Søren Johansen

The Imbedding Problem for Markov Branching Processes

Preprint
October
1973
9

Institute of Mathematical Statistics
University of Copenhagen

Søren Johansen

The imbedding problem for Markov branching processes.

Preprint 1973 No. 9

Institute of Mathematical Statistics
University of Copenhagen.

October 1973

1. Introduction and Summary

The purpose of this paper is to characterize the stochastic matrices that can occur in non-homogeneous Markov branching processes with no deaths.

We find that most of the results from [5] for finite state Markov chains can be generalized to these branching processes with appropriate modifications of the definitions.

The simplest branching process is a Galton-Watson process which is a Markov process with discrete time and state space $S = \{0,1,2,...\}$. The transition probabilities $P = (p_{ik}, i \in S, k \in S)$ satisfy the condition

$$p_{ik} = \Sigma p_{1k_1} \cdot \ldots \cdot p_{1k_i} \tag{1.1}$$

where the summation is over the set of indices $k_1,...,k_i$ for which $k_1 + ... + k_i = k$.

Thus the i'th row is the i-fold convolution of the probability measure $\{p_{1k}, k \in S\}$.

A continuous time Markov branching process also has the properties that the transition probabilities satisfy the condition (1.1), see Harris [5], p. 97.

We shall here only consider branching processes for which $p_{i0} = 0$, $i = 1,2,...$, i.e. processes for which no deaths can occur. Thus we consider the set B of stochastic matrice $P = (p_{ik}, k \in N, i \in N)$ for which conditions (1.1) holds, where $N = (1,2,...)$.

Let us define

$$m(P) = \sum_{k=1}^{\infty} k p_{1k}, \quad P \in B$$

and call P regular if $m(P) < \infty$.

For the present purpose we shall use the following definitions:

1.1. Definition. A Markov branching process is a family

$$\{P(s,t),\ 0 \leq s \leq t < t_0 \leq \infty\} \tag{1.2}$$

of matrices from B which satisfy the Chapman-Kolmogorov equations

$$P(s,t) = P(s,u)P(u,t),\ 0 \leq s \leq u \leq t \tag{1.3}$$

and the initial condition

$$P(s,s) = I \tag{1.4}$$

together with the regularity condition

$$m(P(s,t)) \text{ is finite and continuous} \tag{1.5}$$

1.2. Definition. A stochastic matrix $P \in B$ is called imbeddable if there exists a Markov branching process (1.2) such that

$$P(0,1) = P.$$

The imbedding problem is then to characterize the matrices that can be imbedded.

It is easily seen that B is a semigroup under matrix multiplication and that the mappings

$$P \to \sum_{k=1}^{\infty} kp_{1k} = m(P)$$

and

$$P \to p_{11}$$

are multiplicative functions on B.

Since the matrix P is given by its first row it is also given by the probability generating function

$$f_P(z) = \sum_{k=1}^{\infty} p_{1k} z^k, \quad 0 \leq z \leq 1.$$

The set of probability generating functions C is a semigroup under composition and it is easily seen that

$$f_{P_1 P_2}(z) = f_{P_1}(f_{P_2}(z)) = (f_{P_1} \cdot f_{P_2})(z)$$

which shows that the mapping from B onto C

$$P \to f_P$$

is a homomorphism.

Let $f \in C$ and $f(z) = \sum_{k=1}^{\infty} z^k p_k$ then we define

$$Df(z) = \sum_{k=1}^{\infty} k z^{k-1} p_k$$

and we call f regular if $Df(1) < \infty$.

The mappings

$$f \to Df(0)$$

and

$$f \to Df(1)$$

are multiplicative functions from C to [0,1] and [1,∞] respectively.

Let us also remark that C is an example of a convex semigroup, which B is not, and that the extremepoints have the form

$$f_k(z) = z^k \qquad k = 1,2,3,\ldots$$

In the following we shall work mainly with C but in section 5 we give yet another representation of B.

We can now replace Definition 1.2 by

1.3. Definition. The regular function $f \in C$ is called imbeddable if there exists a family $\{f_{s,t},\ 0 \leq s \leq t < t_0 \leq \infty\}$ of regular functions in C such that

$$f_{s,t} = f_{s,u} \cdot f_{u,t}, \quad 0 \leq s \leq u \leq t, \tag{1.6}$$

$$f_{s,s} = e, \tag{1.7}$$

$$Df_{s,t}(1) \text{ is finite and continuous,} \tag{1.8}$$

$$f_{0,1} = f \tag{1.9}$$

We call f imbeddable in a homogeneous family if $f_{s,t}$ can be chosen to depend only on t - s.

It is convinient to work with uniform convergence of functions in C but we shall also define strong convergence as uniform convergence of the derivatives. The following result will be used repeatedly:

<u>1.4. Theorem</u>. Let $f_n \in C$ converge pointwise to a function f. If f is continuous at z = 1 then $f \in C$ and f_n converges uniformly to f. If further $Df_n(1)$ converges to Df(1) then f_n converges strongly to f.

A function of the form

$$h(z) = \sum_{k=1}^{\infty} a_k z^k, \quad 0 \leq z \leq 1$$

where

$$a_1 \leq 0, \quad a_k \geq 0, \quad k = 2,\ldots, \quad \sum_{k=1}^{\infty} a_k = 0$$

is called an intensity function and it is called regular if

$$Dh(1) < \infty.$$

It is easily seen that h is a convex function and strictly convex unless $h \equiv 0$. We easily find the inequalities

$$Dh(0) \leq h(z) \leq 0 \tag{1.10}$$

$$|Dh(z)| \leq Dh(1). \tag{1.11}$$

It is important to notice that if $f \in C$ then $f - e$ is an intensity function, where $e(z) = z$. We then find from (1.10) and (1.11) that

$$|f - e| \leq 1-Df(0) \leq Df(1) - 1 \tag{1.12}$$

where we use the norm

$$|f - e| = \sup_{0 \leq z \leq 1} |f(z)-z|.$$

we shall finally use the following

1.5. Definition. By a regular measurable (locally integrable) intensity valued function h_t we shall understand a mapping from $[0,t_0[\times [0,1]$ into R which satisfies:

$$h_t(\cdot) \text{ is a regular intensity} \tag{1.13}$$

function for all $t \in [0,t_0[$, and

$$h_{\cdot}(z) \text{ is measurable (locally integrable)} \tag{1.14}$$

for all $z \in [0,1]$.

The set of intensity functions is a convex cone and the extremal elements are proportional to

$$h_k(z) = z^k - z, \; k = 1,2,\ldots . \tag{1.15}$$

We shall now give a short account of the present paper and some of the previous work in this area.

Section 2 contains an extension of the results in [3] and [4] where the idea is to change the time scale of the family imbedding f in such a way that $f_{\cdot,t}$ and $f_{s,\cdot}$ become absolutely continuous and one can then show that they are solutions to the backward and forward Kolmogorov differential equations. The main idea is to use

$$\ln Df_{0,t}(1) \tag{1.16}$$

as the new time scale. The backward Kolmogorov equation was studied by Carathéodory [1] who proved existence and uniqueness of the solution.

In section 3 we construct a product integral of a regular measurable intensity valued function. The approach taken is a generalisation of the method of Dobrušin [2] and the integral obtained is a special case of the integral studied by Neuberger [10].

It is shown that the integral provides a solution to the backward and forward Kolmogorov equation.

Section 4 contains a discussion of the imbedding problem and it is shown that the imbeddable functions can be characterized as being infinitely factorizable or as the limits of triangular null arrays.

It is also shown that finite compositions of functions generated by the extremal intensity functions can approximate any imbeddable function but the problem of a Bang-Bang representation [7] has not been solved.

Finally we find again the relation between imbeddable functions and the Kolmogorov equations, thereby proving again the results of section 2, but without the condition that $f_{s,t} = e$ if and only if $s = t$.

As an application of these characterization we find the result that any imbeddable function satisfies the inequality

$$Df(0)Df(1) \geq 1. \tag{1.17}$$

This inequality easily implies that the truncated binomial and Poisson distributions are not imbeddable.

The results in section 4 are presented in terms of probability generating functions but they clearly have a probabilistic meaning as follows:

If $f_1,\ldots,f_n$ are functions in C then we can interprete

f_i as generating the offspring distribution in the i'th generation, and

$$f_n \cdot f_{n-1} \cdot \ldots \cdot f_1$$

is then the generating function for the distribution of the size of the n'th generation.

The condition

$$0 \leq Df_i(1) - 1 \leq \varepsilon$$

then means that the expected number of offspring in each generation is close to 1 or that the increase in the population is very small.

That f is infinitely factorizable means that for any n we can think of f as the probability generating function of the n'th generation each of which has an expected size close to 1.

Finally the triangular null arrays are the mathematical formulation of the following problem: What happens when one observes a population after a long time, i.e. after many generations, where in each of the generations the expected value is close to 1.

The answer is like in the classical theory of limit theorems for sums of independent random variables that the possible limit distributions can be characterized as being infinitely factorizable. The product integral representation is then an equivalent of the Lévy-Ȟinčin representation.

The last section contains a reformulation of the results. The main idea is to use Choquet's integral representation theorem to represent any stochastic matrix over N as an integral over the extremal stochastic matrices E.

One can easily see that E is a semigroup and it turns out that one can identify the matrices on E that represent matrices in B. Thus the results about imbeddable matrices in B can be translated into results about processes with independent increments on E.

Most of the results of the paper are generalizations of the results on [4] and [6]. The presentation is intended to be selfcontained even though most of the results on the product integral have been given by Neuberger [10]. The characterization of the imbeddable functions as being infinitely factorizable and as solutions to differential equations are given by Loewner [9] and Pommerenke [11] in the context of Schlicht mappings of the unit disc into itself. The derivation of the differential equations by changing the time scale was given by Goodman [12].

2. Change of time scale and differential equations.

We shall in this section consider the family $\{f_{s,t},\ 0 \leq s \leq t < t_0 \leq \infty\}$ of probability generating function satisfying the conditions

$$f_{s,t} = f_{s,u} \cdot f_{u,t} \tag{2.1}$$

$$f_{s,t} = e \Leftrightarrow s = t \tag{2.2}$$

$$Df_{s,t}(1) \text{ is finite and continuous in } (s,t) \tag{2.3}$$

We shall first prove the basic result on change of time scale

2.1. Lemma. Under the assumptions (2.1), (2.2) and (2.3) there exists a change of time scale such that the functions $f_{\cdot,t}(z)$ satisfy a Lipschitz condition.

A possible choice of time scale is

$$\phi(t) = \ln Df_{0,t}(1) \tag{2.4}$$

or

$$\psi(t) = -\ln Df_{o,t}(0). \tag{2.5}$$

Proof. From (2.1) we find that ϕ and ψ are increasing and (2.2) implies that they are strictly increasing whereas (2.3) implies continuity. Thus any of them can serve as a change of time scale.

Now (2.1) implies that

$$f_{s-u,t} - f_{s+v,t} = (f_{s-u,s+v} - e) \cdot f_{s+v,t}, \tag{2.6}$$

for $0 \leq s-u \leq s+v \leq t < t_o$, and the inequality (1.12) gives that

$$|f_{s-u,t} - f_{s+v,t}| \leq 1 - Df_{s-u,s+v}(0) \leq Df_{s-u,s+v}(1) - 1.$$

Now let us change the time scale by means of ϕ, then

$$Df_{s-u,s+v}(1) = e^{u+v}$$

and we get the Lipzchitz condition

$$\left|\frac{f_{s-u,t} - f_{s+v,t}}{u+v}\right| \leq \frac{e^{u+v}-1}{u+v} \leq \frac{e^{t}-1}{t}.$$

A similar condition can be derived if ψ is used as the time scale.

We shall now apply Lemma 2.1 to the Chapman-Kolmogorov equations (2.1) in order to obtain differentiability of $f_{\cdot,t}(z)$ and in order to derive the Kolmogorov equations.

2.2. Theorem. Under the assumptions (2.1), (2.2), and (2.3) there exists a change of time scale such that the intensities

$$h_s = \lim_{u\downarrow 0, v\downarrow 0} (f_{s-u,s+v} - e)(u+v)^{-1} \tag{2.7}$$

exist for $s \notin N$, where N is a null set for Lebesgue measure.

The function h_s is a regular measurable intensity valued function and $Dh.(1)$ is locally integrable.

Further the derivatives

$$\partial_s f_{s,t} = \lim_{u\downarrow 0, v\downarrow 0} (f_{s+v,t} - f_{s-u,t})(u+v)^{-1} \tag{2.8}$$

exist for $s \notin N$ and the function $f_{\cdot,t}(z)$ satisfy the backward Kolmogorov equation:

$$\partial_s f_{s,t}(z) = -h_s(f_{s,t}(z)), \quad s \notin N \tag{2.9}$$

or

$$z - f_{s,t}(z) = - \int_o^t h_u(f_{u,t}(z))du. \tag{2.10}$$

<u>Proof.</u> Let us first choose

$$\phi(t) = \ln Df_{o,t}(1)$$

as our time scale, then we can use the relation

$$Df_{s,t}(1) = e^{t-s}.$$

From Lemma 2.1 we get that $f_{\cdot,t}(z)$ is absolutely continuous and hence that the limit

$$\partial_s f_{s,t}(z) = - \lim_{u\downarrow 0, v\downarrow 0} (f_{s-u,t}(z) - f_{s+v,t}(z))(u+v)^{-1} \tag{2.11}$$

exists for $s \notin N_{t,z}$, where $N_{t,z}$ is a null set.

Now define

$$g_{s,u,v} = (f_{s-u,s+v} - e)(Df_{s-u,s+v}(1)-1)^{-1} + e \tag{2.12}$$

From the inequality (1.12) it follows that $g_{s,u,v}$ is a probability generating function and it is seen that $Dg_{s,u,v}(1) = 2$.

We then consider

$$g_{s,t,u,v} = g_{s,u,v} \cdot f_{s+v,t} \tag{2.13}$$

$$= (f_{s-u,t} - f_{s+v,t})(e^{u+v}-1)^{-1} + f_{s+v,t}.$$

From (2.11) it follows that the limit

$$g_{s,t}(z) = \lim_{u\downarrow 0, v\downarrow 0} g_{s,t,u,v}(z) = -\partial_s f_{s,t}(z) + f_{s,t}(z) \tag{2.14}$$

exists for z rational and $s \notin N = \bigcup N_{t,z}$ where the union is taken over rational z. We now want to prove that (2.14) exists for all $z \in [0,1]$.

Take therefore z' and z'' in [0,1] and use the convexity of g to obtain

$$|g_{s,t,u,v}(z') - g_{s,t,u,v}(z'')|$$

$$\leq D g_{s,t,u,v}(1)|z' - z''|$$

$$= 2Df_{s+v,t}(1)\ |z'-z''| \leq 2e^{t-s}|z'-z''|.$$

Thus for fixed s and t the family

$$\{g_{s,t,u,v},\ u \geq 0,\ v \geq 0,\ u+v \leq t\}$$

is equicontinuous on [0,1] and hence the limit (2.14) exists for all $z \in [0,1]$ and defines a continuous function $g_{s,t}$ on [0,1]. By Theorem 1.4 $g_{s,t}$ is a probability generating function and the convergence in (2.14) is uniform in $z \in [0,1]$.

Now choose a sequence $t_n \to \infty$ and define $N = \bigcup_{n=1}^{\infty} N$.

For $s \notin N$ we can find $t = t_n > s$ and then

$$|g_{s,u,v} \cdot f_{s+v,t} - g_{s,u,v} \cdot f_{s,t}|$$

$$\leq D g_{s,u,v}(1)|f_{s+v,t} - f_{s,t}|$$

$$\leq 2v(e^t-1)t^{-1}$$

which proves that the limit

$$\lim_{u\downarrow 0, v\downarrow 0} g_{s,u,v} \cdot f_{s,t}$$

exists and hence that

$$g_s = \lim_{u\downarrow 0, v\downarrow 0} g_{s,u,v}$$

exists and defines a probability generating function g_s.

If we define the intensity function

$$h_s = g_s - e$$

then (2.7) have been proved.

From the inequality

$$(1-g_{s,u,v}(z))(1-z)^{-1} \leq Dg_{s,u,v}(1) = 2$$

it follows for $u \downarrow 0$, $v \downarrow 0$ that

$$Dh_s(1) + 1 \leq 2$$

and hence that h_s is regular and that $Dh_s(1)$ is locally integrable.

In order to prove (2.8) we consider again (2.6). By dividing with $u + v$ and letting $u \downarrow 0$ and $v \downarrow 0$ we get for $s \notin N$ that (2.8), (2.9), and (2.10) hold.

<u>2.3. Proposition</u>. Let $\{f_{s,t}, 0 \leq s \leq t < t_0\}$ satisfy conditions (2.1), (2.3) and the equation (2.10) for some regular intensity valued function h_s such that $Dh.(1)$ is locally integrable then

$$Df_{s,t}(1) = \exp\int_s^t Dh_u(1)du. \tag{2.15}$$

<u>Proof</u>. From (2.10) we get

$$1 - Df_{s,t}(z) = - D\int_s^t h_u(f_{u,t}(z))du. \tag{2.16}$$

Now

$$|h_u(f_{u,t}(z')) - h_u(f_{u,t}(z''))|$$

$$\leq Dh_u(1)\ Df_{u,t}(1)|z'-z''|$$

$$\leq Dh_u(1)\ e^t|z'-z''|$$

which implies that the differentiation and the integration in (2.16) can be interchanged.

Thus

$$1 - Df_{s,t}(z) = -\int_s^t Dh_u(f_{u,t}(z))Df_{u,t}(z)du.$$

for z = 1 we get

$$1 - Df_{s,t}(1) = -\int_s^t Dh_u(1)Df_{u,t}(1)du$$

which proves (2.15).

2.4 . Corollary. The convergence in (2.7) and (2.8) is strong.

Proof. Since (2.8) follows easily from (2.7) we shall only prove (2.7). By Theorem 1.3 it is enough to prove that

$$Dh_s(1) = \lim_{u\downarrow 0, v\downarrow 0} (Df_{s-u,s+v}(1)-1)(u+v)^{-1}$$

but this clearly follows from (2.15).

3. The product integral.

We shall here give a short account of the product integral that exists in the semigroup C.

The integral is a special case of the integral considered by Neuberger [10] but we shall construct it here in the context of generating functions and prove the properties we need for the imbedding problem.

We shall use the notation

$$\prod_{i=1}^{n} f_i = f_1 \cdot \ldots \cdot f_n .$$

First we prove the basic lemma.

3.1. Lemma. Let $h = h_1 + \ldots + h_n$ be regular intensity functions such that $Dh_i(o) > -1$, $i = 1,\ldots,n$, then

$$\left| \prod_{i=1}^{n} (e+h_i) - (e + \sum_{i=1}^{n} h_i) \right| \leq (-Dh(0))Dh(1). \tag{3.1}$$

Proof. For $n = 1$ this is rather obvious and we assume that it has been proved for $n = k$.

Then

$$a(z) = \left(\prod_{i=1}^{k+1} (e+h_i) - (e + \sum_{i=1}^{k+1} h_i) \right)(z)$$

$$= \prod_{i=1}^{k} (e+h_i)(z+h_{k+1}(z)) - (e+ \sum_{i=1}^{k} h_i)(z+h_{k+1}(z))$$

$$+ \sum_{i=1}^{k} (h_i(z + h_{k+1}(z)) - h_i(z)),$$

since by assumption $0 \leq z + h_{k+1}(z) \leq 1$.

We now use the induction hypothesis and the convexity of h_i to obtain

$$|a(z)| \leq (- \sum_{i=1}^{k} Dh_i(0))(\sum_{i=1}^{k} Dh_i(1))$$

$$+ \sum_{i=1}^{k} Dh_i(1)|h_{k+1}(z)|$$

and from (1.10)we get

$$|a(z)| \leq (- \sum_{i=1}^{k+1} Dh_i(0))(\sum_{i=1}^{k} Dh_i(1))$$

which proves the result for $n = k + 1$.

We want to build the product integral of the regular integrable intensity valued function h_u, $u \in [s,t]$ which we shall denote by

$$\prod_s^t (e + h_u du) \tag{3.2}$$

and we therefore let $T = \{t_1,\ldots,t_n\}$ denote a partition of $[s,t]$,

$$s = t_1 \leq \ldots \leq t_n = t$$

and let $|T| = \max_{1 \leq i \leq n-1} (t_{i+1} - t_i)$.

We then define

$$H(u,v) = \int_u^v h_x dx, \quad s \leq u \leq v \leq t,$$

and

$$f_T = \prod_{i=1}^{n-1} (e + H(t_i, t_{i+1})). \tag{3.3}$$

The set of partitions T form a directed set under refinement and we want to prove that $\lim f_T$ exists, where the limit is taken under refinements. The limit will then be denoted by the symbol (3.2) and will be called the product integral of h_u on the interval $[s,t]$.

We shall first prove the following

<u>3.2. Lemma.</u> If Dh.(1) is integrable on $[s,t]$ then there exists a function $\alpha(T)$ such that

$$\alpha(T) \to 0 \text{ as } |T| \to 0$$

and such that for any refinement U of T we have

$$|f_U - f_T| \leq \alpha(T).$$

Proof. Let $T = \{t_i, i = 1,\dots,n\}$ and let $U = \{u_{ij}, j = 1,\dots,n_i, i = 1,\dots,n\}$ be a refinement of T such that

$$t_i = u_{i1} \le \dots \le u_{in_i} = t_{i+1},$$

and define

$$\begin{aligned} f_i &= e + H(t_i,t_{i+1}) \\ &= e + \sum_{j=1}^{n_i-1} H(u_{ij}, u_{i(j+1)}) \end{aligned}$$

and

$$g_i = \prod_{j=1}^{n_i-1} (e + H(u_{ij},u_{i(j+1)}).$$

For $|T|$ sufficiently small we can apply Lemma 3.1 and we get

$$|f_i - g_i| \le (-\int_{t_i}^{t_{i+1}} Dh_u(0)du)\int_{t_i}^{t_{i+1}} Dh_u(1)du.$$

Using this evaluation we obtain

$$\begin{aligned} |f_U - f_T| &= |\prod_{i=1}^{n-1} f_i - \prod_{i=1}^{n-1} g_i| \\ &\le \sum_{k=1}^{n-1}\left|\prod_{i=1}^{k-1} f_i \cdot f_k \cdot \prod_{i=k+1}^{n-1} g_i - \prod_{i=1}^{k-1} f_i \cdot g_k \cdot \prod_{i=k+1}^{n-1} g_i\right| \\ &= \sum_{k=1}^{n-1}\left|\prod_{i=1}^{k-1} f_i \cdot f_k - \prod_{i=1}^{k-1} f_i \cdot g_k\right| \\ &\le \sum_{k=1}^{n-1}\left(\prod_{i=1}^{k-1} Df_i(1)\right)|f_k - g_k| \\ &= \sum_{k=1}^{n-1} \prod_{i=1}^{k-1} (1+\int_{t_i}^{t_{i+1}} Dh_u(1)du)(\int_{t_k}^{t_{k+1}} Dh_u(1)du)(-\int_{t_k}^{t_{k+1}} Dh_u(0)du) \end{aligned}$$

$$\leq \sup_{1\leq k\leq n-1} \left(-\int_{t_k}^{t_{k+1}} Dh_u(1)du\right)\left[\prod_{k=1}^{n-1}\left(1+\int_{t_i}^{t_{i+1}} Dh_u(1)du\right) - 1\right].$$

Now the last factor converges as a product integral to

$$\exp \int_s^t Dh_u(1)du - 1$$

and the first factor tends to zero since Dh.(1) is integrable on [s,t].

3.3. Theorem. Let h_u be a measurable intensity valued function such that Dh.(1) is integrable on [s,t]. Then the product integral of h_u on [s,t] exists.

Proof. Take a sequence of partitions T_n such that $|T_n| \to 0$. Consider the family

$$\{f_{T_n},\ n = 1,2,\ldots\}. \tag{3.4}$$

From the inequality

$$|f_T(z') - f_T(z'')| \leq Df_T(1)|z' - z''|$$

$$= \prod_{i=1}^{n-1}\left(1 + \int_{t_i}^{t_{i+1}} Dh_u(1)du\right)|z' - z''|$$

it follows that the family (3.4) is equicontinuous. By Arzelá-Ascolis theorem this implies the existence of a probability generating function f and a subsequence n' such that

$$\lim_{n'\to\infty} f_{T_{n'}} = f.$$

We want to prove

$$\lim_{T} f_T = f. \tag{3.5}$$

Take $\varepsilon > 0$. We want to prove that there exists a partition T_o such that for any refinement T of T_o we have

$$|f - f_T| \leq \varepsilon. \tag{3.6}$$

Now take n' so large that

$$|f_{T_{n'}} - f| \leq \varepsilon/2 \tag{3.7}$$

and such that $\alpha(T_{n'}) \leq \varepsilon/2$ and take $T_o = T_{n'}$.

From Lemma 3.2 we find that for any refinement T of T_o we have

$$|f_T - f_{T_o}| \leq \alpha(T_o) \leq \varepsilon/2$$

which together with (3.7) gives (3.6).

<u>3.4. Corollary</u>. Let Dh.(1) be integrable on [s,t], then

$$\left|\prod_s^t(e+h_u du) - e - \int_s^t h_u du\right| \leq \left(-\int_s^t Dh_u(0)du\right)\int_s^t Dh_u(1)du. \tag{3.8}$$

<u>Proof</u>. We let $T = \{t_1,\ldots,t_n\}$ denote a partition of [s,t] and define

$$h_i = \int_{t_i}^{t_{i+1}} h_u du.$$

For |T| sufficiently small we can apply Lemma 3.1 and taking the limit over T we obtain (3.8).

Let in particular $h_u \equiv h$, $u \in [s,t]$ where h is a regular intensity function then we define

$$e(h) = \prod_0^1 (e+hdu). \tag{3.9}$$

The definition of the integral then immediately gives the following results:

<u>3.5 Corollary</u>. Let $h_1,\ldots,h_n$ and $h = h_1+\ldots+h_n$ be regular intensity functions, then

$$\lim_{n\to\infty} (e+\frac{1}{n}h)^n = e(h) \tag{3.10}$$

$$\left| \prod_{i=1}^{n} e(h_i) - e - h \right| \leq -Dh(0)Dh(1) \tag{3.11}$$

and in particular

$$|e(h) - e - h| \leq -Dh(0)Dh(1). \tag{3.12}$$

We also have

$$e((t-s)h) = \prod_{s}^{t}(e+hdu) \tag{3.13}$$

and

$$\lim_{k\to\infty}\left(\prod_{i=1}^{n} e(h_i k^{-1})\right)^k = e(h). \tag{3.14}$$

3.6. Theorem. Let h_t be a regular measurable intensity function such that $Dh.(1)$ is locally integrable and let

$$f_{s,t} = \prod_{s}^{t}(e+h_u du).$$

Then

$$f_{s,t} = f_{s,u} \cdot f_{u,t}, \quad 0 \leq s \leq u \leq t \tag{3.15}$$

and $Df_{s,t}(1)$ is finite and continuous.

Further $f_{\cdot,t}(z)$ is absolutely continuous and satisfies the backward Kemogorov equation

$$\partial_s f_{s,t} = - h_s \cdot f_{s,t}, \quad s \notin N. \tag{3.16}$$

From this follows that

$$Df_{s,t}(1) = \exp \int_{s}^{t} Dh_u(1)du. \tag{3.17}$$

Further $f_{s,\cdot}$ is absolutely continuous and satisfies the forward Kolmogorov equation.

$$\partial_t f_{s,t}(z) = Df_{s,t}(z)h_t(z), \quad t \notin N. \tag{3.18}$$

<u>Proof</u>. The relation (3.15) follows immidiately from the definition of the product integral. We find from (3.3) that

$$1 \leq (1-z)^{-1}(1-f_T(z)) \leq Df_T(1) = \prod_{k-1}^{n-1} (1 + \int_{t_k}^{t_{k+1}} Dh_u(1)du).$$

Taking limits over T and letting $z \uparrow 1$ we get that

$$1 \leq Df_{s,t}(1) \leq \exp \int_s^t Dh_u(1)du,$$

which proves that $Df_{s,t}(1)$ is finite and continuous.

In order to prove (3.16) we consider

$$\left| (e + \int_{s-u}^{s+v} h_x dx) \cdot f_{s+v,t} - f_{s-u,t} \right|$$

$$= \left| (e + \int_{s-u}^{s+v} h_x dx - f_{s-u,s+v}) \cdot f_{s+v,t} \right|$$

$$\leq \left(- \int_{s-u}^{s+v} Dh_x(0)dx \right) \int_{s-u}^{s+v} Dh_x(1)dx.$$

Now choose N such that for $s \notin N$ we have

$$\lim_{u\downarrow 0, v\downarrow 0} (u+v)^{-1} \int_{s-u}^{s+v} h_x dx = h_s \tag{3.19}$$

and

$$\lim_{u\downarrow 0, v\downarrow 0} (u+v)^{-1} \int_{s-u}^{s+v} Dh_x(0)dx = Dh_s(0) \tag{3.20}$$

and

$$\lim_{u\downarrow 0, v\downarrow 0} (u+v)^{-1} \int_{s-u}^{s+v} Dh_x(1)dx = Dh_s(1). \tag{3.21}$$

Then we get that

$$\lim_{u\downarrow 0, v\downarrow 0} (u+v)^{-1} \left[(e + \int_{s-u}^{s+v} h_x dx) \cdot f_{s+v,t} - f_{s-u,t} \right] = 0.$$

Since

$$\lim_{u\downarrow 0, v\downarrow 0} (u+v)^{-1} \int_{s-u}^{s+v} h_x \cdot f_{s+v,t}\,dx = h_s \cdot f_{s,t}$$

we get that

$$\partial_s f_{s,t} = \lim_{u\downarrow 0, v\downarrow 0} (u+v)^{-1}(f_{s+v,t} - f_{s-u,t})$$

exists, and that

$$\partial_s f_{s,t} = - h_s \cdot f_{s,t}, \quad s \notin N,$$

which proves (3.16).

Now consider

$$\left| f_{s,t+v} - f_{s,t-u} \cdot (e + \int_{t-u}^{t+v} h_x dx) \right|$$

$$= \left| f_{s,t-u} \cdot f_{t-u,t+v} - f_{s,t-u} \cdot (e + \int_{t-u}^{t+v} h_x dx) \right|$$

$$\leq Df_{s,t-u}(1) \left| f_{t-u,t+v} - e - \int_{t-u}^{t+v} h_x dx \right|$$

$$\leq Df_{s,t}(1)(- \int_{t-u}^{t+v} Dh_x(0)dx)(\int_{t-u}^{t+v} Dh_x(1)dx)$$

For $t \notin N$ we get that

$$\lim_{u\downarrow 0, v\downarrow 0} (u+v)^{-1}(f_{s,t+v} - f_{s,t-u} \cdot (e + \int_{t-u}^{t+v} h_x dx)) = 0.$$

This is rewritten as follows:

$$(u+v)^{-1}(f_{s,t+v}(z) - f_{s,t-u}(z))$$

$$+ a(u,v)^{-1}(f_{s,t-u}(z) - f_{s,t-u}(z+a(u,v)))a(u,v)(u+v)^{-1},$$

where

$$a(u,v) = \int_{t-u}^{t+v} h_x(z)dx.$$

For each $z < 1$ the differentiation of $f_{s,t}(\cdot)$ at z is uniform in (s,t) and hence

$$\lim_{u\downarrow 0, v\downarrow 0} a(u,v)^{-1}(f_{s,t-u}(z) - f_{s,t-u}(z+a(u,v)) = Df_{s,t}(z).$$

From (3.19) we get that

$$\lim_{u\downarrow 0, v\downarrow 0} a(u,v)(u+v)^{-1} = h_t(z)$$

and hence $f_{s,\cdot}(z)$ is differentiable at t for $z < 1$ and

$$\partial_t f_{s,t}(z) = Df_{s,t}(z)h_t(z), \quad t \notin N.$$

This relation clearly holds for $z = 1$ as well.

Notice that the relation (3.17) implies that the convergence in (3.5) is strong, since

$$Df_T(1) = \prod_{k=1}^{n-1} (1 + \int_{t_i}^{t_{i+1}} Dh_u(1)du)$$

converges to

$$\exp \int_s^t Dh_u(1)du = Df(1).$$

3.7. Theorem. Let h_t be a regular measurable intensity function such that $Dh.(1)$ is locally integrable.

Then the equation

$$z - f_{s,t}(z) = - \int_s^t h_u(f_{u,t}(z))du \tag{3.22}$$

with initial conditions

$$f_{s,s} = e \tag{3.23}$$

has a unique solution given by the product integral (3.2)

Proof. This follows from Carathéodory [1] p.674 since the function $h_s(\cdot)$ satisfies the Lipschitz condition

$$|h_s(z) - h_s(z')| \leq Dh_s(1)|z'-z|$$

where $Dh_{\cdot}(1)$ is locally integrable.

3.8. Theorem. Let h_t and $h_t^{(n)}$, $n = 1,2,\ldots$ denote regular measurable intensity valued functions such that

$$Dh_t^{(n)}(1) \leq c, \quad n = 1,2,\ldots \ .$$

If for all real functions g, such that $\int_0^1 |g(t)|dt < \infty$ we have

$$\lim_{n\to\infty} \int_0^1 h_t^{(n)} g(t)dt = \int_0^1 h_t g(t)dt$$

then

$$\lim_{n\to\infty} \prod_0^1 (e+h_t^{(n)}dt) = \prod_0^1 (e+h_t dt) \tag{3.24}$$

Proof. Let us use the notation

$$f_{s,1}^{(n)} = \prod_s^1 (e+h_t^{(n)}dt).$$

The product integral satisfies equation (3.16) and hence for $s' < s''$

$$\left| f_{s',1}^{(n)}(z) - f_{s'',1}^{(n)}(z) \right| \leq \int_{s'}^{s''} Dh_t^{(n)}(1)dt \leq (s''-s')c.$$

Similarly

$$\left| f_{s,1}^{(n)}(z') - f_{s,1}^{(n)}(z'') \right| \leq Df_{s,1}^{(n)}(1)|z'-z''|$$

$$= |z'-z''| \exp \int_s^1 Dh_t^{(n)}(1)dt \leq |z'-z''| \exp c.$$

Thus the family of functions

$$\{f_{\cdot,1}^{(n)}(\cdot), \ n = 1,2,\ldots\} \tag{3.25}$$

is equicontinuous, and by Arzelà-Ascoli's theorem there exists a probability generating function $f_{s,1}(\cdot)$ and a subsequence n' such that

$$\lim_{n'\to\infty} f_{s,1}^{(n')}(z) = f_{s,1}(z)$$

uniformly in $s \in [0,1]$ and $z \in [0,1]$.

From (3.16) we get

$$f_{s,1}^{(n')}(z) - z = \int_s^1 h_t^{(n')}(f_{t,1}^{(n')}(z))dt,$$

for $n' \to \infty$ get that $f_{\cdot,1}(z)$ satisfies the equation (3.22). By Theorem 3.7 the solution is unique and given by the product integral of h_t on $[0,1]$. Thus any limit point of the set of functions (3.25) must be the product integral and hence the family converges, as was to be proven.

4. The imbedding problem for Markov branching processes.

In order to formulate the results about the imbedding problem we shall define the notions of infinite factorizability and triangular null array.

4.1. Definition. Let f be a regular function in C. Then f is called infinitely factorizable if for all n there exist $f_1,\dots,f_n$ such that

$$f = f_1 \cdot \ldots \cdot f_n$$

$$Df_i(1) - 1 \le \varepsilon, \; i = 1,\dots,n.$$

If for all n there exists f_n such that

$$f = (f_n)^n$$

then f is called infinitely divisible.

4.2. Definition. A triangular null array is a family $\{f_{n,k},\; k = 1,\dots,N_n,\; n = 1,2,\dots\}$ of functions in C, such that

$$\lim_n \sup_k (Df_{n,k}(1)-1) = 0.$$

The marginal product is

$$f_n = f_{n,1} \cdot \ldots \cdot f_{n,N_n} = \prod_{k=1}^{n} f_{n,k}$$

and the limit of the array is the strong limit of f_n.

<u>4.3 Definition</u>. A Poisson generating function is of the form $e(h)$ where h is an extremal intensity function, see (1.15).

The next theorem sums up the results about the imbedding problem and presents various characterizations of the imbeddable functions in C.

<u>4.4. Theorem</u>. Let f be a regular function in C, then the following statements are equivalent:

f is imbeddable (4.1)

f is infinitely factorizable (4.2)

f is the limit of a triangular null array (4.3)

f is the strong limit of a finite composition of generating functions of the form $e(h)$ (4.4)

f is the strong limit of a finite composition of Poisson generating functions (4.5)

f has the representation

$$f = \prod_0^1 (e+h_u du) \tag{4.6}$$

for some regular measurable intensity valued functions h_t for which $Dh.(1)$ is integrable on [0,1].

There exists absolutely continuous functions $f.$ and $f.^*$ and a regular measurable intensity valued function h_t for which $Dh.(1)$ is integrable on [0,1] such that (4.7)

$$\frac{d}{dt} f_t = Df_t \circ h_t, \quad t \notin N$$

$$f_o = e, \ f_1 = f$$

and

$$\frac{d}{dt} f_t = -h_t \cdot f_t, \ t \notin N$$

$$f_0 = f, \ f_1 = e,$$

where N is a null set for Lebesgue measure.

Proof. (4.1) → (4.2). If f is imbeddable in the family $\{f_{s,t}, \ 0 \leq s \leq t \leq 1\}$ then since $Df_{s,t}(1)$ is continuous and multiplicative we can choose a partition $0 = t_o < t_1 < \ldots < t_n = 1$ such that

$$Df_{t_i, t_{i+1}}(1) - 1 \leq \varepsilon, \ i = 0, \ldots, n-1$$

and define

$$f_i = f_{t_i, t_{i+1}}, \ i = 0, \ldots, n-1.$$

This proves that f is infinitely factorizable.

(4.2) → (4.3). If f is infinitely factorizable then for $\varepsilon_n \downarrow 0$ we can choose

$$f_{n,1}, \ldots, f_{n,N_n}$$

such that

$$f = f_{n,1} \cdot \ldots \cdot f_{n,N_n}$$

and

$$Df_{n,k}(1) - 1 \leq \varepsilon_n, \ k = 1, \ldots, N_n.$$

This proves that f is the limit of the triangular array $\{f_{n,k}\}$.

(4.3) → (4.4). Now let f be the limit of the triangular array $\{f_{n,k}\}$. We want to replace it by the array $\{e(f_{n,k}-e)\}$. Now the usual trick gives for $g_{n,k} = f_{n,k} - e$

$$\left| \prod_{k=1}^{N_n} f_{n,k} - \prod_{k=1}^{N_n} g_{n,k} \right|$$

$$\leq \sum_{k=1}^{N_n} \left| \prod_{i=1}^{k-1} f_{n,i} \cdot f_{n,k} - \prod_{i=1}^{k-1} f_{n,i} \cdot g_{n,k} \right|$$

$$\leq \sum_{k=1}^{N_n} \prod_{i=1}^{k-1} Df_{n,i}(1) \left| f_{n,k} - e(f_{n,k}-e) \right|$$

$$\leq \sum_{k=1}^{N_n} \prod_{i=1}^{k-1} Df_{n,i}(1)(Df_{n,k}(1)-1)(1-Df_{n,k}(0))$$

$$\leq \sup_k (1-Df_{n,k}(0))(Df_n(1)-1)$$

The assumptions imply that

$$\lim_n Df_n(1) = Df(1) < \infty$$

and that

$$\lim_n \sup_k (1-Df_{n,k}(0)) = 0$$

and hence that

$$g_n = \prod_{k=1}^{N_n} e(f_{n,k}-e)$$

converges uniformly to f. We want to prove strong convergence, i.e. that $\lim_n Dg_n(1) = Df(1)$ and we therefore evaluate

$$Dg_n(1) = \prod_{k=1}^{N_n} Dg_{n,k}(1) = \exp\left(\sum_{k=1}^{N_n} (Df_{n,k}(1)-1) \right).$$

Now

$$0 \leq \ln Dg_n(1) - \ln Df_n(1) = \sum_{k=1}^{N_n} - \ln Df_{n,k}(1) + Df_{n,k}(1) - 1$$

$$\leq \sum_{k=1}^{N_n} (Df_{n,k}(1))^{-1}(Df_{n,k}(1)-1)^2$$

$$\leq \sup_k (Df_{n,k}(1)-1) \sum_{k=1}^{N_n} (Df_{n,k}(1)-1).$$

The first factor converges to zero and the last factor is bounded since for

$$\sup_k Df_{n,k}(1) \leq 2$$

we get that

$$\sum_{k=1}^{N_n} (Df_{n,k}(1)-1) \leq (\ln 2)^{-1} \ln \prod_{k=1}^{N_n} Df_{n,k}(1) \qquad (4.8)$$

which is convergent.

Hence g_n converges strongly to f and (4.4) is proved.

(4.4) → (4.5). Let h be an intensity function

$$h(z) = \sum_{k=2}^{\infty} a_k (z^k - z)$$

$$= \sum_{k=2}^{\infty} a_k h_k(z),$$

see (1.15).

Now define $h_n(z) = \sum_{k=2}^{n} a_k h_k(z)$ then it follows from the results in Corollary 3.5 that $e(h_n)$ converge strongly to e(h).

From (3.14) it follows that

$$\left(\prod_{k=2}^{n} e(m^{-1} a_k h_k) \right)^m$$

converge strongly to $e(h_n)$ as m tends to infinity. This proves (4.5).

(4.5) → (4.3). This follows easily using the infinite divisibility of the Poisson generating functions.

(4.3) → (4.6). Let f be the limit of a triangular null array. Then it can be approximated by the array $\{g_{n,k} = e(f_{n,k} - e)\}$. If we define

$$g_n = \prod_{k=1}^{N_n} g_{n,k},$$

and

$$p_{n,k} = (Df_{n,k}(1)-1)\Big(\sum_{k=1}^{N_n} (Df_{n,k}(1)-1)\Big)^{-1}$$

and

$$t_{n,k} = \sum_{j=1}^{k} p_{n,j}$$

then

$$0 = t_{n,0} \le t_{n,1} \le \dots \le t_{n,N_n} = 1,$$

and

$$e(f_{n,k}-e) = \prod_{0}^{1} (e + (f_{n,k}-e)dt).$$

Therefore

$$g_n = \prod_{0}^{1} (e + h_t^{(n)} dt)$$

where

$$h_t^{(n)} = (f_{n,k}-e)p_{n,k}^{-1}, \quad t \in [t_{m,k-1}, t_{m,k}[.$$

We know that g_n converges strongly to f and we have to extract a convergent subsequence of the family

$$\{h_{\cdot}^{(n)}, \ n = 1,2,\dots\}$$

and then use Theorem 3.8 to prove the representation of f.

Let us first evaluate

$$Dh_t^{(n)}(1) = \sum_{k=1}^{N_n} (Df_{n,k}(1)-1) \tag{4.9}$$

which by (4.8) is bounded by some constant c uniformly in $t \in [0,1]$, and n, since f_n converges strongly to f.

In order to apply Theorem 1.4 we define the probability generating functions

$$f_t^{(n)}(z) = c^{-1}(h_t^{(n)}(z) + cz)$$

$$= \sum_{k=1}^{\infty} p(k,n,t)z^k.$$

Notice that by (4.9) we have

$$Df_t^{(n)}(1) = c^{-1}\left(\sum_{k=1}^{N_n}(Df_{n,k}(1)-1) + c\right) \le 2. \tag{4.10}$$

In order to find a convergent subsequence we consider first the functions

$$\{p(1,n,.),\ n = 1,2,\ldots\}.$$

These functions belong to the unit ball K of $L_\infty[0,1]$ which is compact in the $L_1[0,1]$ topology. Thus there exists a function $p(1,.) \in K$, and a subsequence $\{n_{1i}\}$ such that for all $g \in L_1[0,1]$ we have

$$\lim \int_0^1 g(t)p(1,n,t)dt = \int_0^1 g(t)p(1,t)dt$$

where the limit is taken over the subsequence $\{n_{1i}\}$.

The functions

$$\{p(2,n_{1i},.),\ i = 1,2,\ldots\}$$

are also in K and hence there exists a function $p(2,.) \in K$ and a subsequence $\{n_{2i}\}$ such that for $g \in L_1[0,1]$ we have

$$\lim \int_0^1 g(t)\ p(2,n,t)dt = \int_0^1 g(t)p(2,t)dt$$

where the limit is taken over the subsequence $\{n_{2i}\}$. Continuing this way we find a family of functions $\{p(k,.),\ k = 1,2,\ldots\} \subset K$ such that

$$\lim \int_0^1 g(t)p(k,n,t)dt = \int_0^1 g(t)p(k,t)dt$$

where the limit is taken over the diagonal sequence $\{n_{ii}\}$.

There clearly exists a null set N such that for $t \notin N$, we have $p(k,t) \geq 0$, $k = 1,2,\ldots$.

From the inequality (4.10) we get

$$\sum_{k=m}^{\infty} p(k,n,t) \leq m^{-1} \sum_{k=1}^{\infty} kp(k,n,t) \leq 2m^{-1}$$

which implies that for $t \notin N$ we get

$$\sum_{k=1}^{\infty} p(k,t) = 1.$$

Now define the probability generating function

$$f_t(z) = \sum_{k=1}^{\infty} p(k,t)z^k, \quad t \notin N.$$

We want to prove that for $g \in L_1[0,1]$ we have

$$\lim \int_0^1 g(t)f_t^{(n)}dt = \int_0^1 g(t)f_t dt$$

where the limit is taken over the subsequence $\{n_{ii}\}$.

If $g(t) \geqq 0$ and $\int_0^1 g(t)dt = 1$ then

$$\lim \int_0^1 g(t)f_t^{(n)}(z)dt = \int_0^1 g(t)f_t(z)dt$$

for each fixed $z \in [0,1]$. But the functions are probability generating functions and hence by Theorem 1.4 the convergence is uniform in $z \in [0,1]$.

This result can easily be extended to all $g \in L_1[0,1]$ and if we define

$$h_t(z) = c(f_t(z)-z)$$

then we get that for all $g \in L_1[0,1]$

$$\lim \int_0^1 g(t)h_t^{(n)}dt = \int_0^1 g(t)h_t dt$$

where the limit is along the sequence $\{n_{ii}\}$ and hence the assumptions of Theorem 3.8 are satisfied and we conclude that

$$f = \lim f_n = \lim \prod_0^1 (e + h_t^{(n)}dt) = \prod_0^1 (e + h_t dt),$$

where the limit is along the sequence $\{n_{ii}\}$. This proves (4.6).

(4.6) $\Rightarrow$ (4.7). This follows from the properties of the product integral given in Theorem 3.6. We define

$$f_t = \prod_0^t (e + h_u du)$$

and

$$f_t^* = \prod_t^1 (e + h_u du).$$

(4.7) → (4.1). If we denote by $f_{s,t}$ the integral

$$f_{s,t} = \prod_s^t (e + h_u du)$$

then this family imbeds f. Notice that by (3.17) we have that $Df_{s,t}(1)$ is finite and continuous.

This completes the proof of Theorem 4.4.

The previous results are all for non-homogeneous chains but we can clearly prove the following theorem using the same techniques:

4.5. Theorem. Let f be a regular function in C, then the following statements are equivalent:

f is imbeddable in a homogeneous chain. (4.11)

f is infinitely divisible. (4.12)

f is the limit of a triangular null array with identical components in each row. (4.13)

f is of the form e(h) for some regular intensity function h. (4.14)

f is the strong limit of a finite composition of Poisson generating functions. (4.15)

There exists a regular intensity function h and absolutely continuous functions $f_.$ and $f_.^*$ such that (4.16)

$$\frac{d}{dt} f_t = Df_t \quad h$$

$$f_0 = e, \; f_1 = f$$

and

$$\frac{d}{dt} f_t^* = - h(f_t)$$

$$f_0 = f, \; f_1 = e.$$

Proof. Omitted.

Karlin and McGregor [8] studied the imbedding problem for homogeneous chains and gave various criteria for non-imbeddability using complex function theory. They proved that at large class of distributions including the Poisson and Binomial distribution could not be imbedded.

One may still ask whether these distributions can be imbedded in non-homogeneous chains.

We shall now give a very simple necessary condition for imbeddability which can be used to exclude certain distributions.

4.6. Theorem. If $f(z) = \sum_{k=1}^{\infty} p_k z^k$ is imbeddable then

$$Df(0)Df(1) \geqq 1 \tag{4.17}$$

or

$$p_1 \sum_{k=1}^{\infty} kp_k \geqq 1 \tag{4.18}$$

Proof. It is easily seen that the set of functions in C satisfying (4.17) is a semigroup and that it is closed under strong convergens. By (4.5), (Theorem 4.4) it is therefore enough to prove (4.17) for the Poisson generating functions.

Let therefore

$$h(z) = \lambda(z^k - z),$$

for some $k = 2,\ldots$ and $\lambda > 0$ then

$$Df(0) = \exp Dh(0) = \exp(-\lambda)$$

$$Df(1) = \exp Dh(1) = \exp \lambda(k-1)$$

which proves (4.17). If $h \equiv 0$ then $f(z) = z$ which also satisfies (4.17).

<u>4.7. Corollary</u>. The truncated binomial distribution given by

$$f_1(z) = (((1-\pi)z+\pi)^n - \pi^n)(1-\pi^n)^{-1}$$

$0 < \pi < 1$, $0 \leq z \leq 1$ is not imbeddable. The truncated Poisson distribution given by

$$f_2(z) = (\exp(-\rho(1-z)) - \exp(-\rho))(1-\exp(-\rho))$$

$0 < \rho < \infty$, $0 \leq z \leq 1$ is not imbeddable.

<u>Proof</u>. We find

$$Df_1(0)Df_1(1) = \left[\frac{n(1-\pi)}{1-\pi^n}\right]^2 \pi^{n-1}.$$

Now

$$\frac{1}{n}\,\frac{1-\pi^n}{1-\pi} = \frac{1}{n}\sum_{k=1}^{n-1}\pi^k > \left(\prod_{k=1}^{n-1}\pi^k\right)^{\frac{1}{n}} = \pi^{\frac{1}{2}(n-1)}$$

which proves that

$$Df_1(0)\; Df_1(1) < 1$$

and hence f_1 is not imbeddable. The corresponding result for f_2 follows by setting $\pi = 1 - \rho n^{-1}$ and letting $n \to \infty$.

The results in Corollary 4.7 are about the truncated distributions because we have only investigated distributions on $N = \{1,2,...\}$.

The results can, however, be interpreted in the context of supercritical branching processes on $S = \{0,1,2,...\}$ as follows.

Let $f(z) = \sum_{k=0}^{\infty} p_k z^k$ and assume $Df(1) = \sum_{k=0}^{\infty} kp_k > 1$ then f is called supercritical. For any supercritical f there exists a unique point $\alpha \in [0,1[$ such that

$$f(\alpha) = \alpha.$$

We shall call α the fixed point of f.

We shall then call such an f with fixed point α imbeddable if there exists a family $\{f_{s,t}, 0 \leq s \leq t \leq t_0\}$ of supercritical functions such that (1.6), (1.7), (1.8), and (1.9) holds together with the condition (4.19)

$$f_{s,t}(\alpha) = \alpha, \ 0 \leq s \leq t. \tag{4.19}$$

Now define

$$\psi(z) = (1-\alpha)z + \alpha$$

and

$$\tilde{f} = \psi^{-1} \circ f \circ \psi.$$

It is easily seen that $\tilde{f}$ is in C and that if f is imbeddable then $\tilde{f}$ is imbeddable in the sense of definition 1.3 Hence by Theorem 4.6 we get

$$D\tilde{f}(0)D\tilde{f}(1) \geqq 1$$

but this is equivalent to

$$Df(\alpha)Df(1) \geqq 1.$$

Thus we have proved.

<u>4.8. Theorem.</u> In order that a supercritical probability generating function f with fixed point α be imbeddable in a family of supercritical functions with fixed point α it has to satisfy the condition

$$Df(\alpha)Df(1) \geqq 1.$$

In this sence Corollary 4.7 implies that the Poisson distribution with $\lambda > 1$ and the binomial distribution with $np > 1$ are not imbeddable.

5. Applications to processes with independent increments.

In this section we shall sketch some applications of the results in section 4 to processes with independent increments.

Let us first note that the set of stochastic matrices on $N' = N \cup \{\infty\}$ is a convex compact set in the topology of entrywise convergence. The extreme points E are the stochastic matrices with entries equal to 0 or 1 and since each row must contain a 1 this is the same as the set of functions mapping N' into N', i.e. $(N')^{N'}$.

A simple application of Choquet's theorem on integral representation of points in convex compact sets then gives the following theorem.

5.1. Theorem. Any stochastic matrix on N' has a representation

$$P = \int_E E\mu(dt)$$

where μ is a probability measure on E and the integration is to be understood as entrywise integration.

It is easily seen that if P is a stochastic matrix on N then μ is concentrated on the set of extreme stochastic matrices on N or on the set of mappings from N into N.

Now observe that E is a semigroup under matrixmultiplication or equivalently that N^N is a semigroup under composition.

It is easily seen that for

$$P(\mu) = \int_E E\mu(dE)$$

we get

$$P(\mu * \nu) = P(\mu)P(\nu)$$

where * denote convolution of the measures on E.

We are therefore led to consider the set of probability measures on N^N, i.e. stochastic processes with discrete time N and state space N.

<u>5.2. Theorem</u>. Let μ denote a random walk on N then $P(\mu) \in B$, i.e. $P(\mu)$ is the transition probability matrix of a branching process.

<u>Proof</u>. Let $X_1,\ldots,X_n,\ldots$ be independent random variable with values in N and with common distribution

$$P\{X_1 = k\} = p_k.$$

Let μ be the distribution of the random walk

$$S_n = \sum_{k=1}^{n} X_i, \quad n = 1,2,\ldots .$$

Then μ is a probability measure on N^N or on E and we therefore compute

$$P(\mu)_{ij} = \int_E E_{ij}\mu(dE)$$

$$= \mu\{E|E_{ij} = 1\}$$

$$= \mu\{S_i = j\} = p^{(i)}(j)$$

Thus the i'th row of $P(\mu)$ is the i-fold convolution of the first row which means that $P(\mu) \in B$.

Notice that this relation between random walks on N and branching processes is one to one and just reflects the fact that given a probability measure p on N one can either construct a random walk or a branching process from p.

Thus there is a relation between the imbedding problem for branching processes and the imbedding problem for processes with increments that are independent and take values in the set of random walks on N. We shall not rewrite theorems 4.4, 4.5, and 4.6 in this new formulation.

To illustrate the connection further we notice that the semigroup operation in N^N is that of composition or subordination, but that is exactly the one used in branching processes. Consider the following array

$$\{X_{ij}, \ j \in N, \ i \in N\}.$$

of independent random variables with values in N. For each i the variables $\{X_{ij}, \ j \in N\}$ have the same distribution p_i.

Now construct the measures μ_i, $i \in N$ as the random walk on N corresponding to p_i, that is induced by the mapping

$$S_{in} = \sum_{j=1}^{n} X_{ij}, \quad n \in N.$$

Let S_i, $i \in N$ denote the sample path

$$\{S_{in}, \ n \in N\},$$

and consider the stochastic process

$$T_k = S_k \cdot S_{k-1} \cdot \ldots \cdot S_1$$

where the composition is that of subordination. Thus

$$T_k(n) = S_k(S_{k-1}(\ldots(S_1(n))\ldots)$$

is the random variable that determines the size of the k'th generation when the zero'th generation had n individuals.

<u>5.3 . Theorem</u>. The process

$$\{T_k(n), \ n \in N\}$$

is for fixed k a random walk on N, and the process

$$\{T_k(n), \ k \in N\}$$

is for fixed n a Galton-Watson process starting with n individuals.

Proof. Obvious from the above remarks and the construction.

Let us finally conclude that we have had four different representations of the semigroup we are working in.

(5.1) The set of probability measures on N with a convolution * given by

$$(p_2*p_1)(n) = \sum_{k=1}^{\infty} p_2(k)p_1^{(k)}(n)$$

corresponding to taking a sum of a random number of random variables.

(5.2) The set of probability generating functions with composition as the semigroup operation.

(5.3) The set of stochastic matrices satisfying condition (1.1) with matrix multiplication.

(5.4) The set of random walk measures on the semigroup N^N with convolution.

Notice that (5.1) and (5.2) are connected with a linear mapping and so is (5.3) and (5.4). In (5.3) and (5.4) we have a non-convex semigroup with a bilinear semigroup operation and in (5.1) and (5.2) we have a convex semigroup with a semigroup operation which is not bilinear. In (5.1) and (5.2) we obtain the infinitessimal generators by just subtracting the identity and this is not the case in (5.3) and (5.4).

6. Acknowledgements

The author expresses his sincere thanks to Gerald Goodman for many discussions on the topic. A lot of the material in section 4 was presented in a lecture at Imperial College in 1972 by Gerald Goodman.

7. References

1. Carathéodory C.: Vorlesungen über reelle funktionen. Leipzig und Berlin: Teubner 1918.

2. Dobrušin, R.L.: Generalization of Kolmogorov's equations for Markov processes with a finite number of possible states. Mat.Sb.N.S.33,(75), 567-596 (1953).

3. Goodman, G.S.: An Intrinsic time forNon-Stationary Finite Markov Chains. Z. Wahrscheinlichkeitstheorie verw. Gebiete 16, 165-180(1970).

4. Goodman, G.S., and Johansen, S.: Kolmogorov's Differential Equations for Non-Stationary, Countable State Markov Processes with Uniformly Continuous Transition Probabilities. Proc. Camb.Phil.Soc. 73, 119-138(1973).

5. Harris, T.E.: The theory of branching processes. Berlin: Springer (1963).

6. Johansen, S.: A central Limit Theorem for Finite Semigroups and Its Application to the Imbedding Problem for Finite Markov Chains. Z. Wahrscheinlichkeitstheorie verw. Gebiete 26, 171-190(1973).

7. Johansen, S.: The Bang-Bang problem for Stochastic Matrices Z.Wahrscheinlichkeitstheorie verw. Gebiete 26, 191-195(1973).

8. Karlin, S. and McGregor J.: Embeddability of discrete Time Simple Branching Processes into Continuous Time Branching Processes. Tran.Amer.Math. Soc. 132, 115-136(1968).

9. Loewner, K.: Untersuchungen über Schlichte Konforme Abbildungen des Einheitskreises. I.Math.Ann.89, 103-121(1923).

10. Neuberger, J.W. Continuous products and nonlinear integral equations. Pacific J. Math. 8, 529-549,(1958).

11. Pommerenke, C.: Über die Subordination analytischer Funktionen. J.Rein Aug.Math. 218, 159-173 (1965).

12. Goodman, G.S. Univalent Functions and optimal control. Thesis.Stanford 1968, 107 pp.