COINCIDENCE PROPERTIES OF BIRTH AND DEATH PROCESSES

SAMUEL KARLIN AND JAMES MCGREGOR

A birth and death process (for brevity referred to henceforth as process B) is a stationary Markov process whose state space is the non-negative integers and whose transition probability matrix

$$(1) \qquad P_{ij}(t) = \Pr\{x(t) = j \mid x(0) = i\}$$

satisfies the conditions (as $t \to 0$)

$$(2) \qquad P_{ij}(t) = \begin{cases} \lambda_i t + o(t) & \text{if } j = i+1 \\ \mu_i t + o(t) & \text{if } j = i-1 \\ 1 - (\lambda_i + \mu_i)t + o(t) & \text{if } j = i \end{cases}$$

where $\lambda_i > 0$ for $i \geq 0$, $\mu_i > 0$ for $i \geq 1$ and $\mu_0 \geq 0$. We further assume that $P_{ij}(t)$ satisfies the foward and backward equation in the usual form. In this paper we restrict attention to the case $\mu_0 = 0$ so that when the particle enters the state zero it remains there a random length of time according to an exponential distribution with parameter λ_0 and then moves into state one etc.

In order to avoid inessential difficulties we assume henceforth that the infinitesimal birth and death rates λ_i and μ_i uniquely determine the process. This is equivalent to the condition $\sum_{n=0}^{\infty}(\pi_n + 1/\lambda_n\pi_n) = \infty$ where

$$\pi_0 = 1 \text{ and } \pi_n = \frac{\lambda_0\lambda_1\lambda_2 \cdots \lambda_{n-1}}{\mu_1\mu_2\mu_3 \cdots \mu_n} \qquad [2].$$

In the companion paper we show that for all $t > 0$

$$(3) \qquad \det(P_{i_\mu, j_\nu}(t)) = P\begin{pmatrix} t\,; & \begin{matrix} i_1, i_2, \cdots, i_n \\ j_1, j_2, \cdots, j_n \end{matrix} \end{pmatrix} \qquad \begin{matrix} i_1 < i_2 < i_3 < \cdots < i_n \\ j_1 < j_2 < j_3 < \cdots < j_n \end{matrix}$$

has the following interpretation: Start n labled particles at time zero in states $i_1, i_2, \cdots, i_n$ respectively, each governed by the transition law (1) and acting independently. The determinant (3) is equal to the probability that at time t particle 1 is located in state j_1, particle 2 is located in state j_2 etc., without any two of these particles having occupied simultaneously a common state at some earlier time $\tau < t$. We refer to this event as a transition in time t of n particles from initial states

Received December 18, 1958. This work was supported in part by an Office of Naval Research Contract at Stanford University.

occur. This problem is completely solved in § 4. By means of trivial arguments it is shown that coincidence is certain if the original birth and death process is recurrent, while coincidence is not certain if the original process is strongly transient. If the original process is weakly transient coincidence may or may not be certain, and this case presents a much more difficult problem. A criterion is given which expresses the necessary and sufficient condition that coincidence be certain, in terms of the constants of the original birth and death process. Finally in § 3 some interesting examples are considered. A technique for computing the distribution of the time until coincidence is developed, and applied to the telephone trunking model and some linear growth models.

1. Positivity properties of $Q\begin{pmatrix} i_1, i_2, \cdots, i_n \\ x_1, x_2, \cdots, x_n \end{pmatrix}$.

Let M, K and L be functions of two variables satisfying

$$M(\xi, \eta) = \int_a^b K(\xi, \zeta)L(\zeta, \eta)d\sigma(\zeta) \tag{8}$$

where ξ traverses X, ζ ranges through Y and η varies over Z all of which are linearly ordered sets and where $\sigma(\zeta)$ denotes a measure defined in Y. X can denote an interval of the real line or a set of discrete points on the line. In the latter case, the set will usually consist of the integers. The same applies to Y and Z. When Y consists of a discrete space then, of course, the integral sign of (8) is interpreted as a sum. We define the Fredholm determinant

$$M\begin{pmatrix} x_1, x_2, \cdots, x_n \\ z_1, z_2, \cdots, z_n \end{pmatrix} = \begin{vmatrix} M(x_1, z_1), & M(x_1, z_2), & \cdots, & M(x_1, z_n) \\ M(x_2, z_1), & M(x_2, z_2), & \cdots, & M(x_2, z_n) \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ M(x_n, z_1), & M(x_n, z_2), & \cdots, & M(x_n, z_n) \end{vmatrix} \tag{9}$$

with $x_1 < x_2 < \cdots < x_n$ and $z_1 < z_2 < \cdots < z_n$ and analogously for K and L.

If the formula (8) is viewed as a continuous version of a matrix product, then the extension of the multiplication rule which evaluates subdeterminants of M in terms those of K and L becomes

$$M\begin{pmatrix} x_1, x_2, \cdots, x_n \\ z_1, z_2, \cdots, z_n \end{pmatrix} = \int_{a \le y_1 < y_2 < \cdots < y_n \le b} \cdots \int K\begin{pmatrix} x_1, x_2, \cdots, x_n \\ y_1, y_2, \cdots, y_n \end{pmatrix} L\begin{pmatrix} y_1, y_2, \cdots, y_n \\ z_1, z_2, \cdots, z_n \end{pmatrix} . \\ d\sigma(y_1)d\sigma(y_2) \cdots d\sigma(y_n) . \tag{10}$$

For the proof of (10) we refer to Pólya and Szëgo I [8 p. 48]

$$(6)\qquad Q\begin{pmatrix} i_1, i_2, \cdots, i_n \\ x_1, x_2, \cdots, x_n \end{pmatrix} = \begin{vmatrix} Q_{i_1}(x_1), Q_{i_1}(x_2), \cdots, Q_{i_1}(x_n) \\ Q_{i_2}(x_1), Q_{i_2}(x_2), \cdots, Q_{i_2}(x_n) \\ \cdots \\ Q_{i_n}(x_1), Q_{i_n}(x_2), \cdots, Q_{i_n}(x_n) \end{vmatrix}$$

where $i_1 < i_2 < \cdots < i_n$ and $x_1 < x_2 < \cdots < x_n$ we obtain by virtue of (5) that

$$(7)\quad P\left(t\,;\, \begin{matrix} i_1, i_2, \cdots, i_n \\ j_1, j_2, \cdots, j_n \end{matrix}\right)$$
$$= \pi_{j_1}\pi_{j_2}\cdots\pi_{j_n} \int\cdots\int_{x_1<x_2<\cdots<x_n} e^{-(x_1+x_2+\cdots+x_n)t} Q\begin{pmatrix} i_1, i_2, \cdots, i_n \\ x_1, x_2, \cdots, x_n \end{pmatrix} Q\begin{pmatrix} j_1, j_2, \cdots, j_n \\ x_1, x_2, \cdots, x_n \end{pmatrix} \cdot$$
$$\cdot\, d\psi(x_1)d\psi(x_2)\cdots d\psi(x_n)\,.$$

(See Paragraph A of Section 1.)

The above formula displays in the simplest possible way the dependence of $P\left(t\,;\, \begin{matrix} i_1, \cdots, i_n \\ j_1, \cdots, j_n \end{matrix}\right)$ on the time t, the initial state $(i_1, \cdots, i_n)$ and final state $(j_1, \cdots, j_n)$. For the birth and death process itself formula (5) has proven to be a very powerful tool in analyzing the statistical properties of the process. It may be expected that formula (7) will be of comparable utility in the study of the compound process. However certain technical details stand in the way of such a study. While the general properties of the orthogonal polynomials $\{Q_n(x)\}$ have been intensively investigated by numerous mathematicians, the somewhat more complicated polynomials $\left\{Q\begin{pmatrix} i_1, \cdots, i_n \\ x_{1i} \cdots, x_n \end{pmatrix}\right\}$ appear to be new objects of study. At the present time we possess numerous interesting theorems about these polynomials but our results are still incomplete. In a separate publication we will elaborate on the structure of this determinantal polynomial system. In the present paper we develop only those properties directly relevant to our analysis.

We investigate two types of problems associated with the compound process. The first problem is concerned with the behavior of the ratio

$$\frac{P\left(t\,;\, \begin{matrix} i_1, \cdots, i_n \\ j_1, \cdots, j_n \end{matrix}\right)}{P\left(t\,;\, \begin{matrix} k_1, \cdots, k_n \\ l_1, \cdots, l_n \end{matrix}\right)}$$

as $t \to \infty$. This requires some knowledge of positivity properties of the polynomials $Q\begin{pmatrix} i_1, \cdots, i_n \\ x_1, \cdots, x_n \end{pmatrix}$. In § 1 these required positivity properties are developed, and in § 2 it is shown that the above ratio converges to a finite positive limit as $t \to \infty$. The second problem is that of determining for which processes coincidence in a finite state is certain to

$i_1, i_2, \cdots, i_n$ to the states $j_1, j_2, \cdots, j_n$ respectively, without coincidence. In particular, for $t > 0$ the expression (3) is always positive. For continuous time discrete state space processes, the converse proposition is also true. Specifically, if (3) is always positive, then $P_{ij}(t)$ is the transition matrix of a birth and death process [6].

In this paper, we investigate certain aspects of the structure of the Markov process describing the transitions of n particles conditioned that no coincidence takes place.

We refer to this process as the compound birth and death process of order n. Frequently, when no ambiguities arise the terms "birth and death" and "order n" will be suppressed. The basis of the subsequent analysis is principally an integral representation for

$$P\left(t;\ \begin{matrix} i_1, i_2, \cdots, i_n \\ j_1, j_2, \cdots, j_n \end{matrix}\right)$$

which is derived from a corresponding representation formula for $P_{ij}(t)$.

Let $Q_n(x)$ denote a sequence of polynomials of degree n defined by the recursive relations

$$\begin{aligned} &-xQ_n(x) = -(\lambda_n + \mu_n)Q_n(x) + \lambda_n Q_{n+1}(x) + \mu_n Q_{n-1}(x) \qquad n \geq 0 \\ &Q_0(x) \equiv 1 \qquad\qquad Q_{-1}(x) \equiv 0\ . \end{aligned} \tag{4}$$

These equations may be written in compact form as

$$-xQ = AQ$$

where Q represents the vector $(Q_0(x), Q_1(x), Q_2(x), \cdots,)$ and A is the infinitesimal matrix of the process

$$A = \left\| \begin{matrix} -\lambda_0 & \lambda_0 & & & & \\ \mu_1 & -(\lambda_1+\mu_1) & \lambda_1 & & & \\ & \mu_2 & -(\lambda_2+\mu_2) & \lambda_2 & & \\ & & \mu_3 & -(\lambda_3+\mu_3) & \lambda_3 & \\ & & & \cdot & \cdot & \cdot \\ & & & & \cdot & \cdot & \cdot \\ & & & & & \cdot & \cdot & \cdot \end{matrix} \right\|$$

Let $\psi(x)$ denote the unique measure on $[0, \infty)$ with respect to which $Q_n(x)$ are orthogonal. (The measure ψ is unique because of the assumption $\sum (\pi_k + 1/\lambda_k\pi_k) = \infty$.) Then

$$P_{nm}(t) = \pi_m \int_0^\infty e^{-xt} Q_n(x) Q_m(x) d\psi(x)\ . \tag{5}$$

Introducing the notation

The relevance and utility of this identity will be abundantly clear. We record several relations which are applications of it.

(A) The derivation of (7) from (5) is a special case of (10).

(B) The identity

$$\sum_{j=0}^{n} \pi_j Q_j(x) = \frac{\lambda_n \pi_n [Q_{n+1}(x) - Q_n(x)]}{-x} = H_n(x) \tag{11}$$

can be expressed in the form (8) with $\xi = n, \zeta = j$, and $\eta = x$

$$M(\xi, \eta) = H_n(x)$$
$$K(\xi, \zeta) = \begin{cases} 1 & \zeta \leq \xi \\ 0 & \zeta > \xi \end{cases}$$
$$L(\zeta, \eta) = Q_j(x)$$

($Q_j(x) \equiv 0$ for j a negative integer and $d\sigma(\zeta) = \pi_j$ when $\zeta = j$.)

Since

$$K\begin{pmatrix} i_1, i_2, \cdots, i_n \\ l_1, l_2, \cdots, l_n \end{pmatrix} = 0$$

unless $0 \leq l_1 \leq i_1, i_1 < l_2 \leq i_2, \cdots, i_{n-1} < l_n \leq i_n$, in which case its value is one, we obtain by applying (10) to (11)

$$\begin{aligned} &H\begin{pmatrix} i_1, i_2, \cdots, i_n \\ x_1, x_2, \cdots, x_n \end{pmatrix} \\ &= \sum_{l_1=0}^{i_1} \sum_{l_2=i_1+1}^{i_2} \cdots \sum_{l_n=i_{n-1}+1}^{i_n} \pi_{l_1}\pi_{l_2} \cdots \pi_{l_n} Q\begin{pmatrix} l_1, l_2, \cdots, l_n \\ x_1, x_2, \cdots, x_n \end{pmatrix} \end{aligned} \tag{12}$$

(C) We shall need to evaluate determinants of the form

$$\begin{vmatrix} Q_{n_0}(0), & Q'_{n_0}(0), & \cdots, & Q^{(k)}_{n_0}(0) \\ Q_{n_1}(0), & Q'_{n_1}(0), & \cdots, & Q^{(k)}_{n_1}(0) \\ \vdots & \vdots & & \vdots \\ Q_{n_k}(0), & Q'_{n_k}(0), & \cdots, & Q^{(k)}_{n_k}(0) \end{vmatrix}, \quad n_0 < n_1 < \cdots < n_k \tag{13}$$

which for convenience of writing we give the name $\alpha(n_0, n_1, \cdots, n_k)$.

We assume tentatively in what follows that Q_n are normalized such that $Q_n(0) = 1$. This can be accomplished with no loss of generality since $Q_n(0)$ are different from zero. The value of the determinant $\alpha(n_0, n_1, \cdots, n_k)$ in the general situation would be altered by the multiplicative factor $1/Q_{n_0}(0)Q_{n_1}(0) \cdots Q_{n_k}(0)$.

A more convenient expression for (13) is obtained as follows: By subtracting the first row from each of the succeeding rows and using the fact that $Q_n(0) = 1$ for all n we have

$$\alpha(n_0, n_1, n_2, \cdots, n_k) = \begin{vmatrix} Q'_{n_1}(0) - Q'_{n_0}(0) & \cdots & Q^{(k)}_{n_1}(0) - Q^{(k)}_{n_0}(0) \\ Q'_{n_2}(0) - Q'_{n_0}(0) & \cdots & Q^{(k)}_{n_2}(0) - Q^{(k)}_{n_0}(0) \\ \vdots & & \vdots \\ Q'_{n_k}(0) - Q'_{n_0}(0) & \cdots & Q^{(k)}_{n_k}(0) - Q^{(k)}_{n_0}(0) \end{vmatrix} .$$

We next observe that relation (11) provided with successive differentiation yields

$$(14) \quad Q^{(r+1)}_{n+1}(0) - Q^{(r+1)}_{n_0}(0) = -(r+1) \sum_{\nu=n_0+1}^{n} \frac{1}{\lambda_\nu \pi_\nu} \sum_{\mu=0}^{\nu} \pi_\mu Q^{(r)}_\mu(0)$$
$$(n > n_0,\ r = 0, 1, 2, \cdots,) .$$

In order to apply (10) to (14), we may identify

$$M(\xi, \eta) = Q^{(r+1)}_{n+1}(0) - Q^{(r+1)}_{n_0}(0)$$
$$K(\xi, \zeta) = \begin{cases} 1 & \zeta \leq \xi \\ 0 & \zeta > \xi \end{cases}$$
$$L(\zeta, \eta) = \sum_{\mu=0}^{\nu} \pi_\mu Q^{(r)}_\mu(0)$$

and $d\sigma(\zeta) = 1/\lambda_\nu \pi_\nu$ where ξ, ζ, η each traverse the set of non-negative integers. By virtue of (10) utilizing the representation (14) we obtain

$$(15) \quad \alpha(n_0, n_1, n_2, \cdots, n_k)$$
$$= (-1)^k (k!) \sum_{l_1=n_0+1}^{n_1} \sum_{l_2=n_1+1}^{n_2} \cdots \sum_{l_k=n_{k-1}+1}^{n_k} \frac{1}{\lambda_{l_1}\pi_{l_1}} \frac{1}{\lambda_{l_2}\pi_{l_2}} \cdots \frac{1}{\lambda_{l_k}\pi_{l_k}} L\begin{pmatrix} l_1, l_2, \cdots, l_k \\ 0, 1, \cdots, k-1 \end{pmatrix}$$

where we have employed the specific evaluations of the Fredholm subdeterminants based on $K(\xi, \eta)$.

Another application of (10) shows that

$$(16) \quad L\begin{pmatrix} l_1, l_2, \cdots, l_k \\ 0, 1, \cdots, k-1 \end{pmatrix}$$
$$= \sum_{\mu_0=0}^{l_1} \sum_{\mu_1=l_1+1}^{l_2} \cdots \sum_{\mu_{k-1}=l_{k-1}+1}^{l_k} \pi_{\mu_0}\pi_{\mu_1} \cdots \pi_{\mu_{k-1}} \alpha(\mu_0, \mu_1, \cdots, \mu_{k-1}) .$$

Putting (15) and (16) together establishes the recursive relation

$$(17) \quad \alpha(n_0, n_1, \cdots, n_k) = (-1)^k k! \sum_{l_1=n_0+1}^{n_1} \cdot \sum_{l_2=n_1+1}^{n_2} \cdots \sum_{l_k=n_{k-1}+1}^{n_k} \frac{1}{\lambda_{l_1}\pi_{l_1}} \frac{1}{\lambda_{l_2}\pi_{l_2}}$$
$$\cdots \frac{1}{\lambda_{l_k}\pi_{l_k}} \sum_{\mu_0=0}^{l_1} \cdot \sum_{\mu_1=l_1+1}^{l_2} \cdots \sum_{\mu_{k-1}=l_{k-1}+1}^{l_k} \pi_{\mu_0}\pi_{\mu_1} \cdots \pi_{\mu_{k-1}} \alpha(\mu_0, \mu_1, \mu_2, \cdots, \mu_{k-1}) .$$

Note that the range of summations guarantee that $\mu_0 < \mu_1 < \mu_2 < \cdots < \mu_{k-1}$.

Furthermore, (17) exhibits determinants $\alpha(n_0, n_1, \cdots, n_k)$ of order $k+1$ in terms of corresponding determinants of order k. Consequently, the procedure may be iterated out of which follows whenever $Q_n(0) > 0$ that

$$(18) \qquad (-1)^{k(k+1)/2} \cdot \alpha(n_0, n_1, n_2, \cdots, n_k) > 0$$

for all choices of n_i provided $n_0 < n_1 < n_2 < \cdots < n_k$. It is also routine to calculate the explicit value of $\alpha(n_0, n_1, n_2 < \cdots < n_k)$ by iteration of (17).

In particular

$$\alpha(n_0, n_1) = -\sum_{k=n_0+1}^{n_1} \frac{1}{\lambda_k \pi_k} \sum_{l=0}^{k} \pi_l$$

$$\alpha(n_0, n_1, n_2) = \sum_{k=n_0+1}^{n_1} \frac{1}{\lambda_k \pi_k} \sum_{l=n_1+1}^{n2} \frac{1}{\lambda_l \pi_l} \sum_{\mu_0=0}^{k} \pi_{\mu_0} \sum_{\mu_1=k+1}^{l} \pi_{\mu_1} \alpha(\mu_0, \mu_1) .$$

The derivation of the identity (17) was predicated upon the fact that $Q_n(0) = 1$ for all n. If all the $Q_n(0)$ are of negative sign then the sign of (13) is altered by the factor $(-1)^{k+1}$ where $k+1$ is the order of the matrix. Indeed, all we need do is replace $Q_n(x)$ by $Q_n(x)/Q_n(0) = P_n(x)$ and apply the argument to $P_n(x)$. The value of $\alpha(n_0, n_1, \cdots, n_k)$ based on $P_n(x)$ differs only by an obvious multiplying factor from that based on $Q_n(x)$.

(D) Following the same lines of argument as above we shall show

$$(19) \qquad (-1)^{(k-1)k/2} Q\begin{pmatrix} n_1, n_2, \cdots, n_k \\ x_1, x_2, \cdots, x_k \end{pmatrix} > 0$$

provided $x_1 < x_2 < x_3 < \cdots < x_k \le a$ where a denote the smallest value in the spectrum of ψ, and where $Q_n(0) > 0$ by our normalization condition. The result expressed in (19) may be regarded as a generalization to the compounded polynomial system of the property that $Q_n(x)$ for $x < a$ is of one sign.

Suppose for definiteness that the polynomials $Q_n(x)$ are orthogonal functions with respect to a measure ψ on $[0, \infty)$. The proof is by induction on k. Since $Q_n(x)$ are normalized to be positive at 0, it follows that $Q_n(x) > 0$ for all $x \le a$ which is the assertion of (19) when $k = 1$. We shall assume that the validity of (19) for kth order determinants has been demonstrated for any system of orthogonal polynomials whose weight function concentrates on the interval $[0, \infty)$, and proceed to show the result is valid for the $k+1$st order determinants. Let $x_1, x_2, \cdots, x_{k+1}$ denote a set of values arranged in increasing order with $x_{k+1} \le a$. Replacing $Q_n(x)$ by $Q_n(x + x_{k+1})/Q_n(x_{k+1})$ we may, without loss of generality assume $x_{k+1} = 0$ and that $Q_n(x_{k+1}) = 1$ for all n. This alters the original determinants by a positive multiplicative factor, provided we evaluate the changed matrix polynomial system at the points $y_i = x_i - x_{k+1}$. Hence

$$Q\binom{n_1, n_2, \cdots, n_{k+1}}{x_1, x_2, \cdots, x_{k+1}} = \begin{vmatrix} Q_{n_1}(x_1), & Q_{n_1}(x_2) & \cdots & Q_{n_1}(x_k) & 1 \\ Q_{n_2}(x_1), & Q_{n_2}(x_2) & \cdots & Q_{n_2}(x_k) & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ Q_{n_{k+1}}(x_1), & Q_{n_{k+1}}(x_2) & \cdots & Q_{n_{k+1}}(x_k) & 1 \end{vmatrix}$$

Subtracting the kth row from the $k+1$ row, the k-1th row from the kth row, etc., and finally the first from the second row we have

$$(20) \qquad Q\binom{n_1, n_2, \cdots, n_{k+1}}{x_1, x_2, \cdots, x_{k+1}}$$
$$= (-1)^k \begin{vmatrix} Q_{n_2}(x_1) - Q_{n_1}(x_1), & Q_{n_2}(x_2) - Q_{n_1}(x_2), & \cdots, & Q_{n_2}(x_k) - Q_{n_1}(x_k) \\ \vdots & & & \\ Q_{n_{k+1}}(x_1) - Q_{n_k}(x_1), & \cdots & & Q_{n_{k+1}}(x_k) - Q_{n_k}(x_k) \end{vmatrix}.$$

Observe that

$$(21) \qquad Q_{n_l}(x) - Q_{n_{l-1}}(x) = -x \sum_{i=n_{l-1}}^{n_l - 1} \frac{1}{\lambda_i \pi_i} H_i(x)$$

where

$$H_i(x) = \frac{\lambda_i \pi_i [Q_{i+1}(x) - Q_i(x)]}{-x}$$

comprise an orthogonal system of polynomials with respect to the measure $x d\psi/\lambda_0$ which concentrates its measure on $(0, \infty)$ since ψ does [2, p 504]. Therefore,

$$(-1)^{k(k-1)/2} H\binom{m_1, m_2, \cdots, m_k}{x_1, x_2, \cdots, x_k} > 0$$

whenever $x_1 < x_2 < x_3 < \cdots < x_k \le 0$ and $m_1 < m_2 < \cdots < m_k$ since $H_i(0) = \sum_{j=0}^{i} \pi_j Q_j(0) > 0$. Inserting (21) into (20) shows that $(-1)^k Q\binom{n_1, n_2, \cdots, n_{k+1}}{x_1, x_2, \cdots, x_{k+1}}$ can be written as

$$(-1)^k x_1 x_2 \cdots x_k \sum_{\mu' s} \gamma_{\mu_1, \mu_2, \cdots, \mu_k} H\binom{\mu_1, \mu_2, \cdots, \mu_k}{x_1, x_2, \cdots, x_k}$$

where the μ's traverse the sets $n_j \le \mu_j \le n_{j+1} - 1$, $j = 1, 2, \cdots, k$ respectively, and $\gamma_{\mu_1, \cdots, \mu_k} > 0$. Taking account of the inequality $x_j < 0$, $j = 1, 2, \cdots, k$ and the induction hypothesis which insures the inequality $(-1)^{k(k-1)/2} H\binom{\mu_1, \cdots, \mu_k}{x_1, \cdots, x_k} > 0$ we obtain

$$(-1)^k (-1)^{k(k-1)/2} Q\binom{n_1, n_2, \cdots, n_{k+1}}{x_1, x_2, \cdots, x_{k+1}} > 0$$

as asserted. This completes the proof.

A little manipulation of (19) will show that

$$(22) \qquad (-1)^{k(k-1)/2} \begin{vmatrix} Q_{n_1}(\gamma) & Q'_{n_1}(\gamma) & \cdots & Q_{n_1}^{(k-1)}(\gamma) \\ Q_{n_2}(\gamma) & & \cdots & \\ \vdots & & & \\ Q_{n_k}(\gamma) & Q'_{n_k}(\gamma) & \cdots & Q_{n_k}^{(k-1)}(\gamma) \end{vmatrix} \geq 0$$

true for every $\gamma \leq a$. This is verified by subtracting the last column of (19) from the next to last and using the mean value theorem. Repeating this k times and afterwards letting all the x_i converge to γ produces (22). Subject to the correct normalization the argument employed in paragraph (C) above shows that these determinants are actually strictly positive.

A further sharpening of the relation (19) and (22) is possible. In order to describe this extension we must assign a special meaning to the

determinant $$Q^*\binom{n_1, n_2, \cdots, n_k}{x_1, x_2, \cdots, x_k}$$

where $n_1 < n_2 < \cdots < n_k$ and $x_1 \leq x_2 \leq \cdots \leq x_k$ and distinguished in that several of the x's can be equal. (The asterisk sign on the Q shall always occur when one or more of the x's are equal and indicates that a special interpretation is to be made.) Let us illustrate by means of an example.

If $x_1 < x_2 = x_3 < x_4 = x_5 = x_6$ then

$$Q^*\binom{n_1, n_2, \cdots, n_6}{x_1, x_2, \cdots, x_6} = \begin{vmatrix} Q_{n_1}(x_1) & Q_{n_1}(x_2) & Q'_{n_1}(x_2) & Q_{n_1}(x_4) & Q'_{n_1}(x_4) & Q''_{n_1}(x_4) \\ Q_{n_2}(x_1) & Q_{n_2}(x_2) & Q'_{n_2}(x_2) & Q_{n_2}(x_4) & Q'_{n_2}(x_4) & Q''_{n_2}(x_4) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ Q_{n_6}(x_1) & Q_{n_6}(x_2) & Q'_{n_6}(x_2) & Q_{n_6}(x_4) & Q'_{n_6}(x_4) & Q''_{n_6}(x_4) \end{vmatrix}$$

In general, when there is a block of equal x values present, the successive columns, corresponding to these x values in forming Q^* are determined by the successive derivatives, *i.e.* (Q_n), (Q'_n), (Q''_n), $\cdots$, $(Q_n^{(r-1)})$ where r is the number of equal x values.

One can show by a more tedious elaboration of the methods in (C) and (D) that generally

$$(23) \qquad (-1)^{k(k-1)/2} Q^*\binom{n_1, n_2, \cdots, n_k}{x_1, x_2, \cdots, x_k} > 0$$

when $x_1 \leq x_2 \leq \cdots \leq x_k \leq a$ with the emphasis on strict inequality in (23).

We do not indicate the details since an analogous argument will be used in the proof of Theorem 1.

(E) With the aid of the results of (C) and (D) we shall deduce determinantal inequalities valid for special choices of positive x's. Let $Q_n(x)$ be a system of orthogonal polynomials normalized as usual so that $Q_n(0)>0$ and ψ its measure on $[0, \infty)$. Let us suppose the measure ψ begins with isolated jumps located at $a_1 < a_2 < \cdots < a_r$ followed by a non-isolated point in the spectrum starting at a_{r+1} where r may be $0, 1, 2, \cdots$. In particular, when $r = 0$ then the first point in the spectrum of ψ is not an isolated jump. On the other extreme if $r = \infty$ then the first portion of the spectrum ψ consists of an infinite number of isolated jumps which could include the full spectrum. It is not necessary, in what follows, to describe more precisely the spectrum beyond a_{r+1}.

THEOREM 1. *Let $0 \leq n_1 < n_2 < n_3 < \cdots < n_k$ and Q_n be normalized as usual such that $Q_n(0) > 0$; then for $k \leq r$,*

$$(-1)^{k(k-1)/2} Q\begin{pmatrix} n_1, n_2, \cdots, n_k \\ a_1, a_2, \cdots, a_k \end{pmatrix} > 0 \tag{24}$$

and for $k > r$,

$$(-1)^{k(k-1)/2} Q^*\begin{pmatrix} n_1, n_2, \cdots, n_r, n_{r+1}, n_{r+2}, \cdots, n_k \\ a_1, a_2, \cdots, a_r, a_{r+1}, a_{r+1}, \cdots, a_{r+1} \end{pmatrix} > 0 \tag{25}$$

where Q^* is defined as above.

Proof. The proof is by induction on the order of the determinant k. The case where $r = 0$ has already been completely examined in paragraph (C). Hence, we assume $r \geq 1$. We suppose furthermore that the theorem has been established with regard to any orthogonal polynomial system whose spectral measure concentrates on the non-negative axis with the number of initial isolated jumps totalling less than r. Let r be fixed and ≥ 1 and suppose we have established the theorem for determinants of size $< k$. Denote by $P_n(x) = Q_n(x + a_1)/Q_n(a_1)$. These polynomials constitute an orthogonal system with respect to the measure $\psi(x + a_1)$ whose first mass points occur at $0, a_2 - a_1, a_3 - a_1, \cdots, a_r - a_1$. Observe that

$$P^*\begin{pmatrix} n_1, n_2, \cdots, n_k \\ b_1, b_2, \cdots, b_k \end{pmatrix} = Q^*\begin{pmatrix} n_1, n_2, \cdots, n_k \\ a_1, a_2, \cdots, a_k \end{pmatrix} C(n_1, n_2, \cdots, n_k)$$

where $C(n_1, n_2, \cdots, n_k) > 0$ and

$$b_1 = 0, b_2 = a_2 - a_1, \cdots, b_r = a_r - a_1, b_j = a_{r+1} - a_1 \quad \text{for} \quad j \geq r + 1 .$$

Subtracting the k-1th row from the kth row, the k-2th row from the k-1th row etc., we obtain

$$P^*\binom{n_1, n_2, \cdots, n_k}{b_1, b_2, \cdots, b_k}$$
$$= \begin{vmatrix} P_{n_2}(b_2) - P_{n_1}(b_2), & \cdots, P_{n_2}(b_{r+1}) - P_{n_1}(b_{r+1}), & P'_{n_2}(b_{r+1}) - P'_{n_1}(b_{r+1}) & \cdots \\ \vdots & \vdots & \vdots & \\ P_{n_k}(b_2) - P_{n_{k-1}}(b_2), & \cdots, P_{n_k}(b_{r+1}) - P_{n_{k-1}}(b_{r+1}), & P'_{n_k}(b_{r+1}) - P'_{n_{k-1}}(b_{r+1}) & \cdots \end{vmatrix}$$

The right-hand side is a determinant of size $k - 1$. Dividing the respective columns by $- 1/b_\nu$, $\nu = 2, 3, \cdots$, (remember $b_2 > 0$ since $r \geq 1$), we have

$$\text{(26)} \qquad (-1)^{k-1} \operatorname{sign} P^*\binom{n_1, n_2, \cdots, n_k}{b_1, b_2, \cdots, b_k}$$
$$= \operatorname{sign} \begin{vmatrix} \dfrac{P_{n_2}(b_2) - P_{n_1}(b_2)}{-b_2}, \cdots, \dfrac{P_{n_2}(b_{r+1}) - P_{n_1}(b_{r+1})}{-b_{r+1}}, \dfrac{P'_{n_2}(b_{r+1}) - P'_{n_1}(b_{r+1})}{-b_{r+1}} \cdots \\ \vdots \\ \dfrac{P_{n_k}(b_2) - P_{n_{k-1}}(b_2)}{-b_2}, \cdots, \dfrac{P_{n_k}(b_{r+1}) - P_{n_{k-1}}(b_{r+1})}{-b_{r+1}}, \dfrac{P'_{n_k}(b_{r+1}) - P'_{n_{k-1}}(b_{r+1})}{-b_{r+1}} \cdots \end{vmatrix}.$$

Let

$$H_r(x) = \lambda_r^* \pi_r^* \left[\frac{P_{r+1}(x) - P_r(x)}{-x}\right]$$

and set

$$M_r^{(l)}(x) = \frac{\lambda_r^* \pi_r^* [P_{r+1}^{(l)}(x) - P_r^{(l)}(x)]}{-x}$$

$l = 0, 1, 2, \cdots$, so that $M_r^{(0)} = H_r$ and λ_r^* and μ_r^* are the parameters corresponding to the polynomial system $P_n(x)$. Finally, for $0 \leq \mu_1 < \mu_2 < \cdots < \mu_{k-1}$ define

$$L\binom{\mu_1, \cdots, \mu_{k-1}}{b_2, \cdots, b_k} = \begin{vmatrix} H_{\mu_1}(b_2) & \cdots H_{\mu_1}(b_{r+1}), & M'_{\mu_1}(b_{r+1}) & \cdots M_{\mu_1}^{(k-1-r)}(b_{r+1}) \\ \vdots & & & \\ H_{\mu_{k-1}}(b_2) & \cdots H_{\mu_{k-1}}(b_{r+1}), & M'_{\mu_{k-1}}(b_{r+1}) & \cdots M_{\mu_{k-1}}^{(k-1-r)}(b_{r+1}) \end{vmatrix}.$$

Expanding the right-hand side of (26), using (21) and an analogous formula for the successive derivatives of $Q_n(x)$, we obtain

$$(27)\quad \operatorname{sign}(-1)^{k-1}P^*\binom{n_1, n_2, \cdots, n_k}{b_1, b_2, \cdots, b_k} = \operatorname{sign}\sum_{\mu}\gamma_{\mu_1,\mu_2,\cdots,\mu_{k-1}}L\binom{\mu_1, \mu_2, \cdots, \mu_{k-1}}{b_2, b_3, \cdots, b_k},$$

where the γ's are positive and $n_i \leq \mu_i < n_{i+1}$, $(i = 1, \cdots, k-1)$. But

$$H_r^{(l)}(x) = M_r^{(l)}(x) + \sum_{t=0}^{l-1} c_t(x)M_r^{(t)}(x)$$

Hence by suitable operations on the columns of L we obtain

$$L\binom{\mu_1, \cdots, \mu_{k-1}}{b_2, \cdots, b_k} = H^*\binom{\mu_1, \mu_2, \cdots, \mu_{k-1}}{b_2, b_3, \cdots, b_k}$$

where the H^* determinant is formed from the polynomial system H_n in the same way that Q^* is constructed in terms of Q_n.

The H system represent orthogonal polynomials with respect to a measure $d\alpha(x) = C\cdot xd\psi(x + a_1)$. The jump at the origin of $d\psi(x + a_1)$ is obliterated due to the factor x. Otherwise, α possesses $r - 1$ initial jumps located at $a_2 - a_1, \cdots, a_r - a_1$ and the non-isolated portion of the spectrum begins at the point $a_{r+1} - a_1$. By the induction hypothesis $(-1)^{(k-1)(k-2)/2}H^*\binom{\mu_1, \mu_2, \cdots, \mu_{k-1}}{b_2, b_3, \cdots, b_k} > 0$. This fact in conjunction with (27) shows that

$$(-1)^{k(k-1)/2}P^*\binom{n_1, n_2, \cdots, n_k}{b_1, b_2, \cdots, b_k} > 0$$

as desired. The proof of the theorem is complete.

What is essential for the validity of (25) is that the first r choices of $y_i > 0$ used in evaluating (25) should coincide with the first spectral points a_i of ψ (here r has the same meaning as in the theorem). Otherwise the values of y_j $(j \geq r+1)$ can be arbitrarily chosen from the interval $a_r < y \leq a_{r+1}$ with the restriction that they are arranged in ascending order even allowing equalities. Actually, more is true. A careful examination of the above arguments shows that

$$(28)\quad (-1)^{k(k-1)/2}Q^*\binom{n_1, n_2, \cdots, n_s, n_{s+1}, \cdots, n_k}{a_1, a_2, \cdots, a_s, y_{s+1}, \cdots, y_k} > 0 \ (s \leq r)$$

where n_i strictly increase and y_j for $j \geq s+1$ satisfy $a_s < y_{s+1} \leq y_{s+2} \leq \cdots \leq y_k \leq a_{s+1}$.

To complete the story we note without proof that it is possible to construct examples which show that $Q\binom{n, n+1}{x, \ y}$ does not possess a fixed sign for all n when x and y satisfy $x < a_1$ and $a_1 < y < a_2$.

2. The compound process. The infinite matrix $P(t)$ satisfies the differential equations

$$\frac{dP(t)}{dt} = AP(t) ,$$

$$\frac{dP(t)}{dt} = P(t)A ,$$

called respectively the backward and the forward equations of the birth and death process. Either equation may be derived from the other when it is known that both $P(t)$ and A satisfy the symmetry relations

$$P_{ij}(t)\pi_i = P_{ji}(t)\pi_j ,\ a_{ij}\pi_i = a_{ji}\pi_j .$$

As a consequence of these equations we deduce the backward and forward equations of the compound process:

$$\begin{aligned}\frac{d}{dt}P\left(t\,;\begin{matrix} i_1, \cdots, i_n \\ j_1, \cdots, j_n\end{matrix}\right) &= \sum_{r=1}^{n}\left\{\mu_{i_r}P\left(t\,;\begin{matrix} i_i, \cdots, i_{r-1}, i_r - 1, i_{r+1}, \cdots, i_n \\ j_1, \cdots, j_{r-1}, j_r, j_{r+1}, \cdots, j_n\end{matrix}\right)\right. \\ &\quad \left. - (\lambda_{ir} + \mu_{ir})P\left(t\,;\begin{matrix} i_1, \cdots, i_n \\ j_1, \cdots, j_n\end{matrix}\right) + \lambda_{i_r}P\left(t\,;\begin{matrix} i_1, \cdots, i_{r-1}, i_r + 1, i_{r+1}, \cdots, i_n \\ j_1, \cdots, j_{r-1}, j_r, j_{r+1}, \cdots, j_n\end{matrix}\right)\right\}\end{aligned}$$

(29)

$$\begin{aligned}\frac{d}{dt}P\left(t\,;\begin{matrix} i_1, \cdots, j_n \\ j_1, \cdots, j_n\end{matrix}\right) &= \sum_{r=1}^{n}\left\{\lambda_{j_r - 1}P\left(t\,;\begin{matrix} i_1, \cdots, i_{r-1}, i_r, i_{r+1}, \cdots, i_n \\ j_1, \cdots, j_{r-1}, j_r - 1, j_{r+1}, \cdots, j_n\end{matrix}\right)\right. \\ &\quad \left. - (\lambda_{j_r} + \mu_{j_r})P\left(t\,;\begin{matrix} i_1, \cdots, i_n \\ j_1, \cdots, j_n\end{matrix}\right) + \mu_{j_r + 1}P\left(t\,;\begin{matrix} i_1, \cdots, i_{r-1}, i_r, i_{r+1}, \cdots, i_n \\ j_1, \cdots, j_{r-1}, j_r + 1, j_{r+1}, \cdots, j_n\end{matrix}\right)\right\} .\end{aligned}$$

Here we employ the natural convention that $P\left(t\,;\begin{matrix} i_1, \cdots, i_n \\ j_1, \cdots, j_n\end{matrix}\right)$ for $i_1 \leq \cdots \leq i_n$ and $j_1 \leq \cdots \leq j_n$ is zero if any two i_ν or any two j_ν are equal or if $i_1 = -1$ or $j_1 = -1$. The first of the above equations (backward equation) follows at once from

$$\begin{aligned}&\frac{d}{dt}P\left(t\,;\begin{matrix} i_1, \cdots, i_n \\ j_1, \cdots, j_n\end{matrix}\right) \\ &= \sum\nolimits_\sigma (\operatorname{sign} \sigma)P_{i_1 j_{\sigma_1}}(t) \cdots P_{i_{r-1} j_{\sigma_{r-1}}}(t)\left\{\frac{d}{dt}P_{i_r j_{\sigma_r}}(t)\right\}P_{i_{r+1} j_{\sigma_{r+1}}}(t) \cdots P_{i_n j_{\sigma_n}}(t)\end{aligned}$$

on applying the backward equation, $P'(t) = AP(t)$. Here $\sum_\sigma$ denotes summation over all permutations $\sigma = \begin{pmatrix} 1, \cdots, n \\ \sigma_1, \cdots, \sigma_n\end{pmatrix}$ of $1, 2, \cdots, n$. The forward equation may be obtained in a similar way from the forward equation of the original process. Alternatively either of the two equations is a consequence of the other one together with the symmetry relations

$$P\left(t\,;\begin{matrix} i_1, \cdots, i_n \\ j_1, \cdots, j_n \end{matrix}\right)\pi_{i_1}\cdots\pi_{i_n} = P\left(t\,;\begin{matrix} j_1, \cdots, j_n \\ i_1, \cdots, i_n \end{matrix}\right)\pi_{j_1}\cdots\pi_{j_n}\,,$$

and $\lambda_r\pi_r = \mu_{r+1}\pi_{r+1}$.

The backward and forward equations of the compound process may also be derived from the representation (7) and the fact the determinantal polynomials $Q\begin{pmatrix} i_1, \cdots, i_n \\ x_1, \cdots, x_n \end{pmatrix}$ satisfy the recurrence formula

$$\begin{aligned} -(x_1+\cdots+x_n)Q\begin{pmatrix} i_1, \cdots, i_n \\ x_1, \cdots, x_n \end{pmatrix} = \sum_{r=1}^{n}\Bigg[\mu_r Q\begin{pmatrix} i_1, \cdots, i_{r-1}, i_r-1, i_r, \cdots, i_n \\ x_1, \cdots\cdots\cdots\cdots\cdots\cdots, x_n \end{pmatrix} \\ -(\lambda_r+\mu_r)Q\begin{pmatrix} i_1, \cdots, i_n \\ x_1, \cdots, x_n \end{pmatrix} + \lambda_r Q\begin{pmatrix} i_1, \cdots, i_{r-1}, i_r+1, i_{r+1}, \cdots, i_n \\ x_1, \cdots\cdots\cdots\cdots\cdots\cdots, x_n \end{pmatrix}\Bigg] \end{aligned} \tag{30}$$

where $Q\begin{pmatrix} j_1, \cdots, j_n \\ x_1, \cdots, x_n \end{pmatrix}$ for $j_1 \leq \cdots \leq j_n$ is taken to be zero if any two j_ν are the same or if $j_1 = -1$. This recurrence formula follows at once by applying the basic recurrence formula $-xQ(x) = AQ(x)$ to the right member of the identity

$$-(x_1+\cdots+x_n)Q\begin{pmatrix} i_1, \cdots, i_n \\ x_1, \cdots, x_n \end{pmatrix}$$

$$= \sum_{r=1}^{n}\sum_\sigma (\operatorname{sign}\sigma)Q_{i_1}(x_{\sigma_1})\cdots Q_{i_{r-1}}(x_{\sigma_{r-1}})[-x_{\sigma_r}Q_{i_r}(x_{\sigma_r})]\cdots Q_{i_n}(x_{\sigma_n})\,.$$

It is not difficult to see that $P\left(t\,;\begin{matrix} i_1, \cdots, i_n \\ j_1, \cdots, j_n \end{matrix}\right)$ converges to zero as $t \to \infty$. In fact if the original birth and death process is either transient or recurrent null then $P_{ij}(t) \to 0$ for each i and j so the determinant $\to 0$. On the other hand if the original birth and death process is recurrent (either ergodic or recurrent null) and $F_{i0}(t)$ is the probability that first passage from state i to state 0 occurs in time $\leq t$ then $F_{i,0}(t) \to 1$ and from probabilistic considerations

$$P\left(t\,;\begin{matrix} i_1, \cdots, i_n \\ j_1, \cdots, j_n \end{matrix}\right) \leq 1 - F_{i_n,0}(t) \to 0 \text{ as } t \to \infty\,.$$

Thus we have two reasons why the determinants may $\to 0$ and at least one of them is always in force.

According to the Doeblin-Chung ratio theorem [1]

$$\lim_{t\to\infty}\frac{\int_0^t P\left(\tau\,;\begin{matrix} i_1, \cdots, i_n \\ j_1, \cdots, j_n \end{matrix}\right)d\tau}{\int_0^t P\left(\tau\,;\begin{matrix} k_1, \cdots, k_n \\ l_1, \cdots, l_n \end{matrix}\right)d\tau}$$

exists and is finite and positive. For the compound process of the birth and death process we are able to make the following considerably sharper statement.

THEOREM 2.

$$\lim_{t\to\infty} \frac{P\left(t\,;\begin{matrix} i_1, \cdots, i_n \\ j_1, \cdots, j_n \end{matrix}\right)}{P\left(t\,;\begin{matrix} k_1, \cdots, k_n \\ l_1, \cdots, l_n \end{matrix}\right)}$$

exists and is finite and positive.

Proof. It is evidently sufficient to consider the case when $(k_1, \cdots, k_n) = (l_1, \cdots, l_n) = (0, 1, \cdots, n-1)$. Let $f(x_1, \cdots, x_n)$ be the polynomial such that

$$Q\begin{pmatrix} i_1, \cdots, i_n \\ x_1, \cdots, x_n \end{pmatrix} Q\begin{pmatrix} j_1, \cdots, j_n \\ x_1, \cdots, x_n \end{pmatrix} = f(x_1, \cdots, x_n)\left[Q\begin{pmatrix} 0, \cdots, n-1 \\ x_1, \cdots, x_n \end{pmatrix}\right]^2 .$$

We wish to show that

$$\text{(31)} \quad \frac{\displaystyle\int\cdots\int_{0\le x_1<\cdots<x_n} e^{-(x_1+\cdots+x_n)t} f(x_1, \cdots, x_n)\left[Q\begin{pmatrix} 0, \cdots, n-1 \\ x_1, \cdots, x_n \end{pmatrix}\right]^2 d\psi(x_1)\cdots d\psi(x_n)}{\displaystyle\int\cdots\int_{0\le x_1<\cdots<x_n} e^{-(x_1+\cdots+x_n)t} Q\left[\begin{pmatrix} 0, \cdots, n-1 \\ x_1, \cdots, x_n \end{pmatrix}\right]^2 d\psi(x_1)\cdots d\psi(x_n)}$$

converges to a finite positive limit as $t \to \infty$. Suppose there are x values $0 \le a_1 < \cdots < a_{r+1}$ such that the function $\psi(x)$ has positive jumps at $a_1, \cdots, a_r$ but no other spectrum in $0 \le x < a_{r+1}$ while ψ has infinitely many points of increase in every interval $a_{r+1} < x < a_{r+1} + \varepsilon$. We consider separately the cases $r \ge n$, $1 \le r < n$ and $r = 0$. The case $1 \le r < n$, which exhibits all the necessary arguments, will be discussed in detail and the other two cases are left as an exercise for the interested reader. When $1 \le r < n$ integrals of the form

$$\int\cdots\int_{0\le x_1<\cdots<x_n} F(x_1, \cdots, x_n) d\psi(x_1)\cdots d\psi(x_n)$$

may be written in the form

$$\int\cdots\int_{\substack{a_1\le x_1\le\cdots\le x_n \\ x_2 \ge a_2 \\ \vdots \\ x_{r+1} \ge a_{r+1}}} F(x_1, \cdots, x_n) d\psi(x_1)\cdots d\psi(x_n) .$$

For large t the main contributions to the integrals in (31) therefore come from the neighborhood of the point $(x_1, \cdots, x_n) = (a_r, \cdots, a_r, a_{r+1}, \cdots, a_{r+1})$. To make this precise we first observe that Theorem 1 shows that $f(a_1, \cdots, a_r, a_{r+1}, \cdots, a_{r+1}) = c$ is positive and that the measure

$$\left[Q\binom{0, \cdots, n-1}{x_1, \cdots, x_n}\right]^2 d\psi(x_1) \cdots d\psi(x_n) = d\theta(x_1, \cdots, x_n)$$

has positive mass in every "right-hand" neighborhood of the point $(a_1, \cdots, a_r, a_{r+1}, \cdots, a_{r+1})$. The expression (31) can be written in the form $c + (I_1/I_2)$ where

$$I_1 = \int \cdots \int e^{-(x_1+\cdots+x_n)t+(a_1+\cdots+a_r+a_{r+1}+\cdots+a_{r+1})t} \cdot [f(x_1, \cdots, x_n) - c] d\theta(x_1, \cdots, x_n) ,$$

$$I_2 = \int \cdots \int e^{-(x_1+\cdots+x_n)t+(a_1+\cdots a_r+a_{r+1}+\cdots+a_{r+1})t} d\theta(x_1, \cdots, x_n) .$$

Given $\varepsilon > 0$ we choose $\delta > 0$ so $|f(x_1, \cdots, x_n) - c| < \varepsilon$ for $|x_1 - a_1| + \cdots + |x_r - a_r| + |x_{r+1} - a_{r+1}| + \cdots + |x_n - a_{r+1}| \leq \delta$. Let R_δ and R'_δ denote the parts of the region $0 \leq x_1 < \cdots < x_n$ where

$x_1 + \cdots + x_n \leq a_1 + \cdots + a_r + (n-r)a_{r+1} + \delta$ and where
$x_1 + \cdots + x_n > a_1 + \cdots + a_r + (n-r)a_{r+1} + \delta$ respectively.

Then

$$\begin{aligned} |I_1| &= \left| \int_{R_\delta} \cdots \int + \int_{R'_\delta} \cdots \int \right| \\ &\leq \varepsilon \int \cdots \int e^{-(x_1+\cdots+x_n)t+(a_1+\cdots)t} d\theta(x_1, \cdots, x_n) \\ &\quad + e^{-\delta t} \int_{R'_\delta} \cdots \int |f(x_1, \cdots, x_n) - c| \, d\theta(x_1, \cdots, x_n) \end{aligned}$$

while

$$\begin{aligned} |I_2| &\geq \int_{R_\delta} \cdots \int e^{-(x_1+\cdots)t+(a_1+\cdots)t} d\theta(x_1, \cdots, x_n) \\ &\geq \int_{R_{\frac{1}{2}\delta}} \cdots \int e^{-(x_1+\cdots)t+(a_1+\cdots)t} d\theta(x_1, \cdots, x_n) \geq Be^{-\frac{1}{2}\delta t} \end{aligned}$$

where $B > 0$. Consequently $\limsup_{t\to\infty} |(I_1)/(I_2)| \leq \varepsilon$ and the theorem follows.

3. Some examples of the probability distribution of the time until coincidence. A random variable of natural interest to the study of the compound process of order n is the time t^* until coincidence. To expedite the discussion we restrict attention to the case of the compound process involving two particles. The obvious extensions are left to the reader. In general, coincidence need not occur with certainty. We define t^* to be the time of first coincidence if this is finite and to be $+\infty$ otherwise. In the next section the condition that coincidence be a certain event is expressed in terms of the parameters of the birth and death process. In this section the explicit distribution of t^* is determined for some important examples.

We begin with a few remarks concerning the general character of this problem. We may consider a two-dimensional birth and death process whose states are all pairs (i, j) with $i \geq 0$, $j \geq 0$ and transition probability law

$$P_{ij;kl} = P_{ik}(t)P_{jl}(t) .$$

In this formulation the problem is to determine the distribution of the time of first hitting the diagonal ray $i = j$.

Alternatively, we may consider the compound process with state space (i, j), $0 \leq i < j$ and transition probability law $P\left(t\,;\begin{smallmatrix} i, & j \\ k, & l\end{smallmatrix}\right)$. In this formulation, coincidence occurs if the particle is in some state $(k, k+1)$ and is then absorbed—the process terminates at $(k, k+1)$. The problem is then to determine the distribution of the time until the process terminates in this manner.

Let $S^{ij}(t)$, $(0 \leq i < j)$ denote the probability distribution of the time until coincidence when the initial states of the particles are respectively i and j; *i.e.*

$$S^{ij}(t) = \Pr\{t^* \leq t \mid x(0) = i,\ y(0) = j,\ i < j\}$$

Because the path functions are continuous (a particle moving from state i to state j in time t must occupy all the intermediate states in the intervening time), coincidence can only occur following a transition from a state $(k, k+1)$ for some k. More exactly, the probability that coincidence happens during the time interval $[t, t+h]$ with h sufficiently small requires that the two particles occupy adjacent states before coincidence at time t and at the next transition the particles meet. The probability of this event is clearly

$$\sum_{k=0}^{\infty} P\left(t\,;\begin{matrix} i, & j \\ k, & k+1\end{matrix}\right)(\lambda_k + \mu_{k+1})h + o(h)$$

and the density function of the time until coincidence is

$$R^{ij}(t) = \frac{dS^{ij}}{dt}(t) = \sum_{k=0}^{\infty} (\lambda_k + \mu_{k+1}) P\left(t\,;\begin{matrix} i, & j \\ k, & k+1\end{matrix}\right) .$$

The method we use to compute $S^{ij}(t)$ consists of determining explicitly the generating function

$$G(z,\ w) = \sum_{0 \le i < j} R^{ij}(t)(z^i w^j - z^j w^i)$$

or sometimes more conveniently

$$H(z,\ w) = \sum_{0 \le i < j} \pi_i \pi_j R^{ij}(t)(z^i w^j - z^j w^i)$$

and then reading off the coefficient of $z^i w^j$, $(i < j)$.

If we have available

$$f_k(z,\ t) = \sum_{l=0}^{\infty} P_{kl}(t) z^l \tag{32}$$

and hence

$$\pi_k f_k(z,\ t) = \sum_{l=0}^{\infty} \pi_l P_{lk}(t) z^k$$

we obtain employing (10) the determinantal identity

$$\begin{aligned} M(k,\ z,\ w,\ t) &= \pi_k \pi_{k+1} \begin{vmatrix} f_k(z,\ t) & f_{k+1}(z,\ t) \\ f_k(w,\ t) & f_{k+1}(w,\ t) \end{vmatrix} \\ &= \sum_{0 \le l_1 < l_2} \pi_{l_1} \pi_{l_2} P\left(t\ ;\ \begin{matrix} l_1, & l_2 \\ k, & k+1 \end{matrix}\right)(w^{l_2} z^{l_1} - w^{l_1} z^{l_2}) \end{aligned}$$

where $z < w$.

Direct summation gives

$$\sum_{k=0}^{\infty} (\lambda_k + \mu_{k+1}) M(k,\ z,\ w,\ t) = \sum_{0 \le l_1 < l_2} \pi_{l_1} \pi_{l_2} R^{l_1 l_2}(t)[w^{l_2} z^{l_1} - w^{l_1} z^{l_2}]\ . \tag{33}$$

In many cases it is possible to recognize the left-hand side of (33) in terms of classical functions and then obtain $R^{l_1 l_2}(t)$ by picking out the proper coefficient in the series expansion. We record several important examples.

EXAMPLE 1. Consider the telephone trunking model $(\lambda_n = \lambda,\ \mu_n = n\mu,\ n \ge 0)$ [4]. The orthogonal polynomials are the Poisson Charlier polynomials. The generating function of the transition probabilities is known to be

$$f_k(z,\ t) = e^{-a(1-z)(1-e^{-\mu t})}[1 - (1 - z)e^{-\mu t}]^k = \alpha_t(z)[\beta_t(z)]^k$$

where $a = \lambda/\mu$ and $\alpha_t(z)$ and $\beta_t(z)$ are defined in the obvious fashion. The preceding calculations in this case yield

$$\sum_{0 \le l_1 < l_2} \pi_{l_1}\pi_{l_2} R^{l_1 \cdot l_2}(t)[w^{l_2}z^{l_1} - w^{l_1}z^{l_2}] \tag{34}$$
$$= a(w - z)e^{-\mu t}\alpha_t(z)\alpha_t(w) \sum_{k=0}^{\infty} [(k + 1)\mu + \lambda]\frac{[a^2\gamma_t(w, z)]^k}{k!\,(k + 1)!}$$

where $\gamma_t(w, z) = \beta_t(z)\beta_t(w)$. This is a combination of Bessel functions viz. $\mu I_0(2\sqrt{a^2\gamma}) + \lambda/\sqrt{a^2\gamma}\; I_1(2\sqrt{a^2\gamma})$ where I_ν denotes the usual Bessel function with imaginary argument. If we specialize to the coefficient of z^0w^1 we get

$$R^{01}(t) = e^{-\mu t}e^{-2a\sigma}\left[I_0(2a\sigma) + \frac{\lambda}{a\sigma}I_1(2a\sigma)\right] \text{ where } \sigma = 1 - e^{-\mu t} .$$

EXAMPLE 2. Consider the linear growth birth and death process where

$$\lambda_n = (n + 1 + \alpha)\kappa \text{ and } \mu_n = n\kappa \qquad n \ge 0$$

and α is real, $\alpha > -1$. The associated orthogonal polynomials are the Laguerre system normalized at the origin equal to 1. Utilizing the generating function of [5 eq. (25)] we obtain

$$\sum_{0 \le l_1 < l_2} \pi_{l_1}\pi_{l_2} R^{l_1 l_2}(t)[w^{l_2}z^{l_1} - w^{l_1}z^{l_2}] \tag{35}$$
$$= \kappa(\alpha + 1)\delta_t(z)\delta_t(w)[\gamma_t(w) - \gamma_t(z)]\{2F(\alpha + 1, \alpha + 2, 1, u_t(z, w))$$
$$+ \alpha F(\alpha + 1, \alpha + 2, 2, u_t(z, w)\}$$

where F denotes the standard hypergeometric function

$$F(a, b, c, t) = \sum_{n=0}^{\infty} \frac{(a)_n(b)_n}{(c)_n n!}t^n$$

and $(a)_n = \Gamma(a + n)/\Gamma(a)$. Here,

$$\delta_t(z) = \left[\frac{1}{(1 + \kappa t)\left(1 - \frac{\kappa t}{1 + \kappa t}z\right)}\right]^{\alpha+1}, \quad \gamma_t(z) = \frac{\kappa t}{1 + \kappa t}\frac{\left(1 + \frac{1 - \kappa t}{\kappa t}z\right)}{\left(1 - \frac{\kappa t}{1 + \kappa t}z\right)}$$

and $u_t(z, w) = \gamma_t(z)\gamma_t(w)$. The coefficient of z^0w^1 in (35) reduces to

$$\frac{\kappa(\alpha + 1)}{(1 + \kappa t)^{2\alpha+4}}\left\{2F\left(\alpha + 1,\ \alpha + 2,\ 1, \left(\frac{\kappa t}{1 + \kappa t}\right)^2\right)\right. \tag{36}$$
$$\left. + \alpha F\left(\alpha + 1, \alpha + 2, 2, \left(\frac{\kappa t}{1 + \kappa t}\right)^2\right)\right\} .$$

The coincidence time density function $R^{01}(t)$ is the expression (36) apart from the constant factor $1/(\alpha + 1)$. When α is a non-negative integer the coincidence time density function reduces to a rational function. In the particular case $\alpha = 0$, we obtain

$$R^{01}(t) = \frac{2\kappa}{(1 + 2\kappa t)^2}$$

which shows that coincidence is certain with the expected time until coincidence infinite. This is true of all the linear growth processes introduced in this example.

If we examine a linear growth process where there exists a permanent absorbing state at -1 then obviously coincidence is never certain. It is of some interest to compute the probability of coincidence before absorption. Let us illustrate by considering the model where $\lambda_n = (n+1)\kappa$ and $\mu_n = (n+1)\kappa$ for $n \geq 0$. A calculation similar to that above gives

$$R^{01}(t) = \frac{2\kappa}{(1 + 2\kappa t)^2(1 + \kappa t)^2} + \frac{\kappa}{(1 + \kappa t)^5} .$$

It is easy to evaluate $\int_0^\infty R^{01}(t)dt = 25/4 - 8 \log 2$. The reader may verify that this lies between 0 and 1 (approximately. 71).

4. The probability of coincidence. In this section we shall determine the exact conditions which imply that coincidence in a finite state is certain to occur. Our results apply to the case of n independent particles moving simultaneously subject to the transition law of the same birth and death process (B). Our methods may be extended in the obvious way to treat the case in which the particles are subject to different independent birth and death laws. Such a generalization is left to the reader.

If the process (B) is recurrent then coincidence is clearly certain. In fact, if two particles originate in states i and $j > i$, respectively then the second particle reaches the state 0 in finite time with probability one and coincidence must precede this event because of "continuity" of paths. Thus it remains to decide the probability of coincidence when the process (B) is transient.

In [3] we classified two kinds of transient processes. A transient birth and death process is said to be "weakly transient" if $\sum_{j=0}^\infty P_{ij}(t) \equiv 1$ for all t and some i. In terms of the birth and death rates this is equivalent to the divergence to infinity of the sequence

$$\sum_{n=0}^{m-1} \frac{1}{\lambda_n \pi_n} \left(\sum_{k=0}^{n} \pi_k \right) = - Q_m^{(\prime)}(0)$$

where Q_m are the associated polynomials of the process (B).

A birth and death process is said to be strongly transient if for some t and i, $\sum_{j=0}^\infty P_{ij}(t) < 1$. A necessary and sufficient condition for the process to be strongly transient is that, for any starting position and for any positive time value t, with positive probability the diffusing particle reaches infinity in time t.

It becomes evident that for strongly transient processes coincidence is not a certain event, since with positive probability one particle may stay in a given state (say i) in any specified length of time while the other particle moves to infinity without touching state i during this same period of time. An analogous argument will prove that the probability of coincidence for the case of n independently moving particles is not a certain event when the process is strongly transient. We shall determine in Theorem 3 the exact condition for coincidence to be certain. It will be clear that the criteria is the same for two, three or n particles.

We concentrate in what follows on the case of two particles. It is tempting to proceed as follows. Let w_{ij} denote the probability of no coincidence in finite time when two particles start respectively in states i and j $(i < j)$. We set $w_{kk} = 0$. Writing out a recursion relation in terms of the first transition, we obtain

$$
\begin{aligned}
w_{ij} = {} & \frac{\lambda_i}{\lambda_i + \lambda_j + \mu_i + \mu_j} w_{i+1,j} + \frac{\mu_i}{\lambda_i + \lambda_j + \mu_i + \mu_j} w_{i-1,j} \\
& \frac{\lambda_j}{\lambda_i + \lambda_j + \mu_i + \mu_j} w_{i,j+1} + \frac{\mu_j}{\lambda_i + \lambda_j + \mu_i + \mu_j} w_{i,j-1}
\end{aligned} \tag{37}
$$

valid for all $0 \leq i \leq j$. A sufficient condition guaranteeing that coincidence is certain is that the only bounded positive solution of the system (37) is the identically zero solution. In the situation of non-certain coincidence it would also be of interest to calculate the probability of no coincidence w_{ij}. The investigation of this problem is complicated by the abundance of positive solutions that (37) possesses.

The study of (37) is interesting in itself and indicative of the difficulties associated with solving two-dimensional difference equation systems even in comparatively simple cases having probabilistic significance.

To illustrate this we exhibit several solutions of (37). Suppose the spectral measure ψ of (B) is located in the interval $[a, \infty)$ where $a \geq 0$. Then

$$
w_{ij}(\alpha) = -\frac{1}{2\alpha} \begin{vmatrix} Q_i(-\alpha), & Q_i(\alpha) \\ Q_j(-\alpha), & Q_j(\alpha) \end{vmatrix} = -\frac{1}{2\alpha} Q\begin{pmatrix} i, & j \\ -\alpha, & \alpha \end{pmatrix} \tag{38}
$$

for each α satisfying $0 \leq \alpha \leq a$ is positive by virtue of Theorem 1. when $\alpha = 0$, $w_{ij}(0)$ is interpreted as $-Q_j'(0) + Q_i'(0) = \sum_{k=i}^{j-1} 1/\lambda_k\pi_k \sum_{r=0}^{k} \pi_r$. The verification that for all α, $w_{ij}(\alpha)$ is a solution of (37) is accomplished by choosing $x_1 = -\alpha$ and $x_2 = \alpha$ in the recursion law (30).

Unfortunately, there is no natural ordering among the solutions $w_{ij}(\alpha)$. We show first that $w_{0j}(\alpha)$ is increasing in α $(0 \le \alpha \le a)$ for each j. To this end, observe that

$$w_{0j}(\alpha) = Q_j(-\alpha) - Q_j(\alpha) = \sum_{k=1}^{j} a_{jk}\alpha^k$$

where a_{jk} is positive for k odd and zero for k even. Hence, $w_{0j}(\alpha)$ increases as asserted. On the other hand, we show that $w_{j,j+1}(\alpha)$ is decreasing in the same range of α. In fact, by virtue of a known representation [9 p. 42] we have

$$w_{j,j+1}(\alpha) = \frac{1}{\lambda_j \pi_j} \sum_{r=0}^{j} \pi_r Q_r(\alpha) Q_r(-\alpha) .$$

Hence, $w'_{j,j+1}(\alpha) = 1/\lambda_j\pi_j \sum_{r=0}^{j} \pi_r[-Q_r(\alpha)Q'_r(-\alpha) + Q'_r(\alpha)Q_r(-\alpha)]$. It is enough to show since $Q_r(\alpha)Q_r(-\alpha)$ is positive that

$$\frac{-Q_r(\alpha)Q'_r(-\alpha) + Q'_r(\alpha)Q_r(-\alpha)}{Q_r(\alpha)Q_r(-\alpha)} = -\frac{Q'_r(-\alpha)}{Q_r(-\alpha)} + \frac{Q'_r(\alpha)}{Q_r(\alpha)} < 0 . \tag{39}$$

But the roots of $Q'_r(x)$ are separated by the roots of $Q_r(x)$ and since $Q_r(x)$ has no roots in $[-\infty, a)$ [9, p. 43] we conclude that $-Q'_r(x)/Q_r(x)$ is increasing.

The lack of order and the multiplicity of natural positive solutions seem to be the main sources of difficulty in proving the non-existence of any bounded positive solutions of (37). The solution $w_{ij}(0)$ should be singled out because it is always present (as $a \ge 0$) and also $\lim_{j\to\infty} w_{0j}(0) = \infty$ is precisely the condition that the process be weakly transient.

It should be added that the one parameter family of solutions, displayed in (38), when α is a positive number, does not exhaust in terms of linear span the totality of solutions. It appears that one can always construct at least a three parameter family of determinantal extremal solutions. The problem of characterizing all solutions of (37) in general remains open and relates to the problem of determining all determinantal polynomial systems satisfying the recursion law of (30).

We now turn to a discussion of the main theorem of this section.

THEOREM 3. *If the process (B) is recurrent or weakly transient then coincidence is certain if and only if*

$$v_n = \sum_{k=1}^{n} \frac{1}{\lambda_k \pi_k} \sum_{r=0}^{k} \pi_r (w_{k+1} - w_r) \to \infty \tag{40}$$

where $w_m = \sum_0^{m-1} 1/\lambda_i\pi_i \sum_{j=0}^{i} \pi_j$ and $w_0 = 0$.

Before embarking on a proof of the theorem, it is necessary to interpret condition (40). To this end, denote by t_i the random variable which represents the length of time for a particle subject to the transition law of the process (B) to move from state i to state $i+1$. In other words t_i denotes the first passage time from state i to state $i+1$. In the same way, since the path functions are continuous, $z_n = t_0 + t_1 + \cdots + t_{n-1}$ represents the first passage time from state 0 to state n. The t_i are evidently independent but not identically distributed random variables.

The Laplace transform $\varphi_n(s)$ of the distribution of z_n is given by

$$\varphi_n(s) = \frac{1}{Q_n(-s)}$$

when Q_n is the nth orthogonal polynomial. More generally, the Laplace transform of the distribution of $t_m + t_{m+1} + \cdots + t_{n-1}$ is

$$\frac{Q_m(-s)}{Q_n(-s)} .$$

These formulae are proved as follows: The well-known Laplace transform formula, which expresses the first passage time distribution from state i to state j in terms of the transition probability function is

$$\hat{F}_{ij}(s) = \frac{\hat{P}_{ij}(s)}{\hat{P}_{jj}(s)} \qquad i \neq j. \tag{41}$$

Inserting the formula of [2 p. 522] in (41) gives the desired result.

From knowledge of the Laplace transform it is routine (successive differentiation of $\varphi_n(s)$ at zero) to determine the moments of z_n. In particular,

$$E(z_n) = \sum_0^{n-1} \frac{1}{\lambda_k \pi_k} \sum_{r=0}^{k} \pi_r = w_n = -\,Q_n'(0)$$

and

$$\text{variance } (z_n) = -\,Q_n''(0) + [Q_n'(0)]^2 . \tag{42}$$

From identity (11), we get

$$+\,\frac{Q_n''(0)}{2!} = \sum_{t=0}^{n-1} \frac{1}{\lambda_t \pi_t} \sum_{r=0}^{t} \pi_r[-Q_r'(0)] = \sum_{t=0}^{n-1} \frac{1}{\lambda_t \pi_t} \sum_{r=0}^{t} \pi_r w_r .$$

Inserting this in (42) leads to

$$\frac{\mathrm{Var}(z_n)}{2} = \frac{w_n^2}{2} - \sum_{t=0}^{n-1} \frac{1}{\lambda_t \pi_t} \sum_{r=0}^{t} \pi_r w_r .$$

But

$$\frac{1}{2}w_n^2 = \sum_{r=0}^{n-1}(w_{r+1} - w_r)w_{r+1} - \frac{1}{2}\sum_{r=0}^{n-1}(w_{r+1} - w_r)^2$$
$$= \sum_{r=0}^{n-1}\left(\frac{1}{\lambda_r\pi_r}\sum_{k=0}^{r}\pi_k\right)w_{r+1} - \frac{1}{2}\sum_{r=0}^{n-1}(w_{r+1} - w_r)^2$$

$$(43) \qquad \frac{1}{2}\operatorname{Var}(z_n) = \sum_{r=0}^{n-1}\frac{1}{\lambda_r\pi_r}\sum_{k=0}^{r}\pi_k(w_{r+1} - w_k) - \frac{1}{2}\sum_{r=0}^{n-1}(w_{r+1} - w_r)^2$$
$$= v_{n-1} - \frac{1}{2}\sum_{r=0}^{n-1}(w_{r+1} - w_r)^2$$

Since w_i is increasing in i

$$v_{n-1} \geq \sum_{r=0}^{n-1}\frac{1}{\lambda_r\pi_r}\sum_{k=0}^{r}\pi_k(w_{r+1} - w_r) = \sum_{r=0}^{n-1}(w_{r+1} - w_r)^2$$

and hence

$$\frac{1}{2}\operatorname{Var}(z_n) \geq \frac{1}{2}\sum_{r=0}^{n-1}(w_{r+1} - w_r)^2 .$$

If the series $\sum_0^\infty (w_{r+1} - w_r)^2$ is divergent then $v_{n-1} \to \infty$ and $\frac{1}{2}\operatorname{Var}(z_n) \to \infty$, but if the series is convergent then $v_{n-1} - \frac{1}{2}\operatorname{Var}(z_n)$ is bounded. In any case $\{v_n\}$ and $\{\operatorname{Var}(z_n)\}$ either both converge or both diverge.

It is possible for w_n to increase to infinity while at the same time v_n stays uniformly bounded. For example, let

$$\pi_r = \frac{e^{e^r}}{r} \text{ and } \frac{1}{\lambda_r\pi_r} = \frac{1}{e^{e^r}} \qquad \text{for } r \geq 1.$$

A straightforward calculation shows that

$$w_{n+1} \sim \sum_{r=0}^{n}\frac{1}{e^{e^r}}\sum_{k=1}^{r}\cdot\frac{e^{e^k}}{k} = \sum_{r=0}^{n}\frac{1}{r} + \text{a convergent series}$$
$$\sim \log n + c .$$

Also

$$v_n \sim \sum_{k=1}^{n-1}\frac{1}{e^{e^k}}\sum_{l=1}^{k}\frac{e^{e^l}}{l}(w_{k+1} - w_l) .$$

The inner sum grows like its largest term and we have

$$v_n = \sum_{k=1}^{n-1}\frac{1}{k}[\log(k+1) - \log k] + a \text{ convergent sequence}$$

which clearly exhibits v_n as uniformly bounded.

A class of examples in which $v_n \to \infty$ can be constructed as follows. Suppose, π_n and $\dfrac{1}{\lambda_n\pi_n}$ obey the asymptotic relations

$$\pi_n \sim n^{\alpha-1}L(n)\ (\alpha \neq 0) \text{ and } \frac{1}{\lambda_n \pi_n} \sim n^{\beta} L^*(n)$$

where $L(n)$ and $L^*(n)$ are slowly oscillating sequences ($L(n)$ is said to be slowly oscillating if for every $c > 1$, $\frac{L([cn])}{L(n)} \to 1$), $\Sigma \frac{1}{\lambda_n \pi_n} < \infty$ and w_n tends to infinity. Under these conditions we show that v_n tends to infinity. In fact

$$w_n \sim \sum_{r=1}^{n} r^{\beta+\alpha} \tilde{L}(r) \sim n^{\alpha+\beta+1} \tilde{L}(n)$$

where $\tilde{L}(n) = c \cdot L(n)L^*(n)$, ($c$ is a constant) and provided $\alpha + \beta + 1 > 0$. (Similar conclusions hold even in the cases $\alpha = 0$ and $\alpha + \beta = -1$ involving iterates of $L(n)$.)

We next observe that

$$(44) \qquad \begin{aligned} v_n &\geq A \sum_{r=0}^{n} r^{\beta} L^*(r) \sum_{k=0}^{r} k^{\alpha-1} L(k)[w_{r+1} - w_k] \\ &\geq A' \sum_{r=0}^{n} r^{\beta} L^*(r) \sum_{k=1}^{r/2} k^{\alpha-1} L(k) \end{aligned}$$

where A and A' stand for fixed constants. The estimate in (44) is valid since w_r grows like $r^{\alpha+\beta+1}\tilde{L}(r)$. Finally,

$$\begin{aligned} v_n &\geq A'' \sum_{r=0}^{n} r^{\beta} L^*(r) r^{\alpha} L\left(\frac{1}{2} r\right) \\ &\geq A''' n^{\alpha+\beta+1} L(n) \end{aligned}$$

and the proof is finished.

Some other useful conditions that assure the validity of (40) are as follows: If the spectral measure ψ of the birth and death procsse (B) has either

(a) positive measure in every neighborhood of the origin, or if

(b) ψ has an infinite number of points of increase, contained in a bounded interval I, then v_n tends to infinity.

The proof of these statements depend on an alternative representation of the quantity Var z_n. To this effect, we observe that the Laplace transform of z_n can be factored in the form

$$(45) \qquad \varphi_n(s) = \frac{1}{Q_n(-s)} = \frac{1}{\prod_{i=1}^{n} \left(1 + \frac{s}{\alpha_{ni}}\right)}$$

where α_{ni} are the roots of Q_n (recall that the α_{ni} are real and positive). A direct calculation shows that

$$\operatorname{Var} z_n = \sum_{i=1}^{n} \frac{1}{\alpha_{ni}^2} \text{ and } w_n = \sum_{i=1}^{n} \frac{1}{\alpha_{ni}}.$$

In case (a), the first root α_{n1} tends to zero and hence Var z_n becomes unbounded. In case (b) as n increases the interval I must contain an unbounded number of roots α_{ni}, and therefore Var z_n is unbounded. Several notable applications may be recorded.

Queueing models, defined by the parameters $\lambda_n = \lambda$, $n \geq 0$, $\mu_n = \mu$, $n \geq k_0$, $\mu_0 = 0$, and k_0 a prescribed positive integer, have the property that coincidence is a certain event. In fact, for these examples case (b) applies (see [4]).

The situation of linear growth, birth and death processes, (*i.e.* $\lambda_n = \lambda n + a$, $n \geq 0$, $\mu_n = \mu n + b$, $n > 0$, $\mu_0 = 0$) with regard to the probability of coincidence is as follows. If $\mu = \lambda$, then coincidence is always certain (case (b) above). If $\mu > \lambda$ then the process is recurrent and coincidence is trivially certain. If $\mu < \lambda$ then the process is weakly transient and coincidence is not certain. This last assertion is proved as follows. The spectral measure is discrete with mass points located essentially at an arithmetic series. The roots of $Q_n(-s)$ for any n are separated by the mass points of ψ and hence always $\Sigma\left(\frac{1}{\alpha_{ni}}\right)^2 \leq k\Sigma\frac{1}{n^2} < C$.

We turn now to the proof of the theorem. The arguments are divided into a series of lemmas.

DEFINITION (Levy [7]). A series of independent random variables $x_1 + \cdots + x_n = s_n$ is *essentially divergent* if there exists no sequence of constants a_n such that $s_n - a_n$ converges almost surely to a finite random variable.

LEMMA 1. *If v_n is divergent then the series of independent random variables $t_0 + t_1 + \cdots + t_{k-1} = z_k$ is essentially divergent. (The meaning of t_r is as before.)*

Proof. Suppose we can find a sequence of constants a_n such that $z_n - a_n$ converges. In particular, its characteristic function

$$\frac{e^{ia_n\lambda}}{Q_n(-i\lambda)} \text{ converges for each real } \lambda$$

to a characteristic function $\varphi(\lambda)$. It follows that the corresponding symmetrized random variable with characteristic function

$$\frac{1}{|Q_n(-i\lambda)|^2} \text{ converges to } |\varphi(\lambda)|^2$$

for each real λ and uniformly in any finite interval. But, by virtue of (45) for $\lambda > 0$

$$|Q_n(-i\sqrt{\lambda})|^2 = \prod_{i=1}^{n}\left(1 + \frac{\lambda}{\alpha_{ni}^2}\right) \geq 1 + \lambda\sum_{i=1}^{n}\frac{1}{\alpha_{ni}^2}$$
$$= 1 + \lambda \mathrm{Var}\, z_n\,.$$

Hence, for $\lambda \neq 0$, $|Q_n(-i\lambda)|^2$ tends to infinity and $|\varphi(\lambda)|^2 \equiv 0$. Thus $\varphi(\lambda)$ is not a characteristic function as required. The contradiction implies that $z_n - a_n$ cannot converge for any sequence of constants and consequently z_n is essentially divergent as was to be shown.

COROLLARY 1. Suppose v_n is divergent and let t_i and t'_i represent independent observations of the first passage time from state i to state $i+1$. Then

$$t_0 + t_1 + \cdots + t_{k-1} - t'_0 - t'_1 - \cdots - t'_{k-1} = z_k - z'_k$$

is essentially divergent.

Proof. This is clear since the characteristic function of $z_k - z'_k - a_k$ ($a_k = a$ sequence of constants) is

$$\frac{e^{i\lambda a_k}}{|Q_k(-i\lambda)|^2}$$

which for real $\lambda \neq 0$ tends to zero as shown in the proof of Lemma 1.

LEMMA 2. *With the same notation as in Corollary* 1, *if v_n diverges then for every fixed r*

(46) $$\Pr\{[t_0 + t_1 + \cdots + t_k] - [t'_r + t'_{r+1} + \cdots + t'_k] < 0 \text{ i.o.}\} = 1$$

(i.o. is an abbreviation of infinitely often).

Proof. With r held fixed it will be sufficient to prove that

(47) $$\Pr\{t'_r + t'_{r+1} + \cdots + t'_k - t_r - t_{r+1} - \cdots - t_k > C \text{ i.o.}\} = 1$$

for every positive constant C. Indeed, the validity of (47) implies that for almost every value of $t_0 + t_1 + \cdots + t_{r-1}$

$$1 = \Pr\{t_0 + \cdots + t_k - (t'_r + \cdots + t'_k) < 0 \text{ i.o.} \mid t_0 + t_1 + \cdots + t_{r-1}\} .$$

Invoking the law of total probabilities leads immediately to the conclusion (46).

We devote ourselves now to the proof of relation (47). Since the series $(t'_r - t_r) + (t'_{r+1} - t_{r+1}) + \cdots + (t'_k - t_k) = T_k$ (the dependence of T_k on r is suppressed since we are keeping r fixed) is essentially divergent we may appeal to a theorem of P. Levy [7 p. 147] and deduce that if A_k is any sequence of constants

(48) $$\Pr\{T_k \geq A_k \text{ i.o.}\}$$

is either 0 or 1. We select for our purpose all $A_k = 0$. Since T_k constitute a series of symmetric random variables the value of the expression (48) is clearly 1. By virtue of a second theorem of P. Levy [7 p. 147],

$$\Pr\{T_k \geq C \text{ i.o.}\} = 1$$

for any constant C and the proof of the lemma is finished.

Proof of the Sufficiency of Theorem 3. Suppose for definiteness that particle labeled (i) starts in state 0 and particle labeled (ii) starts in state r, each independently subject to the same transition law. Let t_i and t'_i, for particles (i) and (ii) respectively, represent as previously the first passage time from state i to state $i+1$. Lemma 2 assures that with probability 1 there is a state k such that the particle labeled (i), having started at zero, reaches k for the first time earlier than the particle labeled (ii) whose initial state was r. Since the path functions are continuous, the two particles necessarily cross and coincidence is certain.

Necessity. The proof of necessity will likewise be written in the form of a series of lemmas.

LEMMA 3. *If v_n is bounded then*

$$\Pr\{t_0 + t_1 + \cdots + t_{k-1} - t'_r - t'_{r+1} - \cdots - t'_k > 0 \quad \textit{for all } k \geq r\} > 0 . \tag{49}$$

Proof. Consider $T_k = (t'_r - t_r) + \cdots + (t'_k - t_k)$, $k = r, r+1, \cdots$, which is a partial sum composed of independent symmetrically distributed random variables. The hypothesis (see (43)) means that the variance of T_k is uniformly bounded. Therefore, invoking the three series theorem (because $t'_i - t_i$ are symmetric only the convergence of the series formed by the variances of the successive terms has to be verified), we may conclude that T_k converges almost surely to a finite valued random

Let t^* denote the limit of T_k. Take any value C such that

$$\Pr\{|t^*| < C\} > 0 .$$

Since T_k converges almost surely to t^* there is a k_0 such that

$$\Pr\{|T_k| < C \text{ for all } k \geq k_0\} > 0 .$$

Making C even larger (say C') if necessary we can assure

$$\Pr\{|T_k| < C' \text{ for all } k = r, r+1, \cdots,\} > 0 . \tag{50}$$

Consider now the random variable $t_0 + t_1 + \cdots + t_{r-1}$ which is independent of all T_k, $k \geq r$. Since t_0 is exponentially distributed it follows that

$$\Pr\{t_0 + t_1 + \cdots + t_{r-1} > C'\} > 0 \tag{51}$$

for any C' sufficiently large. Combining (50) and (51) yields the estimate

$$\Pr\{t_0 + t_1 + \cdots + t_{r-1} - T_k > 0 \text{ for all } k \geq r\}$$
$$\geq \Pr\{t_0 + \cdots + t_{r-1} > C'\} \Pr\{T_k < C' \text{ for all } k \geq r\} > 0$$

for an appropiate positive constant C.

This means that with positive probability a particle starting at zero never reaches a state $k \geq r+1$ for the first time at an earlier time then a particle beginning in state r. The proof of the lemma is finished.

LEMMA 4. *If coincidence is a certain event when the particles have a prescribed pair of initial states r, s $(r < s)$ then coincidence is a certain event for any pair of initial states.*

Proof. This is a direct consequence of the fact that with positive probability any pair of state i, j $(i < j)$ can be attained starting from the initial states r and s without the occurrence of coincidence.

Consequently, if there exists positive probability of no coincidence starting from i and j, respectively then the same is true for r and s contrary to the hypothesis.

LEMMA 5. *Let coincidence be a certain event. Suppose the initial states of the two particles* (i) *and* (ii), *respectively are i_0 and $j_0 > i_0$. Then the event that particle* (ii) *reaches every state k $(k \geq k_0)$ for the first time ahead of particle* (i) *has probability zero.*

Proof. We shall prove the lemma by producing an infinite sequence of states $k_1 < k_2 < \cdots$ with the following properties (called A). If the initial states of the particles (i) and (ii) are any pair r and s where $r < s$ and $s \leq k_i$ then the probability exceeds 1/4 that particle (i) will reach state k_{i+1} ahead of particle (ii).

Let us suppose statement (A) is established and now show how to finish the proof of the lemma. To this end, we have

$$\Pr\{\text{(ii) reaches state } k \text{ prior to (i) for all } k \geq k_0\}$$
$$\leq \Pr\{\text{(ii) reaches state } k_i \text{ prior to (i) for all } k_i \geq k_0\}$$
$$\leq \prod_{k_i > k_0} (1 - \Pr\{\text{(i) reaches state } k_{i+1} \text{ prior to (ii)} \mid \text{(ii) reaches state}$$
$$k_i \text{ prior to (i)})\} .$$

The infinite product is zero since on account of statement (A) infinitely many factors are $\leq 3/4$.

It remains to prove statement (A).

Suppose we have already constructed $k_1, k_2, \cdots, k_i$. Since coincidence is a certain event regardless of the pair of initial states r and s, $(r < s$

and $s \leq k_i$) there exists a time value t_0 so that with probability $\geq 1 - \varepsilon$ coincidence occurs sometime earlier then t_0. The value of t_0 may be determined for each pair of initial states r and s. However, since there are only a finite number of possibilities r, s ($r < s$) where $s \leq k_i$ we can choose t_0 large enough so that the same value of t_0 applies for any of these pairs of starting states. By further reducing to a subset of paths of probability $\geq 1 - 2\varepsilon$ (ε can be specified in advance as small as desired) we can determine a state $k_{i+1} > k_i$ which is not entered by either particle in the time duration $(0, t_0)$. The existence of k_{i+1} is guaranteed since the hypothesis of the lemma postulates that coincidence is certain and hence the process cannot be strongly transient. Restricting consideration to this set of paths we note that at the first instance of coincidence the two particles are indistinguishable and hence, with probability 1/2, particle (i) will enter state k_{i+1} ahead of particle (ii). Let E_i denote the event that (i) reaches state k_{i+1} ahead of (ii) when the initial states respectively are any pair r and s, $s \leq k_i$.

The above argument establishes that $\Pr\{E_i\} \geq (1 - 2\varepsilon)/2 > 1/4$ and the proof is hereby complete.

Proof of Necessity. This is immediate by comparing Lemmas 3 and 5.

The problem of computing the probability of coincidence for the case when v_n is bounded remains open.

We close with some observations regarding the problem of determining criteria which guarantee finite expected time for coincidence. First it is evident that for an ergodic birth and death process the expected time until coincidence is finite. To decide when the event of coincidence has a finite expected time is in general an open question.

The following two examples are of some interest. In the case of the linear growth processes associated with the Laguerre polynomials, we were able to determine a double generating function for the explicit distribution of the coincidence time (33). Here, it is easy to show by direct calculation that the expected coincidence time is infinite.

We shall now prove that for the recurrent null or transient queueing model (labeled B) the expected coincidence time is infinite. For definiteness $\lambda_n = \lambda$, $n \geq 0$ and $\mu_n = \mu$, $n \geq 1$, $\mu_0 = 0$.

We consider for the situation of two particles starting in states i and j, $0 < i < j$, the following induced random walk W whose state space is composed of the non-negative integers. We say that W is in state r if $j - i = r$. Transitions in W are engendered whenever one of the particles of process B changes its state. Explicitly a transition of W occurs from state r to $r - 1$ if and only if after the first change the state labels of the two particles, undergoing the process B, are either $(i + 1, j)$ or $(i, j - 1)$. A movement from r to $r + 1$ occurs in the contrary case. The motion on W, thus induced by the birth and death process will be understood to apply only when $i > 0$. The homogeneity of the queueing model implies that the changes engendered in W are independent of the specific states occupied by the two particles of the process B and only depend on their distance $(j - i)$ apart provided $i > 0$. Hence

$$\Pr_W\{r \to r - 1\} = \Pr_W\{r \to r + 1\} = 1/2 \text{ for } r > 0 .$$

It is well known that for this random walk the time until first passage into the state 0 from any non-zero initial state has an infinite expected value [3]. Moreover, first passage into 0 obviously corresponds to the event of coincidence for the original birth and death process. There is one slight complication in the above argument arising from the fact that when one of the particles of process B starting at i reaches zero, the transition probabilities of the induced random walk do not agree with the probabilities of the changes in distance between the particles. This is due to the reflecting character of state zero, *i.e.* when one of the particle of its process is in state 0 then this particle can only move to state 1. We will show that this complication is of no consequence in deciding whether coincidence in B occurs with finite expected time.

Let the particles begin in states i and $j, (i < j)$. Since coincidence is certain let us consider all those paths E where coincidence occurs without either particle ever reaching zero. Conditioned in this way the induced random walk describes the changes of the "distance" (number of states separating the two particles) until coincidence. But, for the random walk W the expected number of transitions for the first passage into zero is infinite. Since the expected time between transitions for the birth and death process is $1/(\lambda + \mu)$, the expected time until coincidence averaged over the paths of E is infinite. Next, let F denote the set of paths in the process B where the particle, starting in state $i < j$, reaches state zero before coincidence. Since the process B is null recurrent or transient, again the expected time length of the paths of F is infinite. Hence, under either circumstance the expected time until coincidence is infinite.

The above argument may be extended to prove that if a birth and death process is null recurrent of transient with certain coincidence, then the expected time of coincidence is infinite provided $\sum_{i=1}^{\infty} 1/(\lambda_i + \mu_i) = \infty$, and

$$1 - p_n = q_n = \max_{i>0} \frac{\lambda_i + \mu_{i+n}}{\lambda_i + \mu + \lambda_{i+n} + \mu_{i+n}}, \qquad p_0 = 1$$

are the transition parameters of a recurrent null or transient random walk W on the integers (*i.e.* $\Pr_W \{n \to n + 1\} = p_n$).

On the other hand, the expected coincidence time is finite whenever

$$1 - p_n = q_n = \min_{i>0} \frac{\lambda_i + \mu_{i+n}}{\lambda_i + \mu_i + \lambda_{i+n} + \mu_{i+n}}$$

describes an ergodic random walk W and $\max_j 1/(\lambda_j + \mu_j)$ is bounded.

REFERENCES

1. K. L. Chung, *Foundations of the theory of continuous parameter Markoff chains*, Proc. Third. Berkeley Symposium, 2, (1956), 29-40.
2. S. Karlin and J. L. McGregor, *The differential equations of birth and death processes and the Stieltjes moment problem*, Trans. Amer. Math. Soc., **86** (1957), 489-546.
3. Samuel Karlin and J. L. McGregor, *The classification of birth and death processes*, Trans. Amer. Math. Soc., **86** (1957), 366–401.
4. ———, *Many server queueing processes with Poisson input and exponential service times*, Pacific J. Math., **8** (1958), 87–118.
5. ———, *Linear growth, birth and death processes*, Journal of Math. and Mechanics, **7** (1958), 643-662.
6. ———, *A characterization of the birth and death process*, Proc. Nat. Acad. of Sciences **45** (1959), 375-379.
7. P. Levy, *Théorie de l'addition des variables aléatorires*, Paris, Gauthier-Villars, 1937.
8. G. Pólya and G. Szegö, *Aufgaben und lehrsätze aus der analysis*, I, Berlin, 1925.
9. G. Szegö, *Orthogonal polynomials*, Amer. Math. Soc. Colloquium Publications, 23 (1939).

STANFORD UNIVERSITY