see that from $c_3(\cdot)$ we may identify functions f with unrestricted range, a situation not previously considered.

By stationarity $c_3(\cdot)$ is a non-negative bivariate function integrable over the plane. Its transform $c_3^*(\omega_1,\omega_2)$ is the bispectrum studied by Rosenblatt and Van Ness [74], who discuss estimation for both $c_3^*(\cdot)$ and $c_3(\cdot)$. Now take $\omega_2 = k\omega_1 \equiv k\omega$ for all real ω and $k \neq 1$. Then (5.12) becomes, neglecting the known $\kappa_3 < \infty$,

$$c_3^*(\omega,\ k) = \phi(\omega)\phi(k\omega)\ \overline{\phi[(k+1)\omega]}$$

and

$$c_3^*(\theta) = \phi(\theta)\phi(\beta_1\theta)\phi(\beta_2\theta)$$

where

$$\beta_1 = -\frac{1}{k+1},\ \beta_2 = -\frac{k}{k+1},\quad \theta = -\frac{\omega}{k+1}\quad .$$

Let $\Phi(\theta) = \log\phi(\theta)$, which exists in some interval about $\theta = 0$ because ϕ is proportional to a characteristic function. Then

$$\log c_3^*(\theta) = \Phi(\theta) + \Phi(\beta_1\theta) + \Phi(\beta_2\theta),$$

so to see whether $c_3^*(\cdot)$ uniquely determines Φ we must find the solutions of

$$\Phi(\theta) + \Phi(\beta_1\theta) + \Phi(\beta_2\theta) \equiv 0 \qquad (5.14)$$

in some interval about zero. This is achieved by

Lemma 5.4 (Rao [71]) If $\Phi(\theta)$ is given, in an interval of θ about zero, by

$$\sum_{i=1}^{n-1} \alpha_i\, \Phi(\beta_i\theta) + \Phi(\theta) \equiv 0$$

where $\Sigma_{i=1}^{n-1} \alpha_i\beta_i = -1$, $|\beta_i| < 1$, $\alpha_i\beta_i < 0$ $i=1,\ldots,n-1$, and $\Phi(\theta)$ has a derivative continuous at zero, then

$$\Phi(\theta) = c\theta.$$

By suitable choice of $k (k > 0$ for instance), our β_i satisfy the conditions of Lemma 5.4 and we get

Theorem 5.6 A linear stochastic Poisson process with ψ known, $\kappa_3 < \infty$, and

$$\int |x| f(x)dx < \infty \qquad (5.15)$$

determines $f \in B_3$ uniquely from $c_3(\cdot)$ up to a location factor.

Proof. Condition (5.15) ensures that Φ has a continuous derivative everywhere (Feller [18] p.485). Then Lemma 5.4 shows that the indeterminacy in f is at most a location factor; for if Φ_1, Φ_2 are two possible solutions of (5.14) we see that $\Phi_1(\theta) - \Phi_2(\theta) = c\theta$ in an interval about zero and hence, by analytic continuation, over the whole line, where c is a purely imaginary constant. ▼

We emphasise that Theorem 5.6 applies to functions over the whole real line. If we consider knowing $c_4(\cdot)$, $c_5(\cdot)\ldots$ exactly the same procedure will produce the functional equation (5.14) with additional terms, and under the conditions of the theorem it will have the same solution in all these cases. Thus no new identifiability criteria emerge. It is striking that the second factorial cumulant density gives such a different set of conclusions from all the higher order densities.

If f in fact vanishes on $(-\infty, 0)$ then Theorem 5.6 is not as general as Theorem 5.5. To see this, take f as the symmetric stable density function of order $\frac{1}{2}$ (Feller [18] p.170). It is bounded and integrable so is certainly in B_2 and B_3 (Lemma 5.1). Its transform satisfies (5.13), as it must, and also the condition of Theorem 5.5, but its mean is infinite so that (5.15) does not hold.

Suppose now that ψ is unknown also. Then for two linear stochastic Poisson processes with parameters (f_1,ψ_1), (f_2,ψ_2) and the same p.g.fl we see, on integration of (5.10), that for $m,n = 1,2,\ldots$

$$(\kappa_n^{(1)})^m(\kappa_m^{(2)})^n = (\kappa_n^{(2)})^m(\kappa_m^{(1)})^n \tag{5.16}$$

Consequently we can use any of the preceding results for f to show that, under suitable conditions, f_1 and f_2 differ only in location and scale (by a factor $\sigma^2_{(1)}/\sigma^2_{(2)}$). If however at least one of the cumulant pairs $\kappa_n^{(1)}$, $\kappa_n^{(2)}$ is equal then all the cumulants are equal and f_1, f_2 differ only in location. We conclude that dropping the assumption of known ψ leads to indeterminacy up to a constant multiple as well as in location.

It is easy to apply these results to more general linear processes. Consider a linear process with non-negative f and signed increments for $X(t)$. Then by the results of Shiryaev [77] its characteristic functional has a unique Taylor expansion very like (2.15), in terms of a set of cumulant functions $s_k(t_1,\ldots,t_k)$. From the characteristic functional analogue of (5.3) it is easily shown that $s_k(t_1,\ldots,t_k) = c_k(t_1,\ldots,t_k)$ and so all the theorems of this section carry over directly. In particular, Theorem 5.4 shows that a linear process belonging to $T^{(3)}$ of [77] uniquely determines $f \varepsilon B_3$ up to a location factor for known ψ. Again it is not possible to carry the theory over directly to signed f as then the result of Kendall and Lewis [35] used in Theorem 5.4 may not hold and (5.14) may also be invalidated. Simple additional conditions will remove these difficulties.

REFERENCES

In references to translated Russian journals, the pagination given is that of the translation.

1 Agnew, R. Transformations of Uniform and Stationary Point Processes Ph.D. Thesis, Northwestern University, 1968.

2 Badrikian, A. Remarques sur les théorèmes de Bochner et P. Lévy Lecture Notes in Mathematics 31, Springer, 1967, 1-19.

3 Bartlett, M.S. Processus stochastiques ponctuels. Ann. Inst. Henri Poincaré 14, 1955, 35-60.

4 ——— The spectral analysis of point processes. J. Roy. Statist. Soc. Ser. B 25, 1963, 264-296.

5 ——— The spectral analysis of two-dimensional point processes. Biometrika 51, 1964, 299-311.

6 ——— Stochastic Processes 2nd ed. Cambridge University Press, 1966.

7 ——— and D.G. Kendall. On the use of the characteristic functional in the analysis of some stochastic processes occurring in physics and biology. Proc. Camb. Phil. Soc. 47, 1951, 65-76.

8 Baxter, G. and J.M. Shapiro. On bounded infinitely divisible distributions. Sankhyā 22, 1960, 253-260.

9 Belyaev, Yu. V. Limit theorems for dissipative flows. Th. Prob. Appl. 8, 1963, 165-173.

10 Beutler, F.J. and O.A.Z. Leneman. The theory of stationary point processes. Acta Math. 116, 1966, 159-197.

11 Bochner, S. Stochastic processes. Ann. Math. 48, 1947, 1014-1061.

12 Breiman, L. The Poisson tendency in traffic distribution. Ann. Math. Statist. 34, 1963, 308-311.

13 Cox, D.R. Some statistical methods connected with series of events. J.Roy. Statist. Soc. Ser. B 17, 1955, 129-164.

14 ——— and P.A.W. Lewis. The Statistical Analysis of Series of Events. Methuen 1966.

15 Daley, D.J. Spectral properties of weakly stationary point processes. 37th I.S.I. Conference, London, 1969, 344-346.

16 Dobrushin, R.L. On Poisson laws for distributions of particles in space (Russian). Ukrain. Mat. 8, 1956, 127-134.

17 Doob, J.L. Stochastic Processes Wiley, 1953.

18 Feller, W. An Introduction to Probability Theory and its Applications. Vol. II. Wiley, 1966.

19 Fieger, W. Eine für beliebige Call-Prozesse geltende Verallgemeinerung der Palmschen Formeln. Mat. Scand. 16, 1965, 121-147.

20 Franken, P., A. Liemant and K. Matthes. Stationäre zufällige Punktfolgen III. J-ber. Deutsch. Math. Verein. 67, 1964, 183-202.

21 ——— and G. Richter. Über eine Klasse von zufälligen Punktfolgen. Wiss. Z. Friedrich-Schiller Univ. Jena 14, 1965, 247-249.

22 Goldman, J.R. Stochastic point processes; limit theorems. Ann. Math. Statist. 38, 1967, 771-779.

23 ——— Infinitely divisible point processes in R^n. J.Math. Anal. Appl. 17, 1967, 133-146.

24 Grenander, U. Probabilities on Algebraic Structures Almqvist and Wiksell, 1963.

25 Grigelionis, B. On the convergence of random step processes to a Poisson process. Th. Prob. Appl. 8, 1963, 177-182.

26 Halmos, P.R. Measure Theory Van Nostrand, 1950.

27 Hardy, G., J.E. Littlewood and G. Pólya. Inequalities Cambridge University Press, 1934.

28 Harris, T.E. The Theory of Branching Processes Springer, 1963.

29 Hoffman, K. Banach Spaces of Analytic Functions Prentice-Hall, 1962.

30 Ibragimov, I.A. Some limit theorems for stationary processes. Th. Prob. Appl. 7, 1962, 349-382.

31 Jacobs, K. On Poincaré's recurrence theorem. Proc. 5th Berkeley Symp. 2, II, 1965, 375-404.

32 Jiřina, M. Branching processes with measure-valued states. Trans. 3rd Prague Conf. 1964, 333-357.

33 Jiřina, M. Asymptotic behaviour of measure-valued branching processes. Rozpravy Českosl. Akad. Véd. R. Mat. Přivodnich Véd 76, 3, 1966.

34 Kendall, D.G. Stochastic processes and population growth. J.Roy. Statist. Soc. Ser. B 11, 1949, 230-264.

35 ——— and T. Lewis. On the structural information contained in the output of GI/G/∞. Z. Wahrscheinlichkeitstheorie verw. Geb. 4, 1965, 144-148.

36 Kersten, J. and K. Matthes. Stationäre zufällige Punktfolgen II. J-ber. Deutsch. Math. Verein. 66, 1963, 106-118.

37 ——— and ——— Ergodische unbegrenzt teilbare stationäre zufällige Punktfolgen. Trans. 4th Prague Conf. 1965, 399-415.

38 Khintchine, A. Ya. Sequences of chance events without aftereffect. Th. Prob. Appl. 1, 1956, 1-15.

39 ——— Mathematical Methods in the Theory of Queueing 2nd ed. Griffin, 1969.

40 Kingman, J.F.C. On doubly stochastic Poisson processes. Proc. Camb. Phil. Soc. 60, 1964, 923-932.

41 ——— and S.J. Taylor. Introduction to Measure and Probability Cambridge University Press, 1966.

42 Kolmogorov, A.N. La transformation de Laplace dans les espaces lineáres. C.R. Acad. Sci. Paris. 200, 1935, 1717-1718.

43 Leadbetter, M.R. On three basic results in the theory of stationary point processes. Proc. Am. Math. Soc. 19, 1968, 115-117.

44 Le Cam, L. Un instrument d'étude des fonctions aléatoires : la fonctionnelle charactéristique. C.R. Acad. Sci. Paris 224, 1947, 710-711.

45 Lee, P.M. Infinitely divisible stochastic processes. Z. Wahrscheinlichkeitstheorie verw. Geb. 7, 1967, 147-160.

46 ——— Some examples of infinitely divisible point processes. Stud. Sci. Math. Hung. 3, 1968, 219-224.

47 Leonov, V.P. Applications of the characteristic functional and semi-invariants to the ergodic theory of stationary processes. Dokl. Akad. Nauk SSSR, 1960, 523-526.

48 Leonov, V.P. Some Applications of Higher Cumulants to the Theory of Stationary Stochastic Processes (Russian). 'Nauka' Moscow, 1964.

49 Lewis, P.A.W. A branching Poisson process model for the analysis of computer failure patterns. J.Roy. Statist. Soc. Ser.B 26, 1964, 398-456.

50 ——— Implications of a failure model for the use and maintenance of computers. J. Appl. Probab.1, 1964, 347-368.

51 ——— Asymptotic properties and equilibrium conditions for branching Poisson processes. J. Appl. Probab. 6, 1969, 355-371.

52 Lewis, T. The intervals between regular events displaced in time by independent random deviations of large dispersion. J. Roy. Statist. Soc. Ser. B 23, 1961, 476-483.

53 Lukacs, E. Characteristic Functions. Griffin, 1960.

54 Matthes, K. Unbeschränkt teilbare Verteilungsgesetze stationärer zufälliger Punktfolgen. Wiss. Z. Hochschule Elektro. Ilmenau 9, 1963, 235-238.

55 ——— Stationäre zufällige Punktfolgen I. J-ber. Deutsch.Math. Verein. 66, 1963, 66-79.

56 Mecke, J. Stationäre zufällige Masse auf lokalkompakten Abelschen Gruppen. Z.Wahrscheinlichkeitstheorie verw. Geb. 9, 1967, 36-58.

57 ——— Eine charakteristische Eigenschaft der doppelt stochastichen Poissonschen Prozesse. Z.Wahrscheinlichkeitstheorie verw. Geb. 11, 1968, 74-81.

58 Milne, R.K. Simple proofs of some theorems on point processes. Submitted for publication.

59 Moran, P.A.P. Dams in series with continuous release. J. Appl. Probab. 4, 1967, 380-388.

60 ——— An Introduction to Probability Theory. Oxford University Press, 1968.

61 Moyal, J.E. The general theory of stochastic population processes. Acta Math. 108, 1962, 1-31.

62 ——— Multiplicative population chains. Proc.Roy.Soc. A, 266, 1962, 518-526.

63 Nawrotzki, K. Eine Grenzwersatz für homogene zufällige Punktfolgen. Math. Nachr. 24, 1962, 201-217.

64 Nelsen, R.B. and T. Williams. Randomly delayed appointment streams. Nature 219, 1968, 573-574.

65 Neyman, J. On finiteness of the process of clustering. Sankhyā A 25, 1963, 69-74.

66 ——— and E.L. Scott. A theory of the spatial distribution of galaxies. Astrophys. J. 116, 1952, 144-163.

67 ——— and ——— Statistical approach to the problems of cosmology. J.Roy. Statist. Soc. Ser.B 20, 1958, 1-43.

68 Ososkov, G.A. A limit theorem for flows of similar events. Th. Prob. Appl. 1, 1956, 248-255.

69 Paley, R.E.A.C. and N. Wiener. Notes on the theory and applications of Fourier transforms I, II. Trans. Am. Math. Soc. 35, 1933, 348-355.

70 Prohorov, Yu. V. The method of characteristic functionals. Proc. 4th Berkeley Symp. 2, 1960, 403-419.

71 Rao, C.R. On some characterisations of the Normal Law. Sankhyā A 29, 1967, 1-14.

72 Rao, J.S. An application of stationary point processes to queueing theory and textile research. J. Appl. Probab. 3, 1966, 231-246.

73 Rosenblatt, M. Random Processes. Oxford University Press 1962.

74 ——— and J.W. Van Ness. Estimation of the bispectrum. Ann.Math. Statist. 36, 1965, 1120-1136.

75 Ryll-Nardzewski, C. Remarks on the Poisson stochastic process (III). Studia Math. 14, 1954, 314-318.

76 ——— Remarks on processes of calls. Proc. 4th Berkeley Symp. 2, 1960, 455-465.

77 Shiryaev, A.N. Some problems in the spectral theory of higher order moments I. Th. Prob. Appl. 5, 1960, 265-284.

78 Slivnyak, I.M. Some properties of stationary flows of homogeneous random events. Th. Prob. Appl. 7, 1962, 336-341. (see also 9, 1964, 168).

79 Smith, W.L. Renewal theory and its ramifications J.Roy. Statist. Soc. Ser. B 20, 1958, 243-302.

80 Stone, C. On a theorem of Dobrushin. Ann. Math. Statist. 39, 1968, 1391-1401.

81 Szász, D. On the general branching process with continuous time parameter. Stud. Sci. Math. Hung. 2, 1967, 227-247.

82 Thedéen, T. A note on the Poisson tendency in traffic distribution. Ann. Math. Statist. 35, 1964, 1823-1824.

83 ——— Convergence and invariance questions for point systems in R_1 under random motion. Arkiv för Mat. 7, 1967, 211-239.

84 Thompson, H.R. Spatial point processes, with applications to ecology. Biometrika 42, 1955, 102-115.

85 Titchmarsh, E.C. Introduction to the Theory of Fourier Integrals. 2nd ed. Oxford University Press 1948.

86 Verbitskaya, I.N. On conditions for the strong law of large numbers to be applicable to second-order stationary processes. Th. Prob. Appl. 9, 1964, 325-331.

87 Vere-Jones, D. Some applications of probability generating functionals to the study of input-output streams. J.Roy. Statist. Soc. Ser. B 30, 1968, 321-333.

88 ——— Stochastic models for earthquake occurrence. J.Roy. Statist. Soc. Ser. B. 32 (to appear)

89 Volkonskii, V.A. An ergodic theorem on the distribution of duration of fades. Th. Prob. Appl. 5, 1960, 323-326.

90 Von Wanderfels, W. Charakteristische Funktionale zufällige Masse. Z. Wahrscheinlichkeitstheorie verw. Geb. 10, 1968, 279-283.

91 Westcott, M. Identifiability in linear processes. Submitted for publication.

92 Wold, H. Sur les processus stationnaires ponctuels. Colloques Internationaux du C.N.R.S. 13, 1948, 75-86.

93 Zitek, F. A note on a theorem of Korolyuk (Russian) Czech. Math. J. 7, 1957, 318-319.

94 Zolotarev, V.M. Distribution of the superposition of infinitely divisible processes. Th. Prob. Appl. 3, 1958, 185-188.