J. Appl. Prob. **11**, 493-503 (1974)
Printed In Israel

A CLUSTER PROCESS REPRESENTATION OF A SELF-EXCITING PROCESS

ALAN G. HAWKES*, *University of Durham*
DAVID OAKES**, *Imperial College London*

Abstract

It is shown that all stationary self-exciting point processes with finite intensity may be represented as Poisson cluster processes which are age-dependent immigration-birth processes, and their existence is established. This result is used to derive some counting and interval properties of these processes using the probability generating functional.

CLUSTER PROCESS; POINT PROCESS; SELF-EXCITING PROCESS

1. Introduction

Let $P = \{t_j : j = 0, \pm 1, \pm 2, \cdots\}$ be an orderly point process with no fixed atoms on the real line R. Here we take $\cdots < t_{-1} < 0 \leqq t_0 < t_1 < \cdots$, and it is assumed that the set P has no limit points. The *counting measure* $N(\cdot)$ is defined for each bounded Borel subset A of R to be the cardinal of the set $P \cap A$. It is well known (e.g., Daley and Vere-Jones (1971), Theorem 2.5) that the so-called *finite dimensional distributions*, i.e., the joint distributions of $\{N(I_1), N(I_2), \cdots, N(I_k)\}$ for every integer $k > 0$ and all bounded intervals $I_1, I_2, \cdots, I_k$ in R, characterise P. However other methods of defining point processes are often useful in discussing particular models. For example we may consider the *interval sequence* $\cdots, X_{-1}, Y_{-1}, Y_0, X_1, X_2, \cdots$, where $X_i = t_i - t_{i-1}$ $(i = \pm 1, \pm 2, \cdots)$ and $Y_{-1} = -t_{-1}$, $Y_0 = t_0$. If the joint distribution of $\{Y_{-1}, Y_0, X_1, X_{-1}, \cdots, X_n, X_{-n}\}$ is given for every n then P is completely determined. Here we must have

$$\sum_{-n}^{-1} X_i \to \infty \text{ and } \sum_{1}^{n} X_i \to \infty \text{ as } n \to \infty$$

to ensure that P is well-defined.

Another approach is through the *complete intensity function*. This was discussed in the context of multivariate point processes by Cox and Lewis (1971). The *history* H_t of P at time t is defined by

$$H_t = \{t_j : t_j \in P \text{ and } t_j < t\}.$$

Received in revised form 23 July 1973.
* Now at University College Swarsea.
** Now at Harvard University. Supported by the Science Research Council and N. S. F. grant GS32327X.

We do not exclude the possibility that $H_t = \varnothing$ for some t. If $H_0 = \varnothing$ with probability one then P will be called transient. The complete intensity function $\Lambda(t, H_t)$ is defined by

$$\Lambda(t,H_t) = \lim_{\delta\to 0} \delta^{-1}P[\text{Event in } [t,t+\delta) \,|\, H_t]. \tag{1}$$

In this paper we shall only be concerned with processes for which $\Lambda(t, H_t)$ is well-defined and finite with probability one. A disadvantage of the complete intensity function is that it does not necessarily characterise the process and it is usually difficult to demonstrate the existence of stationary point processes which have a particular complete intensity function. However we do have the following lemma.

Lemma. If H_0 has a specified distribution, then there exists at most one orderly point process P^+ in $t \geqq 0$ which satisfies (1) for a given function $\Lambda(\cdot, \cdot)$.

Proof. $\Lambda(\cdot,\cdot)$ determines the joint interval distributions as follows:

$$P\,[Y_0 \leqq y_0 \,|\, H_0] = 1 - \exp\left[-\int_0^{y_0} \Lambda(u, H_u)du\right],$$

where $H_u = H_0$ for $0 \leqq u \leqq y_0$, and, for $n \geqq 1$,

$$P[X_n \leqq x_n \,|\, H_0, Y_0 = y_0, X_1 = x_1, \cdots, X_{n-1} = x_{n-1}]$$
$$= 1 - \exp\left[-\int_{t_{n-1}}^{t_n} \Lambda(u, H_u)du\right],$$

where $H_u = H_0 \cup \{t_0, \cdots, t_{n-1}\}$ for $t_{n-1} < u \leqq t_n$, and,

$$t_i = y_0 + \sum_{j=1}^{i} x_j, \qquad (0 \leqq i \leqq n).$$

The *intensity* of P is $f(t) = E\{\Lambda(t, H_t)\}$. We assume that this exists. Then (cf. Daley and Vere-Jones (1971), Chapter 6) we have

$$E\{N[a,b)\} = \int_a^b f(u)du. \tag{2}$$

If P is stationary, $f(t) = \lambda$, the (constant) rate of P, and (2) is Korolyuk's theorem.

Hawkes (1971a) has introduced a new class of point processes which he calls *self-exciting processes*. These are formally defined as point processes with a complete intensity function of the form

$$\Lambda(t, H_t) = \nu + \int_{-\infty}^{t} \gamma(t-u)dN(u), \tag{3}$$

where $\nu > 0$, $\gamma(u) \geqq 0$ for $u \geqq 0$, and

(4) $$0 < m = \int_0^\infty \gamma(u)du < 1.$$

In this case the above interval distributions can be written down explicitly and provide in principle a means of predicting future events in terms of the history of the process, see Jowett and Vere-Jones (1971). For example the time to the first event after zero has the conditional distribution

$$P(Y_0 \leqq y_0 \mid H_0) = 1 - \exp\left\{ - \nu y_0 - \int_{-\infty}^{0} [\Gamma(y_0 - u) - \Gamma(-u)]dN(u) \right\}$$

where $\Gamma(v) = \int_0^v \gamma(u)du$. Hawkes (1971a), (1971b) has given the second order counting properties of these and related multivariate processes, on the assumption of stationarity. The purpose of this paper is to demonstrate the existence of stationary self-exciting processes, and to indicate their relation with other point process models. We show that any stationary self-exciting process with finite intensity is a generalised branching Poisson process (Lewis (1969)) — here called a Poisson cluster process. Lewis's results are used to derive some properties of the self-exciting process.

2. A cluster process representation

We first prove that a Poisson cluster process exists satisfying the axioms of a self-exciting process, and then show that there is only one finite intensity orderly process satisfying (3) and (4) for a given function $\gamma(u)$.

Lemma 1. If $\nu > 0$ and $m < 1$ there exists a stationary Poisson cluster process with rate $\lambda = \nu/(1 - m)$ satisfying (3) and (4).

Proof. Kendall (1949) introduced an age-dependent birth and death process such that for any individual of age x alive at time t, then for the next interval $(t, t + dt)$ there are probabilities $\lambda(x)dt$ of a birth and $\mu(x)dt$ of a death (up to order $o(dt)$), independently for each individual. If we take the death rate as zero, change the notation from $\lambda(x)$ to $\gamma(x)$ and in addition allow immigration at a rate ν per unit time, then we may write the complete intensity of the process as

$$\Lambda(t, H_t) = \nu + \sum_{t_i < t} \gamma(t - t_i)$$

where t_i are the instants at which births or immigrations occur (immigrants are assumed to have age zero on arrival). This result may be written equivalently in the form (3) where $N(t)$ is the counting measure corresponding to the times of immigration or birth.

This process may also be considered as a cluster process in which the process of cluster centres $N_c(t)$ is the Poisson process of rate ν formed by the arrival of immigrants. Associated with each event of $N_c(t)$ we have a cluster of subsidiary events formed by the births of all the descendants of all generations of the immigrant. These clusters, which are mutually independent by construction, may be considered as generalised branching processes, see Harris (1963), Chapter III.7.

If we consider one immigration or birth at time $t = 0$ say, then the first generation offspring are generated by a non-stationary ($\gamma(t)$) Poisson process with probability generating functional (p.g.fl.) given by Vere-Jones (1970) as

$$(5) \qquad \begin{aligned} H(z(\cdot)) &= E\left(\exp\left\{\int \log z(t)dN(t)\right\}\right) \\ &= \exp\left\{\int_{-\infty}^{\infty} [z(t)-1]\gamma(t)dt\right\}. \end{aligned}$$

where in this case $\gamma(t) = 0$, for $t < 0$. Vere-Jones also shows that this process is equivalent to a Neyman-Scott cluster process in which the number of events in a cluster has a Poisson distribution with mean m, and the distances of each point of the cluster from the cluster centre are independent random variables with p.d.f. $\gamma(t)/m$. Now our cluster contains all generations so that the total number of events in a cluster, including the original immigrant, is equivalent to the total size of the Galton-Watson process in which the distribution of the number of offspring per generation has a Poisson distribution with mean m. From Harris ((1963) p. 32) we then see that the cluster size has p.g.f. $\pi(z)$ satisfying

$$(6) \qquad \pi(z) = z\exp\{m[\pi(z) - 1]\}.$$

From the Galton-Watson theory we know that the clusters are almost surely finite if $m < 1$, and from (6) we easily find that the mean cluster size is $1/(1-m)$. This ensures the existence of this stationary cluster process with rate $\lambda = \nu/(1-m)$ by applying Theorem 3 of Westcott (1971). We have already seen that this process is also a self-exciting process satisfying (3) and (4). It may be noted that Westcott's results imply that the integral in (3) is almost everywhere finite, with probability one, and is indeed the limit in (1), again with probability one.

It also follows from known results (e.g., Kerstan and Matthes (1965)) that the process is infinitely divisible and mixing.

It is clear that a transient self-exciting process satisfying (3) and (4) can also be constructed by taking the process of cluster centres as a Poisson process for $t > T$ only for some arbitrary constant T.

Lemma 2. There exists at most one stationary orderly point process of finite rate whose complete intensity is given by (3) and (4).

Proof. Let P_1 be any such process of finite rate $\lambda_1 < \infty$. Let $P_1^{(1)}$ be a process identical to P_1 on $(-\infty, 0]$ but with complete intensity

$$\Lambda_1^{(1)}(t, H_{1,t}^{(1)}) = \int_{-\infty}^{t} \gamma(t-u)dN_1^{(1)}(u)$$

on $(0, \infty)$. Consider the process P_1^*, defined as the superposition of $P_1^{(1)}$ and $P_1^{(2)}$, where $P_1^{(2)}$ is a transient immigration-birth process satisfying (3) and (4) starting at time zero and independent of $P_1^{(1)}$. Then P_1^* coincides with P_1 on $(-\infty, 0]$ and has the same complete intensity as P_1 on $(0, \infty)$, so that by the lemma of Section 1 it has the same distributions as P_1 everywhere.

The intensity $f_1^{(1)}(t) = E(\Lambda_1^{(1)}(t, H_{t,t}^{(1)}))$ of $P_1^{(1)}$ satisfies

$$f_1^{(1)}(t) = \begin{cases} \lambda_1 & (t < 0), \\ \lambda_1 \displaystyle\int_{-\infty}^{0} \gamma(t-u)du + \int_0^t \gamma(t-u) f_1^{(1)}(u)du & (t > 0). \end{cases}$$

It follows easily that $f_1^{(1)}(t) \to 0$ as $t \to \infty$. Consequently, for any bounded Borel set B,

$$E(N_1^{(1)}(T_t B)) = \int_{T_t B} f_1^{(1)}(u)du \to 0 \qquad (t \to \infty),$$

where T_t is the translation operator: $x \in T_t B \Leftrightarrow x - t \in B$. Since $N_1^{(1)}(T_t B)$ is a non-negative integer valued random variable this implies that it converges in probability to zero.

Now consider the cylinder sets $C(A, k)$ of the form $\{N(A_i) = k_i,\ 1 \leqq i \leqq r\}$ where the A_i are bounded Borel sets, $k_i \geqq 0$ and $r > 0$. Let $\boldsymbol{P}$ represent the probability measure of such sets when the counting measure N corresponds to a point process P. Then, by the independence of $P_1^{(1)}$ and $P_1^{(2)}$,

$$\begin{aligned} \boldsymbol{P}_1(C(T_t A, \boldsymbol{k})) &= \boldsymbol{P}_1^{(1)}(C(T_t A, \boldsymbol{0}))\boldsymbol{P}_1^{(2)}(C(T_t A, \boldsymbol{k})) \\ &\quad + \sum_{\boldsymbol{l} \neq 0} \boldsymbol{P}_1^{(1)}(C(T_t A, \boldsymbol{l}))\boldsymbol{P}_1^{(2)}(C(T_t A, \boldsymbol{k} - \boldsymbol{l})). \end{aligned}$$

The previous result with $B = UA_i$ implies that each term in the (finite) summation tends to zero as $t \to \infty$, while $\boldsymbol{P}_1^{(1)}(C(T_t A, \boldsymbol{0})) \to 1$. Also, since P_1 is stationary the left-hand side of this equation does not depend on t, so that as $t \to \infty$,

$$\boldsymbol{P}_1^{(2)}(C(T_t A, \boldsymbol{k})) \to \boldsymbol{P}_1(C(A, \boldsymbol{k})).$$

To prove uniqueness we note that a process P_2 of the type constructed in Lemma 1 has a similar decomposition into components $P_2^{(1)}$ and $P_2^{(2)}$ and that

$$\boldsymbol{P}_2^{(2)}(C(T_t A, \boldsymbol{k})) \to P_2(C(A, \boldsymbol{k})).$$

But $P_1^{(2)}$ and $P_2^{(2)}$ have the same structure and so also the same distributions. Hence

$$P_1(C(A, \boldsymbol{k})) = P_2(C(A, \boldsymbol{k}))$$

for all cylinder sets, completing the proof.

Lemmas 1 and 2 combined make the following theorem.

Theorem 1. If $\nu > 0$ and the non-negative function $\gamma(u)$ satisfies

$$0 < m = \int_0^\infty \gamma(u)du < 1,$$

then there is precisely one stationary orderly process of finite rate whose complete intensity function satisfies (3). This process may be represented as a Poisson cluster process which is an age-dependent immigration-birth process having rate $\nu/(1-m)$. The limiting distributions of the corresponding transient process starting at zero exist and are the same as the distributions of the stationary process.

The existence and uniqueness are proved in Lemmas 1 and 2. The result on the limiting distributions of the transient process $P^{(2)}$ follows immediately from the proof of Lemma 2. Similar results can be obtained for sets of mutually exciting processes as defined by Hawkes (1971a, b).

3. Some counting and interval properties

The representation of a self-exciting process as a cluster process enables us to obtain an equation for the probability generating functional of the process. We use this together with results of Lewis (1969) on Poisson cluster processes to obtain some properties of the process. As far as possible we retain his notation.

Theorem 2. The p.g.fl. of the self-exciting process has the form

$$G(z(\cdot)) = \exp\left\{\int_{-\infty}^{\infty} \nu[F(z_t(\cdot)) - 1]dt\right\}, \tag{7}$$

where $F(z(\cdot))$ is the p.g.fl. of a cluster generated by an immigrant arriving at time zero, and including that immigrant, while $z_t(\cdot) = z(t+\cdot)$ is simply the translation of $z(\cdot)$. The functional $F(z(\cdot))$ satisfies the functional equation

$$F(z(\cdot)) = z(0)\exp\left\{\int_0^\infty [F(z_t(\cdot)) - 1]\gamma(t)dt\right\}. \tag{8}$$

Proof. Vere-Jones (1970) gives the p.g.fl. of a cluster process as

$$G(z(\cdot)) = G_0(F(z(\cdot)\,|\,t)), \tag{9}$$

where $G_0(z(\cdot))$ is the p.g.fl. of the process of cluster centres and $F(z(\cdot)\,|\,t)$ is the p.g.fl. for a cluster given that the cluster centre occurs at time t. If the cluster process is time homogeneous we may write $F(z(\cdot)\,|\,t) = F(z_t(\cdot))$ where $F(z(\cdot))$ corresponds to a cluster centre at time zero. In this case we have as a special case of (5)

$$G_0(z(\cdot)) = \exp\left\{\int_{-\infty}^{\infty} \nu[z(t)-1]dt\right\}. \tag{10}$$

To investigate the clusters we define $F_n(z(\cdot))$ to be the p.g.fl. of the process which consists of all births in all generations up to and including the nth generation descended from one individual who originates the cluster at $t = 0$, and is included as part of the cluster. Then treating the first generation as a process of cluster centres each of which generates further sub-clusters we use (7) to obtain a backward equation for the branching process

$$\begin{aligned} F_n(z(\cdot)) &= z(0)H\{F_{n-1}(z_t(\cdot))\} \\ &= z(0)\exp\left\{\int_0^{\infty} [F_{n-1}(z_t(\cdot))-1]\gamma(t)dt\right\}, \end{aligned} \tag{11}$$

with $F_0(z(\cdot)) = z(0)$ representing the original ancestor. In the limit as $n \to \infty$ we include the whole cluster consisting of all generations of the family tree, and the p.g.fl. satisfies Equation (8). This together with (9) and (10) completes the proof.

Note that one could obtain Equation (8) by a transformation and limiting operation on a special case of Equation (26) of Bartlett and Kendall (1951), which is concerned with a characteristic functional for the age structure at time t of the birth and death process. The same operations applied to Equation (8) of Bartlett ((1955), p. 109) lead to our Equation (11) which may be regarded as a formal iterative solution of (8). No explicit solutions are known but useful equations may be obtained by choosing particular functions $z(\cdot)$. For example if we let $z(x) = z$ for all x then $F(z(\cdot)) = \pi(z)$ is the p.g.f. for the total size of the cluster and (8) reduces to (6) which we have already used. If we take $z(x) = z$ for $y \leqq x \leqq y + l$ and $z(x) = 1$ elsewhere then $F(z_t(\cdot)) = \pi(y - t, l, z)$ say is the p.g.f. for the number of events in the interval $[y - t,\ y - t + l]$ for a cluster whose originating event occurs at time zero. Then (8) becomes

$$\begin{aligned} \pi(y,l,z) &= \exp\left\{\int_0^{y+l}[\pi(y-t,l,z)-1]\gamma(t)dt\right\}, && y > 0, \\ &= z\exp\left\{\int_0^{y+l}[\pi(y-t,l,z)-1]\gamma(t)dt\right\}, && -l \leqq y \leqq 0, \\ &= 1\ , && y < -l, \end{aligned} \tag{12}$$

the last result arises because if the interval precedes the cluster centre there can be no events of the cluster in it. Now let $z(x) = z$ for $0 \leqq x \leqq l$ and $z(x) = 1$ elsewhere in Equation (7). Then $G(z(\cdot)) = Q_l(z)$, say, is the p.g.f. for the number of events in $[0, l]$ for the equilibrium self-exciting process and (7) becomes

$$Q_l(z) = \exp\left\{\int_{-\infty}^{l} \nu[\pi(-t,l,z)-1]dt\right\}. \tag{13}$$

In principle Equations (12) and (13) determine the distribution of counts but are rather intractable. However we establish an asymptotic result for counts in Theorem 4 and meanwhile use a special case of these equations to establish some interval properties of the process.

Theorem 3. (a) The forward recurrence time L of the equilibrium self-exciting process has survivor function $R_L(l) = P(L > l)$ given by

$$R_L(l) = \exp\left\{ -\nu l - \nu \int_0^\infty [1 - \phi_J(t, l)]dt \right\}, \tag{14}$$

where $\phi_J(y, l)$ satisfies the equation

$$\begin{aligned} \phi_J(y, l) &= \exp\left\{ \int_0^{y+l} [\phi_J(y - t, l) - 1]\gamma(t)dt \right\}, && y > 0, \\ &= 0, && -l \leqq y \leqq 0, \\ &= 1, && y < -l. \end{aligned} \tag{15}$$

(b) The equilibrium distribution of T, the interval between successive events, has survivor function

$$R_T(t) = P(T > t) = -[dR_L(t)/dt]/\lambda, \tag{16}$$

where λ is the process rate.

Proof. This follows simply by taking $z = 0$ in Equations (12) and (13) so that $R_L(l) = Q_l(0) = P$ (no events in $[0, l]$), and $\phi_J(y, l) = \pi(y, l, 0)$ is the probability of no events in $[y, y + l]$ for a cluster originating at zero. Part (b) is a consequence of the well-known form for the p.d.f. of the recurrence time $f_L(t) = R_T(t)/E(T)$, see Equation 4.4.1 of Lewis (1964), and the fact that $E(T) = 1/\lambda$.

Equation (14) is equivalent to the limit of Equation 4.3.2 of Lewis (1964). Equation (15) may be solved in principle by repeated numerical integration using the recurrence relation (11), which in this case takes the form

$$\begin{aligned} \phi_{J,n}(y, l) &= \exp\left\{ \int_0^{y+l} [\phi_{J,n-1}(y - t, l) - 1]\gamma(t)dt \right\}, && y > 0, \\ &= 0, && -l \leqq y \leqq 0, \\ &= 1, && y < -l, \end{aligned}$$

and with the initial function

$$\begin{aligned} \phi_{J,0}(y, l) &= 0 \text{ if } -l \leqq y \leqq 0, \\ &= 1 \text{ otherwise.} \end{aligned}$$

Returning now to study further the structure of a cluster we take $z(x) = z$ for

$0 \leqq x \leqq u$ and 1 otherwise, and find that $F(z(\cdot)) = \pi_u(z)$ say is the p.g.f. for the number of events in $[0,u]$ of a cluster originating at time zero. Equation (8) becomes

$$\pi_u(z) = z\exp\left\{\int_0^u [\pi_{u-t}(z) - 1]\gamma(t)dt\right\}, \qquad u \geqq 0.$$

If $S(u)$ is the number of subsidiary events in $[0,u]$, i.e., excluding the originating event, if $H(u) = E(S(u))$ we have

$$1 + H(u) = [d\pi_u/dz]_{z=1} = 1 + \int_0^u [1 + H(u-t)]\gamma(t)dt.$$

The Laplace transform $H^*(s)$ of $H(u)$ is easily obtained from this as

$$H^*(s) = \gamma^*(s)/[s(1-\gamma^*(s))]. \tag{17}$$

Lewis points out that in general $H(u)/H(\infty)$ has the form of a distribution function. In this case we see from (17) that this corresponds to a random variable representable as the sum of K independent variables having p.d.f. $\gamma(u)/m$, where K has a geometric distribution $P(K = k) = (1-m)m^{k-1}$, $k = 1,2,\cdots$. This result may be obtained alternatively by using the fact that the expected number of kth generation descendants is m^k. The above result enables us to prove the asymptotic normality of the counting process in certain circumstances.

Theorem 4. If $\int_0^\infty u\gamma(u)du < \infty$ then the asymptotic distribution of $N(0,t]$, the number of events in $(0,t]$ for the transient or for the stationary self-exciting process, is given by

$$P\left\{\frac{N(0,t] - vt/(1-m)}{[vt/(1-m)^3]^{\frac{1}{2}}}\right\} \leqq y \to \Phi(y) \text{ as } t \to \infty,$$

where $\Phi(y)$ is the distribution function for the unit normal distribution.

Proof. This follows from Theorem 3.3 of Lewis (1969) substituting $E(S+1) = 1/(1-m)$, $E((S+1)^2) = 1/(1-m)^3$ which is easily obtained from (6). $1 + S = 1 + S(\infty)$ is the total cluster size. Sufficient conditions for the validity of this theorem are $\int_0^\infty [E(s) - H(u)]du < \infty$. This is assured in this case because from (17) we find that

$$\int_0^\infty [E(s) - H(u)]du = \mu m/(1-m)^2$$

where $\mu m = \int_0^\infty u\gamma(u)du$, so that μ represents the mean of the random variable with p.d.f. $\gamma(u)/m$.

Notice that this result is consistent with the asymptotic slope of the variance time curve

$$\operatorname{Var}\{N(0,t]\}/t \to v/(1-m)^3 \text{ as } t \to \infty$$

which is given by the ordinate at the origin of the point spectrum given by Equation (15) of Hawkes (1971b).

Finally we find some results for the length J of a cluster, i.e., the time between the first and last events of the cluster.

Theorem 5. The distribution function $D_J(y) = P(J \leqq y)$ of the length of a cluster satisfies

$$(18) \qquad \begin{aligned} D_J(y) &= \exp\left\{ -m + \int_0^y D_J(y-t)\gamma(t)dt \right\}, \qquad & y \geqq 0, \\ &= 0, & y < 0. \end{aligned}$$

Proof. This follows once again from (8) by taking $z(x) = 1$ for $x \leqq y$ and 0 for $x > y$. Then $F(z(\cdot))$ is the probability of no events in a cluster after time y, i.e., $P(J \leqq y)$, because

$$F(z(\cdot)) = E\left(\exp\left\{\int_0^\infty \log z(t) dN(t)\right\}\right)$$

and the random variable $\exp\{\int_0^\infty \log z(t)dN(t)\}$ takes the value 1 if there are no events after t and the value 0 otherwise. This equation may also be solved in principle by repeated numerical integration as (11) now takes the form

$$D_{J,n}(y) = \exp\left\{ -m + \int_0^y D_{J,n-1}(y-t)\gamma(t)dt \right\}, \qquad y \geqq 0,$$

with initially $D_{J,0}(y) = 1, y \geqq 0$, and 0 otherwise.

We conclude by noting three small points. Firstly $D_J(0) = e^{-m}$ is the probability of no subsidiary events. Secondly it is easy to see that

$$\mu m e^{-2m} = \mu P(S=1) \leqq E(J) \leqq \mu E(S) = \mu m/(1-m)$$

and it follows that $E(J) < \infty$ if and only if $\mu < \infty$. Thirdly in the case $E(J) < \infty$ it follows from Theorem 2.1 of Lewis (1969) that for the equilibrium process the number of cluster processes running at any time has a Poisson distribution with mean $vE(J)$.

Acknowledgements

D.O. is grateful to his supervisor, Professor D. R. Cox, and to Dr. M. Westcott, for helpful discussion. Both authors would like to thank the referee for helpful suggestions which have improved the presentation, particularly with respect to Lemma 2. The joint paper arose after some of the results had been derived independently by the two authors at about the same time.

References

[1] Bartlett, M. S. (1955) *An Introduction to Stochastic Processes.* Cambridge University Press.

[2] Bartlett, M. S. and Kendall, D. G. (1951) On the use of the characteristic functional in the analysis of some stochastic processes occurring in physics and biology. *Proc. Camb. Phil. Soc.* **47**, 65–67.

[3] Cox, D. R. and Lewis, P. A. W. (1971) Multivariate point processes. *Proc. Sixth. Berk. Symp. on Math. Statist. and Prob.*

[4] Daley, D. J. and Vere-Jones, D. (1971) A summary of the theory of point processes. Paper presented to the stochastic point process conference, New York, 1971.

[5] Harris, T. E. (1963) *The Theory of Branching Processes.* Springer-Verlag, Berlin.

[6] Hawkes, A. G. (1971a) Spectra of some self-exciting and mutually exciting point processes. *Biometrika* **58**, 83–90.

[7] Hawkes, A. G. (1971b) Point spectra of some mutually exciting point processes. *J. R. Statist. Soc.* B **33**, 438–443.

[8] Jowett, J. and Vere-Jones, D. (1971) The prediction of stationary point processes. Paper presented to the stochastic point process conference, New York, 1971.

[9] Kendall, D. G. (1949) Stochastic processes and population growth. *J. R. Statist. Soc.* B **11**, 230–264.

[10] Kerstan, J. and Matthes, K. (1965) Ergodische unbegrenzt teilbare stationare zufallige Punktfolgen. *Trans. Fourth Prague Conference on Information Theory, Statistical Decision Functions and Random Processes* 399–415. Czechoslovakia Academy of Sciences, Prague.

[11] Lewis, P. A. W. (1964) A branching Poisson process model for the analysis of computer failure patterns. *J. R. Statist. Soc.* B **26**, 398–456.

[12] Lewis, P. A. W. (1969) Asymptotic properties and equilibrium conditions for branching Poisson processes. *J. Appl. Prob.* **6.** 355–371.

[13] Vere-Jones, D. (1970) Stochastic models for earthquake occurrence. *J. R. Statist. Soc.* B **32**, 1–62.

[14] Westcott, M. (1971) On existence and mixing results for cluster point processes. *J. R. Statist. Soc.* B **33**, 290–300.