Z. Wahrscheinlichkeitstheorie verw. Geb. 7, 271—288 (1967)

# The Limit of a Sequence of Branching Processes

JOHN LAMPERTI*

Received October 26, 1966

## 1. Introduction

In this paper, the term 'branching process' will mean a Markov chain whose states are the nonnegative integers and whose transition probabilities are given by

$$P_{ij} = \text{coefficient of } x^j \text{ in } f(x)^i, \tag{1.1}$$

where $f(x) = \sum_{k=0}^{\infty} p_k x^k$ is the generating function of a set of probabilities $\{p_k\}$. These basic probabilities can be interpreted as governing the number of 'offspring' which an individual in a population will contribute to the following 'generation'. Thus (1.1) expresses the idea that if the $n$'th generation contains $i$ individuals, the $n+1$'st generation will consist of the sum of $i$ independent random variables representing their immediate descendents, each variable having distribution $\{p_k\}$. Background on these simple branching or 'Galton-Watson' processes is given in chapter 1 of [4]. In particular, we will repeatedly need the fact that

$$P_{ij}^{(n)} = \text{coefficient of } x^j \text{ in } f_n(x)^i, \tag{1.2}$$

where $f_n$ is the $n$'th functional iterate of the basic generating function $f(x) = f_1(x)$.

In a previous paper [6] the following situation was examined: Let $\{Z_n\}$ denote the random variables of a branching process governed by certain fixed probabilities $\{p_k\}$. Let us *suppose* that the processes

$$x_t^{(r)} = \frac{Z_{[rt]} - a_r}{b_r}, \quad \text{where } Z_0 = c_r, \tag{1.3}$$

converge as $r \to \infty$ to a limiting random process $\{x_t\}$. Here the $c_r$ are positive integers $\to \infty$, $b_r > 0$ and $a_r$ are normalizing constants, and 'converge' means that the finite-dimensional joint distributions of $\{x_t^{(r)}\}$ converge to those of $\{x_t\}$ in the usual (weak) sense. The problem is to determine all the possible limit processes which can arise under various choices of $\{p_k\}$, $a_r$, $b_r$, and $c_r$. It was shown in [6] that in case there is no centering ($a_r \equiv 0$), the limiting processes form, apart from scaling, only a one-parameter family. The most familiar example occurs when $\{p_k\}$ has mean one (always a necessary condition) and finite variance. The choices $b_r = c_r = r$ then result in a limiting diffusion process on $[0, \infty)$ whose backward equation is of the form

$$\frac{\partial u}{\partial t} = \frac{\beta x}{2} \frac{\partial^2 u}{\partial x^2}, \quad \beta > 0. \tag{1.4}$$

* This research was in part supported by the (U. S.) National Science Foundation.

This is the only diffusion in the family; the other limiting processes are also martingales but do not have continuous paths. When nontrivial centering ($a_r/b_r \to \infty$) is allowed, however, it was only shown that any limiting process must have independent increments; the precise class which arises was not determined in this case. If the variance of $\{p_k\}$ is finite, the limit is the Brownian motion process.

The above situation does not include all known examples of limits for branching processes. If the set-up in (1.3) is generalized by allowing the basic probabilities $\{p_k\}$ to depend on $r$, essentially different cases arise. For example, in [2] W. FELLER considers a limit with scaling and without centering, where the distribution $\{p_k\}$ has a fixed (finite) variance, but its mean is of the form $1 + \alpha/r$. He then obtains a limiting diffusion corresponding to the (backward) equation

$$\frac{\partial u}{\partial t} = \frac{\beta x}{2} \frac{\partial^2 u}{\partial x^2} + \alpha x \frac{\partial u}{\partial x} . \tag{1.5}$$

(See also [7] for some related discussions.) Unless $\alpha = 0$, processes satisfying (1.5) cannot occur as limits when $\{p_k\}$ is held fixed during the limiting process. Another relevant example, of a quite different sort involving centering but no scaling, was discussed by STRATTON and TUCKER in [8]. (But see the remark at the end of section 3 below.)

The purpose of the present paper is to study the possible limiting processes for a sequence $\{x_t^{(r)}\}$ under these more general conditions (that is, when $\{p_k\}$ may depend on $r$), and we will succeed in obtaining a description of the class of limits $\{x_t\}$ which can occur. The limiting processes are of quite different sorts in the two cases $a_r = 0$, $c_r = b_r$, and $a_r/b_r \to \infty$, $a_r = c_r$. (We will see that no others need be considered.) In the first case, $\{x_t\}$ retains essentially the character of a branching process but has the nonnegative reals, rather than the integers, as its state-space. Such processes were introduced and studied in their own right by M. JIŘINA in [5]. In the second case, with nontrivial centering, $\{x_t\}$ must be an additive process whose canonical (Lévy) measure is supported on $[0, \infty)$. (There is an exception to the last statement, in case $b_r \not\to \infty$.) Conversely, in each case we will see that every process of the sort indicated does in fact arise as the limit of a suitable sequence of normalized branching processes $\{x_t^{(r)}\}$.

The plan of the paper is as follows: In section 2 we define and discuss the 'continuous state branching processes' (C.B. processes) which form the class of limits in case one above. A deeper study of this interesting class is planned for a future publication[1]. Next, in section 3 we formulate all of the limit theorems, for both case one and case two (that is, both without and with nontrivial centering). The proofs follow in sections 4 and 5 for the two situations respectively.

One preliminary remark may as well be located here. It would apparently be possible to generalize (1.3) further by using $Z_{[d_r t]}$ in place of $Z_{[rt]}$, where $d_r$ could be any sequence tending to $+\infty$. However, it will be quite clear from the proofs of Theorems 1 and 3 that these results, which characterize the limiting processes,

[1] An announcement of some of the results will appear shortly in Bull. Amer. Math. Soc.

are still valid without change in this situation. Theorems 2 and 4, moreover, are formally stronger with the specific choice $d_r = r$ than if stated for arbitrary $d_r$. Hence in order to simplify our already sufficiently cumbersome notation, and with no real loss, we will keep to the choice $d_r = r$ throughout.

## 2. Continuous State Branching Processes

We will consider a class of functions $P_t(x, E)$ as follows:

(i) $P_t(x, E)$ is defined for $t \geqq 0$, $x \geqq 0$, and $E$ a Borel subset of the half line $[0, \infty)$.

(ii) For fixed $t$ and $x$, $P_t(x, \cdot)$ is a probability measure on the Borel sets of $[0, \infty)$, and for fixed $E$, $P_t(x, E)$ is jointly measurable in $x$ and $t$.

(iii) The Chapman-Kolmogorov equation holds:

$$\int_0^\infty P_t(u, E)\, P_s(x, du) = P_{t+s}(x, E)\,. \tag{2.1}$$

(iv) For each $x$, $y$, $t \geqq 0$, $P_t$ satisfies

$$P_t(x + y, \cdot) = P_t(x, \cdot) * P_t(y, \cdot)\,, \tag{2.2}$$

where the symbol '$*$' denotes convolution.

(v) There exist $t > 0$ and $x > 0$ such that $P_t(x, \{0\}) < 1$.

**Definition.** A function $P_t$ satisfying (i) to (v) above will be called a 'C.B. function'; a Markov process with such a function for its transition probabilities will be a 'C.B. process'.

Remark 1. It is evident that conditions (i) to (iii) are simply the definition of a Markov transition function on a certain state space. Postulate (iv) is a special condition representing the 'branching property'. Indeed, if the states were the nonnegative integers instead of the whole of $[0, \infty)$, then (2.2) would be the characteristic property of 'Markov branching processes' in the usual sense [4]. Condition (v) simply rules out an undesirable trivial case.

Remark 2. Continuous branching processes were introduced by M. Jiřina in [5]. His definition was more general than ours, in that time homogeneity was not assumed and the dimension of the state space could exceed one. However, the analysis in [5], when dealing with the continuous parameter case, was confined to processes of the 'purely discontinuous' type, and we will not want such a restriction.

If $P_t(x, E)$ is a C.B. function, it is clear from (2.2) that the distributions $P_t(x, \cdot)$ are infinitely divisible. It also follows, arguing in the usual way, that their Laplace transforms can be written in the form

$$\int_0^\infty e^{-\lambda y}\, P_t(x, dy) = e^{-x\psi_t(\lambda)} \qquad (\lambda \geqq 0)\,. \tag{2.3}$$

Using (2.3) it is easy to translate the Chapman-Kolmogorov equation (2.1) into the equivalent condition that

$$\psi_{t+s} = \psi_t\{\psi_s(\lambda)\}\,. \tag{2.4}$$

This is analogous to the functional iteration property of the generating functions of a Galton-Watson process.

We will now establish several facts of a technical nature for future use. The first of these is an easy consequence of the definition and we state it without proof:

**Lemma 2.1.** *A C.B. function satisfies*

$$P_t(x, \{0\}) < 1 \quad \text{for all} \quad t, x > 0; \quad P_t(0, \{0\}) = 1\,. \tag{2.5}$$

**Lemma 2.2.** *If $P_t$ is a C.B. function, the operators*

$$T_t f(x) = \int_0^\infty f(y)\, P_t(x, dy) \tag{2.6}$$

*define a contraction semigroup on the space $C_0$ of all continuous functions on $[0, \infty)$ which vanish at $\infty$.*

*Proof.* It is only necessary to show that $T_t$ takes $C_0$ into itself. For finite $x$ this is an immediate consequence of (2.3) (and so of (2.2)). Indeed, the right hand side of (2.3) is certainly continuous in $x$, and by the continuity theorem for Laplace-Stieltges transforms this fact — the continuity of $T_t f$ in the special cases $f(x) = e^{-\lambda x}$ — implies continuity for any $f \in C_0$. It remains to see that $T_t f(\infty) = 0$. But from Lemma 2.1 it follows that $\psi_t(\lambda) > 0$ for each $\lambda > 0$, so that as $x \to \infty$ the Laplace transform of $P_t(x, \cdot)$ tends to 0 for $\lambda > 0$. This implies that the measure $P_t(x, \cdot)$ tends to have its mass 'at $\infty$', and hence $T_t f(x) \to 0$ and the proof is complete.

**Lemma 2.3.** *A C.B. function is stochastically continuous.*

*Proof.* 'Stochastic continuity' means that, for each $x$, the measures $P_t(x, \cdot)$ converge weakly to a unit mass at $x$ when $t \to 0+$. We begin with the remark that, because of (ii), the function

$$g(t) = \int_0^\infty d\mu(x) \int_0^\infty f(y)\, P_t(x, dy)$$

is measurable in $t > 0$ for any finite Borel signed measure $\mu$ on $[0, \infty)$ and any $f \in C_0$. This is equivalent to the assertion that the semigroup $T_t$ is 'weakly measurable'. It is well known that weak measurability implies that $T_t$ is in fact strongly continuous for $t > 0$. (See, for instance, p. 35 of [1].)

Specializing to the case $f(x) = e^{-\lambda x}$, $\lambda > 0$, strong continuity of $T_t$ implies that the function $\psi_t(\lambda)$ is continuous in $t > 0$ for each (fixed) $\lambda$. Using (2.4) this can be written as

$$\lim_{h \to 0+} \psi_h(\psi_t(\lambda)) = \psi_t(\lambda)$$

for each $t > 0$, $\lambda > 0$. In other words,

$$\lim_{h \to 0+} \psi_h(u) = u$$

for each number $u$ which can be represented as $\psi_t(\lambda)$ for some $t$, $\lambda$. But because of (2.3) and (2.5) the set of such $u$ certainly contains some interval $[0, a]$, $a > 0$. Thus for each $x$ and each $\lambda \in [0, a]$, the Laplace transform of $P_h(x, \cdot)$ converges as $h \to 0+$ to $e^{-x\lambda}$, and that is sufficient for the continuity theorem to apply and yield the desired conclusion.

The rest of this section will be devoted to some examples and further elementary facts concerning C.B. processes; what follows is not logically necessary for the sequel. We begin with the observation that (2.3) and (2.4) can be used to compute moments of the process, if they exist. For example, since $\psi_t(0) = 0$ we have from (2.3) that

$$\int_0^\infty y\, P_t(x, dy) = x\psi_t'(0+), \tag{2.7}$$

where the prime means differentiation with respect to $\lambda$. But differentiating (2.4) yields the functional equation

$$\psi_{t+s}'(0+) = \psi_t'(0+)\,\psi_s'(0+),$$

and so $\psi_t'(0+)$, if finite, must be of the form $e^{\alpha t}$ for some constant $\alpha$. As a result, we have

$$\int_0^\infty y\, P_t(x, dy) = x\, e^{\alpha t}. \tag{2.8}$$

The second moment can be evaluated similarly if it exists. From (2.4) and the result above we obtain

$$\psi_{t+s}''(0+) = \psi_t''(0+)\, e^{2\alpha s} + \psi_s''(0+)\, e^{\alpha t}. \tag{2.9}$$

If $\alpha = 0$ the functional equation (2.9) has only the solution $\psi_t''(0+) = \beta t$. When $\alpha \neq 0$, we note that interchanging $s$ and $t$ in (2.9) and subtracting yields

$$\psi_t''(0+)\,(e^{\alpha s} - e^{2\alpha s}) = \psi_s''(0+)\,(e^{\alpha t} - e^{2\alpha t}).$$

Upon setting $s = 1$ this takes the form

$$\psi_t''(0+) = \frac{-\beta}{\alpha}(e^{2\alpha t} - e^{\alpha t}), \quad \beta = \text{constant}.$$

In view of the fact that

$$\int_0^\infty y^2\, P_t(x, dy) = x^2\, \psi_t'(0+)^2 - x\psi_t''(0+) \tag{2.10}$$

we finally obtain the formula

$$\int_0^\infty y^2\, P_t(x, dy) = \begin{cases} x^2 + \beta x t & \text{if } \alpha = 0, \\ x^2\, e^{2\alpha t} + \dfrac{\beta}{\alpha} x (e^{2\alpha t} - e^{\alpha t}) & \text{if } \alpha \neq 0, \end{cases} \tag{2.11}$$

where the constant $\alpha$ is that appearing in (2.8).

If both (2.8) and (2.11) hold, it is very easy to obtain

$$\lim_{t \to 0} t^{-1} \int_0^\infty (y - x)\, P_t(x, dy) = \alpha x, \tag{2.12}$$

$$\lim_{t \to 0} t^{-1} \int_0^\infty (y - x)^2\, P_t(x, dy) = \beta x. \tag{2.13}$$

This strongly suggests that *if* the C.B. function $P_t$ is the transition probability of a diffusion process, it must be one with the backward equation (1.5). We will prove this in the future paper which was promised in the introduction.

The transition probability function for the diffusion processes satisfying (1.5) was given explicitly in [2]. The Laplace transform (2.3) can be obtained from that formula, although it is simpler to get it directly by solving (1.5). In any case the result is

$$\psi_t(\lambda) = \begin{cases} \dfrac{\lambda}{1+\dfrac{\beta t \lambda}{2}} & \text{if} \quad \alpha = 0, \\ \dfrac{\lambda e^{\alpha t}}{1-\dfrac{\beta\lambda}{2\alpha}(1-e^{\alpha t})} & \text{if} \quad \alpha \neq 0. \end{cases} \tag{2.14}$$

It is trivial to verify directly that (2.4) is satisfied. From (2.14) it is easy to derive many facts about the process, such as the value of $P_t(x, \{0\})$ and that ultimate absorption at $x = 0$ is certain if and only if $\alpha \leqq 0$.

As discussed in [6], when $\alpha = 0$ there is a class of examples with infinite variance which are related to (2.14). Their transforms are

$$\psi_t(\lambda) = \frac{\lambda}{\left[1 + t\dfrac{\beta}{2}\lambda^p\right]^{1/p}}, \quad 0 < p \leqq 1. \tag{2.15}$$

There are also solutions of (2.4) which are similarly related to (2.14) when $\alpha \neq 0$:

$$\psi_t(\lambda) = \frac{\lambda e^{\alpha t/p}}{\left[1 - \dfrac{\beta\lambda^p}{2\alpha}(1-e^{\alpha t})\right]^{1/p}}, \quad 0 < p \leqq 1. \tag{2.16}$$

An interpretation for the processes whose transition functions satisfy (2.16), explaining their relation to stable laws, will be given in the forthcoming paper already mentioned.

We conclude this section with a brief summary of the relevant results in [5]. Suppose that in addition to being a C.B. function (i.e., to satisfying (i) to (v)), $P_t$ also is the transition function of a purely discontinuous process. This is taken to mean the existence of the limits

$$\lim_{t\to 0} \frac{P_t(x, E) - X_E(x)}{t} = \pi(x, E) \tag{2.17}$$

for all $x \geqq 0$ and Borel sets $E \subset [0, \infty)$; $X_E$ is the indicator function of $E$. JIŘINA shows that because of (iv) we have

$$\pi(x, E) = x\,\pi(1, E - x + 1) \tag{2.18}$$

and

$$\pi(x, [0, x)) = 0 \quad \text{for all} \quad x > 0. \tag{2.19}$$

The first of these conditions implies that the expected waiting time before jumping from $x$ is inversely proportional to $x$ and that the magnitude of the jump, when it occurs, has a distribution not depending on $x$. The second condition implies that the process cannot move to the left.

JIŘINA also considers the converse question of constructing the most general C. B. process satisfying (2.17). It is easy to find the corresponding $\pi$; it must have the form

$$\pi(x, E) = a\,x\,\mu(E - x + 1) \quad (a > 0)\,, \tag{2.20}$$

where $\mu$ is a set function determined by combining any Borel probability measure on $(1, \infty)$ with a point mass of value $-1$ at $x = 1$. The minimal transition probability determined by (2.20) and (2.17) then satisfies all the requirements for a C. B. function except that $P_t(x, [0, \infty)) < 1$ can occur (the process may 'explode'). These facts are analogous to those for the discrete state Markov branching process [4]. Setting

$$\varphi(\lambda) = \int_0^\infty e^{-\lambda x}\, d\mu(x)\,, \tag{2.21}$$

a necessary and sufficient condition that $P_t(x, [0, \infty)) = 1$ for all $t \geqq 0$ is

$$\int_0^\delta \frac{d\lambda}{\varphi(\lambda)} = -\infty\,, \qquad \delta > 0\,. \tag{2.22}$$

It is sufficient but not necessary that the probability measure from which $u$ is constructed have a finite mean. These results allow the construction of a great variety of C. B. processes, and show decisively that this class is far wider than the family of limiting processes found in [6] whose transforms were given in example (2.15) above.

One final remark. Any C. B. process with the property that absorption at 0 is possible (i.e., such that $P_t(x, \{0\}) > 0$ for some $t, x > 0$) obviously is able to move to the left. Hence no such process can be of the purely discontinuous type. In particular, the examples determined by (2.14), (2.15) and (2.16) have this property, as does any C. B. function such that $P_t(x, \cdot)$ is a compound Poisson law.

## 3. Statement of Limit Theorems

We consider a sequence of processes $\{x_t^{(r)}\}$ defined as in (1.3); for each $r$, $\{Z_n\}$ is a branching process governed by probabilities $\{p_k\}$, with generating function $f(x)$. *These may depend on* $r$, but this dependence will not be explicitly indicated in order to avoid complicating the notation. Our basic assumption is that there exists a stochastic process $\{x_t\}$, $t \geqq 0$, which is the limit of $\{x_t^{(r)}\}$ as $r \to \infty$ in the sense that the finite-dimensional joint distributions of $\{x_t^{(r)}\}$ converge to those of $\{x_t\}$. To be specific, this means that for every finite set $0 = t_0 < t_1 < \cdots < t_h$,

$$\begin{aligned} &\lim_{r\to\infty} P(x_{t_0}^{(r)} \leqq y_0, \ldots, x_{t_h}^{(r)} \leqq y_h) \\ &= P(x_{t_0} \leqq y_0, \ldots, x_{t_h} \leqq y_h) \end{aligned} \tag{3.1}$$

in the usual sense of weak convergence of measures on $R^{h+1}$.

We first take up the case $a_r = 0$ for all $r$. It will be required that $c_r \to \infty$, for if this is not done any continuous-time, discrete-state Markov branching process $\{z_t\}$ can be trivially obtained as a limit by setting $b_r = c_r = 1$ and $Z_n = z_{n/r}$.

**Theorem 1.** *Suppose that* (3.1) *holds with* $a_r = 0$, $c_r \to \infty$, *and* $P(x_t = 0) < 1$ *for some* $t > 0$. *Then* $\{x_t\}$ *is a C. B. process in the sense of section* 2.

Remark. It would actually be possible to state the result under the weaker assumptions that the limiting distribution functions

$$\lim_{r\to\infty} P(x_t^{(r)} \leqq y) = G_t(y) \tag{3.2}$$

exist for each $t \geqq 0$, with $G_t(0+) < 1$ for some $t > 0$. Then, as the proof will show, $c_r/b_r$ tends to a finite, positive limit $c$, and there is a C. B. process $\{x_t\}$ with $x_0 = c$ such that $P(x_t \leqq y) = G_t(y)$. The stronger convergence (3.1) must then hold as well. A similar formal generalization applies to Theorem 3 below, but we shall not bother to state it in detail.

The converse is perhaps more surprising than the result above.

**Theorem 2.** *Let* $\{x_t\}$ *be any C. B. process with* $x_0 = c > 0$. *Then there exists a sequence of branching proceses* $\{Z_n\}$ *and of positive integers* $c_r \to \infty$ *such that* (3.1) *holds when* $\{x_t^{(r)}\}$ *is defined by* (1.3) *using* $a_r = 0$ *and* $b_r = c_r/c$.

Next we remove the assumption that $a_r = 0$. It is evident, looking at (1.3), that if $a_r/b_r$ tends to a finite limit we are really still in the above situation, and the limit process will be simply a translate of a C. B. process. If even a subsequence of $a_r/b_r$ has a limit the same conclusion holds. (Recall the remark at the end of section 1.) Hence these cases must be avoided if we are to obtain anything new.

**Theorem 3.** *Suppose again that* (3.1) *holds, where now it is assumed that* $a_r/b_r \to +\infty$. *Then* $\{x_t\}$ *is a process with (stationary) independent increments. If* $b_r \to \infty$, *the canonical measure governing the distribution of the increments of* $\{x_t\}$ *has support contained in* $[0, \infty)$. *If* $b_r \not\to \infty$, *however, the canonical measure is supported on the set* $\{n/b\colon n = -1, 1, 2, 3, \ldots\}$ *for some positive* $b$, *and so the law of* $x_t$ *is of the compound-Poisson type. There is no translation term in this case.*

**Theorem 4.** *Let* $\{x_t\}$ *be any additive process with* $x_0 = 0$ *and with distributions of either form described in Theorem* 3. *Then there exists a sequence of branching processes* $\{Z_n\}$, *of integers* $c_r \to \infty$, *and of constants* $b_r > 0$ *such that* (3.1) *holds when* $\{x_t^{(r)}\}$ *is defined by* (1.3) *using* $a_r = c_r$.

Remark. The example of STRATTON and TUCKER mentioned in the introduction does not quite fit into our scheme as it stands, for they assume a continuous parameter for $\{Z_n\}$ and do not contract the time scale while passing to the limit. Also, they consider processes which do not necessarily have stationary transition probabilities. However, in the stationary case it is easy to recast their limiting set-up into our form by first expanding the time axis, and so it is no accident that their limiting processes are of the sort described in Theorem 3 (the last case).

## 4. Proof of Theorems 1 and 2

In particular, the assumptions of Theorem 1 imply the existence of the limiting (one-dimensional) distributions

$$\lim_{r\to\infty} P(Z_{[rt]} \leqq y\, b_r \mid Z_0 = c_r) = G_t(y) \tag{4.1}$$

for all $t \geqq 0$, where for some $t > 0, G_t(y) = P(x_t \leqq y)$ does not have all its mass at the origin. We will first show that a limit like (4.1) must still hold when $c_r$ is replaced by $[x c_r]$, $x > 0$. In fact, (4.1) is equivalent to

$$\lim_{r\to\infty} E(e^{-\lambda Z_{rt}/b_r} \mid Z_0 = c_r) = \varphi_t(\lambda)\,, \tag{4.2}$$

where $\varphi_t$ is the Laplace-Stieltjes transform of $G_t$. (From now on we will always write $Z_{rt}$ in place of $Z_{[rt]}$.) But because $\{Z_n\}$ is a branching process (specifically, because of (1.2)), we know that

$$E(e^{-\lambda Z_{rt}/b_r} \mid Z_0 = c_r) = f_{[rt]}(e^{-\lambda/b_r})^{c_r}\,. \tag{4.3}$$

It is clear from (4.2) and (4.3) that

$$\lim_{r\to\infty} E(e^{-\lambda Z_{rt}/b_r} \mid Z_0 = [x c_r]) = \varphi_t(\lambda)^x \tag{4.4}$$

for each $x > 0$; the distribution $G_t$ is of course infinitely divisible. It is obvious that the convergence holds for $x = 0$ as well, and that it is uniform in $x$ for each fixed $\lambda$.

Next we consider the relation between $b_r$ and $c_r$. From (4.1) with $t = 0$ it is immediate that

$$\lim_{r\to\infty} c_r/b_r = c \quad \text{exists,} \qquad 0 \leqq c < \infty\,. \tag{4.5}$$

We will show that $c > 0$ unless, contrary to assumption, $G_t(0+) = 1$ for every $t > 0$.

Choose a $t > 0$ such that $G_t(0+) < 1$; accordingly $\varphi_t(\lambda) < 1$ for all $\lambda > 0$. By (4.2) and the Markov property of $\{Z_n\}$ we can write

$$\begin{aligned} \varphi_{2t}(\lambda) &= \lim_{r\to\infty} E(e^{-\lambda Z_{2rt}/b_r} \mid Z_0 = c_r) \\ &= \lim_{r\to\infty} \int_0^\infty E(e^{-\lambda Z_{rt}/b_r} \mid Z_0 = c_r y)\, dP(Z_{rt} \leqq c_r y \mid Z_0 = c_r)\,. \end{aligned} \tag{4.6}$$

But the integrand converges to $\varphi_t(\lambda)^y$ uniformly in $y$, and this passage to the limit can be interchanged with the integration:

$$\begin{aligned} \varphi_{2t}(\lambda) &= \lim_{r\to\infty} \int_0^\infty \varphi_t(\lambda)^y \, dP(Z_{rt} \leqq c_r y \mid Z_0 = c_r) \\ &= \lim_{r\to\infty} E(\exp[\log \varphi_t(\lambda)\, Z_{rt}/c_r] \mid Z_0 = c_r)\,. \end{aligned} \tag{4.7}$$

In other words, the Laplace transform of the distribution of $Z_{rt}/c_r$, given $Z_0 = c_r$, has a limit for each number $\delta$ of the form $\delta = -\log \varphi_t(\lambda)$. Because these values fill up an interval $[0, \varepsilon]$ (we use here the nondegeneracy of $G_t$), the distribution of $Z_{rt}/c_r$ must be convergent by the continuity theorem. The limiting transform is continuous as $\delta \to 0+$, so the limiting distribution has total variation 1. But $Z_{rt}/b_r$ also has a limiting distribution — one which is not concentrated at the origin — and these facts combined show that $c_r/b_r \to 0$ is impossible.

It is now clear that we can assume $c_r = c b_r$ without any change in the limiting distributions, and so we can set

$$\lim_{r\to\infty} P(Z_{rt} \leqq y b_r \mid Z_0 = [x b_r]) = P_t(x, [0, y]) . \tag{4.8}$$

We will show that $P_t$ (more accurately, the extension of $P_t$ to a measure) is a C. B. function, and that the joint distributions which it generates, taking $c$ as the initial state, are those of the limiting process $\{x_t\}$. Properties (i), (ii), and (v) of the definition are already evident. As we have seen, (4.8) is equivalent to

$$\lim_{r\to\infty} E(e^{-\lambda Z_{rt}/b_r} \mid Z_0 = [x b_r]) = \varphi_t(\lambda)^{x/c} = e^{-x\,\psi_t(\lambda)} \tag{4.9}$$

(this expression defines $\psi_t(\lambda)$), and property (iv) is apparent from the fact that $x$ appears as an exponent.

To complete the proof, we will show that $P_t$ satisfies the Chapman-Kolmogorov equation and that the two-dimensional laws of $\{x_t\}$ are those generated by $P_t$. (The verification for dimensions exceeding two is easy and will be omitted.) We attack the latter in its Laplace transformed version; it is required to prove that

$$E(e^{-\lambda x_t - \delta x_{t+s}}) = \int_0^\infty\int_0^\infty P_t(c, du)\, e^{-\lambda u}\, P_s(u, dv)\, e^{-\delta v} . \tag{4.10}$$

But since

$$\int_0^\infty P_t(x, dy)\, e^{-\lambda y} = e^{-x\psi_t(\lambda)} , \tag{4.11}$$

we easily obtain

$$\int_0^\infty\int_0^\infty P_t(c, du)\, e^{-\lambda u}\, P_s(u, dv)\, e^{-\delta v} = e^{-c\psi_t(\lambda + \psi_s(\delta))} . \tag{4.12}$$

If (4.10) holds, then, putting $\lambda = 0$ and using (4.12) yields

$$e^{-\psi_{t+s}(\delta)} = e^{-\psi_t(\psi_s(\delta))} , \tag{4.13}$$

which, as we have remarked in section 2, is the transform of the Chapman-Kolmogorov equation in this case.

It remains to prove (4.10). By definition and (4.5) the left side is

$$\lim_{r\to\infty} E(\exp[-(\lambda Z_{rt} + \delta Z_{r(t+s)})/b_r] \mid Z_0 = c b_r) , \tag{4.14}$$

and the expectation in question can be written as

$$\int_0^\infty E(e^{-\delta Z_{rs}/b_r} \mid Z_0 = x b_r)\, e^{-\lambda x}\, dP(Z_{rt} \leqq x b_r \mid Z_0 = c b_r) . \tag{4.15}$$

Arguing as before we can replace the integrand by its limit, namely

$$\exp[-x\,\psi_s(\delta) - \lambda x] ,$$

and then use (4.9) to obtain finally

$$E(e^{-\lambda x_t - \delta x_{t+s}}) = e^{-c\,\psi_t(\psi_s(\delta) + \lambda)} . \tag{4.16}$$

Because of (4.12) this is the same as (4.10) and so completes the proof. The joint laws of more than two variables can be handled in a very similar way.

We turn now to the converse, Theorem 2, in which it is required to construct a branching process such that $\{x_t^{(r)}\}$ converges to a given C. B. process $\{x_t\}$. We begin with

**Lemma 4.1.** *Any infinitely divisible distribution on* $[0, \infty)$ *is the weak limit of a sequence of distributions of the form*

$$F_d(x) = G_d^{(d)}(x\,d)\,, \tag{4.17}$$

*where* $d$ *is a positive integer,* $G_d$ *is a distribution concentrated on the nonnegative integers, and the 'exponent' means convolution.*

*Proof.* It is well known that any infinitely divisible law can be approximated by compound-Poisson distributions; that is, by distributions of the form

$$H(x) = P\left(\sum_{j=1}^{N} X_j \leqq x\right), \tag{4.18}$$

where $N, X_1, X_2, \ldots$ are independent, the $X_j$ have a common distribution and $N$ is a Poisson random variable. Perhaps the simplest way to actually construct such an approximation is to let the law of $X_j$ be the n'th root of the original law, and to choose $E(N) = n$; then it is easy to see using characteristic functions that $H$ tends to the given law as $n \to \infty$. (See [3], chapter XVII, for a particularly nice discussion.) This method has the advantage that $X_j$ is supported in $[0, \infty)$ if the given law has the same property.

Now let $d$ be a positive integer, and define the 'discretized' random variables and the distribution function

$$X_j^{(d)} = \max\left\{\frac{k}{d} : \frac{k}{d} \leqq X_j\right\}; \quad F_d(x) = P\left(\sum_{j=1}^{N} X_j^{(d)} \leqq x\right). \tag{4.19}$$

It is easy to see that as $d \to \infty$ the distribution $F_d(x)$ tends to $H$, and so $F_d$ can also be made an arbitrarily good approximation to the original i.d. law. Moreover, $F_d$ is of the form (4.17). To see this, note that $F_d$ has a '$d$'th root' given by

$$F_d^{(1/d)}(x) = P\left(\sum_{j=1}^{M} X_j^{(d)} \leqq x\right), \tag{4.20}$$

where $M$ is a Poisson variable independent of the $X_j$ with mean equal to $E(N)/d$. But since the sum of any number of $X_j^{(d)}$ is still supported on the set $\{k/d\}$, the distribution $F_d^{(1/d)}(x/d)$ is concentrated on the nonnegative integers and so can serve as $G_d$ in (4.17). This completes the proof.

Suppose the i.d. distribution of Lemma 4.1 has Laplace transform $e^{-\psi(\lambda)}$, and that the distribution $G_d$ has generating function $g(x)$ (depending on $d$, of course). Then the conclusion of the lemma is equivalent, by the continuity theorem, to

$$\lim_{d \to \infty} g(e^{-\lambda/d})^d = e^{-\psi(\lambda)}\,, \quad \lambda \geqq 0\,. \tag{4.21}$$

**Lemma 4.2.** *If* (4.21) *holds, so does*

$$\lim_{d \to \infty} g(g(e^{-\lambda/d}))^d = e^{-\psi(\psi(\lambda))}\,, \tag{4.22}$$

*as well as the corresponding statements for more than two iterations of the functions $g$ and $\psi$.*

*Proof.* From (4.21) we have

$$g(e^{-\lambda/d}) = \exp\left[\frac{-\psi(\lambda) + o(1)}{d}\right], \tag{4.23}$$

and so

$$g_2(e^{-\lambda/d})^d = g\left(\exp\left[\frac{-\psi(\lambda) + o(1)}{d}\right]\right)^d. \tag{4.24}$$

But the convergence in (4.21) is necessarily uniform in $\lambda$, and so applying (4.21) to (4.24) we obtain (4.22). The extension to more than two iterations by means of mathematical induction is routine.

We are ready to prove Theorem 2. Let $P_t$ denote the (C. B.) transition function of $\{x_t\}$, and fix an integer $r$. The distribution $P_{1/r}(1, \cdot)$ is infinitely divisible, and so can be approximated by a sequence of laws of the form (4.17). By Lemma 4.2 and the continuity theorem, the laws obtained as in (4.22) upon iterating the generating function of $G_d$ $k$ times converge to $P_{k/r}(1, \cdot)$. For any fixed number, say $r^2$, we can choose $G_d$ to make the approximation hold to within the distance $\varepsilon = 1/r$ in the Lévy metric[2] for each $k \leqq r^2$.

Let $\{p_k\}$ be probabilities corresponding to the distribution $G_d$ on the non-negative integers chosen as above, and let $\{Z_n\}$ denote the branching process having $\{p_k\}$ for its basic distribution. Let $b_r$ be the number $d$ of (4.17) for the approximating distribution we have taken. Then our construction yields

$$\mathscr{L}\{P(Z_{rt} \leqq x b_r \mid Z_0 = b_r);\ P_t(1, [0, x)]\} \leqq 1/r \tag{4.25}$$

for $t = 0, 1/r, 2/r, \ldots, r^2/r$, where '$\mathscr{L}$' denotes the Lévy distance between distribution functions. We can construct such a $\{Z_n\}$ for each $r$, and it only remains to see that (4.25) does imply the desired convergence (3.1).

The last step is easy. Fix $t$, and let $j$ be the largest integer such that $j/r \leqq t$. Since $Z_{rt}$ is constant in the time interval $(j/r, t)$ we have for all large $r$

$$\begin{aligned} &\mathscr{L}\{P(Z_{rt} \leqq x b_r \mid Z_0 = b_r);\ P_t(1, [0, x])\} \leqq \\ &\leqq 1/r + \mathscr{L}\{P_{j/r}(1, [0, x]);\ P_t(1, [0, x])\}, \end{aligned} \tag{4.26}$$

using (4.25) and the triangle inequality for $\mathscr{L}$. But for each $t$ the last term tends to 0 with $r$ because $j/r \to t$ and $P_t$ is stochastically continuous by Lemma 2.3. This proves (4.1) with $G_t(y) = P_t(1, [0, y])$ and $c_r = b_r$. As we have seen in the proof of Theorem 1, the extension to $c_r = c b_r$ and the convergence of the joint distributions now follow automatically. This completes the proof of Theorem 2.

## 5. Proof of Theorems 3 and 4

From the hypotheses of Theorem 3 we have in particular that the weak limits

$$\lim_{r\to\infty} P(Z_{rt} - a_r \leqq y b_r \mid Z_0 = c_r) = H_t(y) \tag{5.1}$$

---

[2] The argument to follow uses only the fact that the topology of weak convergence for distribution functions *can* be metrized; it does not matter how it is done.

exist for each $t \geqq 0$ with $H_t(y) = P(x_t \leqq y)$ a probability distribution, and $a_r/b_r \to +\infty$. Setting $t = 0$ we see at once that

$$\lim_{r\to\infty} \frac{c_r - a_r}{b_r} = c \tag{5.2}$$

exists, and that $c = x_0$. Clearly, adding $o(b_r)$ to $a_r$ has no effect on the limits, while adding a multiple of $b_r$ simply represents a translation of the limiting process. It follows that there is no loss of generality in assuming $a_r = c_r$ and we will do so below; consequently, of course, we always have $x_0 = 0$.

The proof that $\{x_t\}$ has independent increments is very similar to that of Theorem 5.1 in [6]; the assumption made throughout that paper that $\{p_k\}$ is independent of $r$ played no role in the case of that theorem. However, both because one point in that proof was treated with insufficient care and because we wish to go further, a complete discussion here seems in order. Expressing (5.1) in terms of characteristic functions, and writing the characteristic function of $Z_n$ in terms of $f(x)$, we obtain

$$\varphi_t(\lambda) = E(e^{i\lambda x_t}) = \lim_{r\to\infty} e^{-i\lambda c_r/b_r} f_{[rt]}(e^{i\lambda/b_r})^{c_r}. \tag{5.3}$$

This can at once be recast in the form

$$f_{[rt]}(e^{i\lambda/b_r}) = e^{i\lambda/b_r}[\varphi_t(\lambda) + o(1)]^{1/c_r}; \tag{5.4}$$

since $\varphi_t$ is an infinitely divisible characteristic function the meaning of

$$\log[\varphi_t(\lambda) + o(1)]$$

is uniquely determined for large $r$ in the usual way.

We will need (as in [6]) a simple extension of (5.4), namely

$$f_{[rt]}\left(\exp\left[\frac{i\lambda}{b_r} + \frac{\delta_r}{c_r}\right]\right) = \exp\left[\frac{i\lambda}{b_r} + \frac{\delta + \log\varphi_t(\lambda) + o(1)}{c_r}\right], \tag{5.5}$$

where $\delta_r$ is any sequence of complex numbers with nonpositive real parts[3] which tend to a limit $\delta$. Assuming this for the moment, we will verify that

$$\lim_{r\to\infty} E\left(\exp\left[\frac{i\lambda(Z_{rt} - c_r)}{b_r} + \frac{i\sigma(Z_{r(t+s)} - Z_{rt})}{b_r}\right]\middle| Z_0 = c_r\right) = \varphi_t(\lambda)\,\varphi_s(\sigma). \tag{5.6}$$

A corresponding formula is similarly derived for the joint distributions of more than two variables, and the result shows that the limiting process $x_t$ is indeed an additive one. To obtain (5.6) we express the left side in terms of the generating function $f$ using the fact, readily obtained from (1.2), that

$$E(x^{Z_n} y^{Z_{n+m}} | Z_0 = j) = f_n(x f_m(y))^j. \tag{5.7}$$

The result is

$$E(\exp[i\lambda x_t + i\sigma(x_{t+s} - x_t)]) = \lim_{r\to\infty} e^{-i\lambda c_r/b_r} f_{[rt]}(e^{i(\lambda-\sigma)/b_r} f_{[rs]}(e^{i\sigma/b_r}))^{c_r}. \tag{5.8}$$

---

[3] The condition $\mathrm{Re}(\delta_r) \leqq 0$ was inadvertently omitted from the discussion in [6], and the proof there of the analogue of (5.5) was incomplete.

But replacing $f_{[rs]}$ by its expression in (5.4), letting

$$\delta_r = \log E(\exp[i\lambda x_t^{(r)}]) = \log[\varphi_t(\lambda) + o(1)]$$

and using (5.5), the limit is easily seen to be just the right hand side of (5.6). We omit details of the extension to the higher dimensional distributions.

We will obtain formula (5.5), among other things, using the following lemma, whose straightforward proof is omitted.

**Lemma 5.1.** *Suppose that, for each $n$, a pair $(A_n, B_n)$ of (possibly dependent) random variables is defined on some probability space. Suppose too that $A_n$ is real, that for each real $\lambda$*

$$\lim_{n\to\infty} E(e^{i\lambda A_n}) = \varphi(\lambda) \tag{5.9}$$

*exists, $\varphi$ a characteristic function, that* $\mathrm{Re}(B_n) \leqq M$ *a. s. for all $n$, and that $B_n \to 0$ in distribution. Then*

$$\lim_{n\to\infty} E(e^{i\lambda A_n + B_n}) = \varphi(\lambda). \tag{5.10}$$

To derive (5.5) from Lemma 5.1, set

$$A_r = \frac{Z_{rt} - c_r}{b_r}, \qquad B_r = \delta_r \frac{Z_{rt} - c_r}{c_r}, \tag{5.11}$$

where $Z_n$ is the $r$'th branching process with $Z_0 = c_r$. The convergence (5.1) (with $a_r = c_r$) then yields (5.9), as well as the fact that $B_r \to 0$ in distribution. (We use here the assumption that $c_r/b_r \to \infty$.) Finally, $Z_{rt} \geqq 0$ and $\delta_r \to \delta$ with $\mathrm{Re}(\delta_r) \leqq 0$ mean that $\mathrm{Re}(B_r)$ is bounded above uniformly in $r$. It follows, then, that (5.10) holds for the choices (5.11) with $\varphi_t(\lambda)$ in the role of $\varphi(\lambda)$, and this almost immediately implies that

$$\lim_{r\to\infty} E\left(\exp\left[i\lambda \frac{Z_{rt} - c_r}{b_r} + \delta_r \frac{Z_{rt}}{c_r}\right] \middle| Z_0 = c_r\right) = e^{\delta}\varphi_t(\lambda). \tag{5.12}$$

Expressing the expectation in terms of the generating function $f$, (5.12) can be transformed into (5.5) in the same way that (5.3) became (5.4). This completes the proof that $\{x_t\}$ is an additive process; as yet the question of whether or not $b_r \to \infty$ has played no role.

Supose next that $b_r \to \infty$ does hold. To study the laws of the process $\{x_t\}$ it is only necessary to consider one value of $t$, say $t = 1$. We know that

$$\begin{aligned} P(x_1 \leqq y) &= \lim_{r\to\infty} P(Z_r - c_r \leqq y b_r \mid Z_0 = c_r) \\ &= \lim_{r\to\infty} P\left(\sum_{j=1}^{c_r} \frac{X_j - 1}{b_r} \leqq y\right), \end{aligned} \tag{5.13}$$

where the $X_j$ are independent variables representing the number of individuals in the $r$'th generation of the process $\{Z_n\}$ who are 'descended' from each of the $c_r$ original 'ancestors'. From the theory of limit laws for triangular arrays of random variables (for which now the most convenient source is [3]), a necessary condition for (5.13) is the 'proper' convergence

$$\lim_{r\to\infty} c_r x^2 G_r(dx) = M(dx), \tag{5.14}$$

where $G_r$ is the common distribution of $(X_j - 1)/b_r$ and $M$ is Feller's form of canonical measure for the limit law. But $c_r x^2 G_r(dx)$ has no mass whatever to the left of $1/b_r$, and so if $b_r \to \infty$ it is obvious that $M$ must vanish on $(-\infty, 0)$. (Although $M$ is not the same as Khintchine's or Lévy's canonical measure, the fact of vanishing on the negative axis is equivalent for the three forms.)

Finally, suppose that $b_r \not\to \infty$, and select a subsequence converging to $b$. Clearly $b = 0$ is impossible if the limit in (5.13) is a probability distribution except in the degenerate case, which we now exclude. But then the measures $c_r x^2 G_r(dx)$ have support $\{-1/b_r, 1/b_r, 2/b_r, \ldots\}$ and so the limit $M$ must have support $\{k/b\colon k = -1, 1, 2, 3, \ldots\}$; the limit law must, accordingly, be of the compound Poisson type. It is also clear, looking at the subsequence such that $b_r \to b$, that the limit laws $P(x_t \leqq y)$ have support on $\{k/b\colon k = 0, \pm 1, \pm 2, \ldots\}$ for all $t$, so there can be no steady drift term. This completes the proof of Theorem 3.

We turn to the converse, Theorem 4; the pattern of the proof will resemble that for Theorem 2. We begin with two lemmas which jointly will play a role analogous to that of Lemma 4.1.

**Lemma 5.2.** *Let $F$ be any infinitely divisible law with its canonical measure supported on $[0, \infty)$. Then $F$ is the weak limit of a sequence of laws of the form*

$$F_d(x) = G_d^{(d)}(b_d x + d), \tag{5.15}$$

*where $b_d > 0$, $d$ is a positive integer, $d/b_d \to \infty$, and $G_d$ is a distribution concentrated on the nonnegative integers.*

*Proof*. The characteristic function of $F$ can be written in the form

$$\log \varphi(\lambda) = i a \lambda - \frac{\sigma^2 \lambda^2}{2} + \int_{0+}^{\infty} \frac{e^{i\lambda x} - 1 - i\lambda \sin x}{x^2}\, dM(x), \tag{5.16}$$

where the measure $M$ is finite on finite intervals, and satisfies $\int_1^\infty x^{-2} dM < \infty$ [3]. The right side of (5.16) can be approximated arbitrarily well by extending the integral only from $\varepsilon > 0$ to $\infty$. In that case, the approximating characteristic function is the product of a normal part (having variance $\sigma^2$ and mean $\mu = a - \int_\varepsilon^\infty x^{-2} \sin x\, dM$) with a compound Poisson part. Therefore $F$ is the limit of distributions of the form

$$K(x) = P\left(\sum_{j=1}^{N} X_j \leqq x\right) * \Phi_{\mu,\sigma}(x), \tag{5.17}$$

where '$*$' means convolution, $X_j$ and $N$ are as in (4.18), and $\Phi_{\mu,\sigma}$ is the normal distribution function.

Choose a 'large' constant $b > 0$. The first factor in (5.17) — the compound Poisson part — can be 'discretized' just as in the proof of Lemma 4.1. We write the result

$$H_b(x) = P\left(\sum_{j=1}^{N} X_j^{(b)} \leqq x\right), \tag{5.18}$$

and note that $H_b$ is supported on the set $\{k/b\colon k \geqq 0\}$ and that $H_b$ has a $d$'th root of the same form as itself for any $d$. Hence $H_b^{(1/d)}(x/b) = G_1(x)$ is a distribution on the nonnegative integers with the property that $G_1^{(d)}(bx)$ is a good approximation to the first factor of $K(x)$.

Next we approximate $\Phi$ by a binomial distribution. We choose one which can be thought of as the law of the sum of $d$ independent random variables taking the values $\pm 1/b$. The number $d$ and the probabilities with which each variable chooses $+$ or $-$ are then taken in such a way that the mean and variance of the sum tend to $\mu$ and $\sigma^2$ respectively; this can always be done and entails $d/b \to \infty$ as $b \to \infty$. We denote by $G_2(x)$ the distribution function of one of the summands multiplied by $b$ and translated so that its range becomes $\{0, 2\}$. The binomial which approximates $\Phi$ can then be written as $G_2^{(d)}(bx + d)$.

Finally we set $G_1 * G_2 = G_d$; since both $G_1$ and $G_2$ are supported on the nonnegative integers, $G_d$ has the same property. But

$$G_d^{(d)}(bx + d) = G_1^{(d)}(bx) * G_2^{(d)}(bx + d), \tag{5.19}$$

and the two factors approximate respectively the factors of $K$ in (5.17). Since convolution is continuous, $F_d(x) = G_d^{(d)}(bx + d)$ approximates $K$ which in turn is close to $F$; this proves the lemma.

**Lemma 5.3.** *Let $F$ be a compound Poisson law of the form* (4.18), *where the variables $X_j$ take values $\{-1, 1, 2, 3, \ldots\}$. Then $F$ is the weak limit of a sequence of laws of the form*

$$F_d(x) = G_d^{(d)}(x + d), \tag{5.20}$$

*where $G_d$ is concentrated on the nonnegative integers and $d \to \infty$.*

*Proof.* The characteristic function of $F$ is of the form

$$\log \varphi(\lambda) = \mu\left[\sum_{n=-1}^{\infty} e^{in\lambda} a_n - 1\right], \tag{5.21}$$

where $\mu = E(N)$ and $a_n$ are probabilities with $a_0 = 0$. On the other hand, the characteristic function of a law of the form $F_d$ must satisfy

$$\log \varphi_d(\lambda) = d \log \psi_d(\lambda), \tag{5.22}$$

where $\psi_d(\lambda) = \sum_{n=-1}^{\infty} e^{in\lambda} \alpha_n$ is the characteristic function of $G_d(x + 1)$. Let us choose $\alpha_n = \mu\, a_n/d$ for $n \neq 0$; $\alpha_0 = 1 - \sum_{n \neq 0} \alpha_n$. (When $d$ is large $\alpha_0 > 0$, so $\psi_d$ is a characteristic function of the desired form.) We then have

$$d[\psi_d(\lambda) - 1] = \mu\left[\sum_{n=-1}^{\infty} e^{in\lambda} a_n - 1\right],$$

and because $\log \psi_d(\lambda)$ is asymptotic to $\psi_d(\lambda) - 1$ as $d \to \infty$ it follows that (5.22) converges to (5.21). This proves the lemma.

The next step is to obtain a suitable replacement for Lemma 4.2.

**Lemma 5.4.** *Suppose that, for each $r$, $X_j$ are independent random variables taking integer values $\geqq 0$ and having a common generating function $f(x)$. Suppose also that for some sequences $b_r > 0$, $c_r \to \infty$ with $c_r/b_r \to \infty$,*

$$\lim_{r\to\infty} E\left(\exp\left[i\lambda \frac{\sum_{j=1}^{c_r}(X_j - 1)}{b_r}\right]\right) = \varphi(\lambda) \tag{5.23}$$

*exists, $\varphi$ a characteristic function. Let $Y_j$ be independent variables with generating function $f_k(x)$ for any $k \geqq 1$. Then*

$$\lim_{r\to\infty} E\left(\exp\left[i\lambda \frac{\sum_{j=1}^{c_r}(Y_j - 1)}{b_r}\right]\right) = \varphi(\lambda)^k. \tag{5.24}$$

*Proof.* We will consider the case $k = 2$; the extension by induction is easy. Suppose that $\delta_r \to \delta$, where $\mathrm{Re}(\delta_r) \leqq 0$. Then using Lemma 5.1 and arguing exactly as in the proof of (5.5) we can extend (5.23) to

$$\lim_{r\to\infty} E\left(\exp\left[i\lambda \frac{\sum_{j=1}^{c_r}(X_j - 1)}{b_r} + \delta_r \frac{\sum_{j=1}^{c_r} X_j}{c_r}\right]\right) = e^{\delta}\varphi(\lambda). \tag{5.25}$$

This result can be written in terms of $f$ as

$$f(e^{i\lambda/b_r} e^{\delta_r/c_r}) = e^{i\lambda/b_r} \exp\left[\frac{\delta + \log\varphi(\lambda) + o(1)}{c_r}\right]. \tag{5.26}$$

But in terms of $f$ the expectation on the left in (5.24) (with $k = 2$) becomes

$$e^{-i\lambda c_r/b_r} f(f(e^{i\lambda/b_r}))^{c_r}, \tag{5.27}$$

and expressing first the inner function and then the outer by means of (5.26), with $\delta_r = 0$ and $\delta_r = \log\varphi(\lambda) + o(1)$ in the two cases respectively, it is very easy to see that the limit (5.24) is $\varphi(\lambda)^2$.

Finally we come to Theorem 4 itself. The rest of the argument is entirely similar to the proof of Theorem 2, and will only be sketched. Let $H_t(x)$ be the law of the position at time $t$ of any additive process $\{x_t\}$ which has $x_0 = 0$ and satisfies either of the two conditions on its canonical measure which were specified in Theorem 3. (In the second case let $b = 1$.) Fix a 'large' integer $r$, and approximate $H_{1/r}$ by the sort of distribution provided in either Lemma 5.2 or 5.3, whichever is applicable. Then use the distribution $G_d$ of the lemma as the basic distribution in constructing a branching process $\{Z_n\}$. By Lemma 5.4 the branching approximation, if made good enough at $t = 1/r$, will be good at $2/r, \ldots, r^2/r$ as well. Since an additive process is stochastically continuous, the convergence

$$\lim_{r\to\infty} P(Z_{rt} - c_r \leqq y b_r \mid Z_0 = c_r) = H_t(y) \tag{5.28}$$

follows for all $t$ by the same reasoning as before. Finally we only need remark that, by the proof of Theorem 3, (5.28) is enough to establish automatically that the joint distributions of $\{x_t^{(r)}\}$ converge to those of $\{x_t\}$. This completes the proof of the theorem.

## Bibliography

1. DYNKIN, E. B.: Markov processes I. Berlin-Heidelberg-New York: Springer 1965.
2. FELLER, W.: Diffusion processes in genetics. Proc. Second Berkeley Sympos. math. Statist. Probability, 227–246 (1951).
3. — An introduction to probability theory and its applications II. New York: Wiley 1966.
4. HARRIS, T. E.: The theory of branching processes. Berlin-Göttingen-Heidelberg: Springer 1963.
5. JIŘINA, M.: Stochastic branching processes with continuous state space. Czechosl. math. J. 8, 292–312 (1958).
6. LAMPERTI, J. W.: Limiting distributions for branching proceses. Proc. Fifth Berkeley Sympos. math. Statist. Probability, in press.
7. SEVASTYANOV, B. A.: Transient phenomena for branching stochastic processes. Theor. Probab. Appl. 4, 113–128 (1959).
8. STRATTON, H. H., and H. G. TUCKER: Limit distributions of a branching stochastic process. Ann. math. Statistics 35, 557–565 (1964).

Dartmouth College and University of Cambridge