On a limit theorem for non-stationary branching processes.

TETSUYA HATTORI and HIROSHI WATANABE

0. Introduction.

The purpose of this paper is to give a limit theorem for a certain class of discrete-time multi-type non-stationary branching processes (multi-type varying environment Galton-Watson processes). Let $X_N = {}^t(X_{N,1}, X_{N,2}, \cdots, X_{N,d})$, $N = 0, 1, 2, \cdots$ be a discrete-time branching process with d types. The discrete-time branching process X_N is determined by its generating functions $\phi_{n,N,i}(z)$. For $i \in \{1, 2, \cdots, d\}$, define $e_i = {}^t(e_{i,1}, e_{i,2}, \cdots, e_{i,d}) \in \mathbf{Z}_+{}^d$ by $e_{i,i} = 1$, and $e_{i,j} = 0$, if $i \neq j$. Then

$$
\begin{aligned}
(1) \qquad & \phi_{n,N,i}(z) \stackrel{d}{=} \sum_{\alpha \in \mathbf{Z}_+{}^d} \left(\prod_{i=1}^{d} z_i^{\alpha_i} \right) \mathbf{Prob}\{X_N = \alpha \mid X_n = e_i\} , \\
& n \in \mathbf{Z}_+ ,\ N \in \mathbf{Z}_+ ,\ n \leq N ,\ i = 1, 2, \cdots, d ,\ z \in \mathbf{C}^d .
\end{aligned}
$$

The generating functions $\phi_{n,N} = {}^t(\phi_{n,N,1}, \phi_{n,N,2}, \cdots, \phi_{n,N,d})$ are determined by recursion relations (see, for example, [2] on branching processes). For a non-stationary branching process, we have

$$
\begin{aligned}
(2) \qquad & \phi_{n,N}(z) = \phi_{n,N-1}(\phi_{N-1,N}(z)) = \phi_{n,n+1}(\phi_{n+1,N}(z)) ,\ n < N , \\
& \phi_{n,n}(z) = z ,
\end{aligned}
$$

Compared with the stationary case, $\phi_{n,N} = \phi_{0,N-n}$, or the one-type non-stationary case, $d = 1$, not many studies seem to have been done for multi-type non-stationary branching processes. Perhaps one of the difficulties with such studies lies in finding a class of processes where the non-stationarity and multi-typedness enter in a non-trivial way, and yet a clear limiting behavior is obtained.

As an approach to this problem, we consider in this paper the case where the non-stationarity enters in a simple way. Namely, we assume

$$
(3) \qquad \phi_{N-1,N}(z) = D_{N-1}^{-1} F(D_N z) ,\ N = 1, 2, 3, \cdots ,
$$

where $D_N = \mathrm{diag}(D_{N,1}, D_{N,2}, \cdots, D_{N,d})$, $N \in \mathbf{Z}_+$, are diagonal matrices, and F is a $\mathbf{C}^d$-valued function in d variables.

From eq. (2) and eq. (3) it follows that

$$
\phi_{n,N}(z) = D_n^{-1} F^{N-n}(D_N z) ,
$$

which is similar to the corresponding formula for the stationary case. In fact, if the matrices D_N are unit matrices, this equation reduces to the formula for the stationary case.

If the limit $\hat{F}(z) = \lim_{N\to\infty} D_{N-1}^{-1} F(D_N z)$ exists and is analytic at $z=1$, then the asymptotic behavior of the branching process will be essentially that of a stationary branching process defined by $\phi_{N-1,N} = \hat{F}$. However, if $\hat{F}$ is singular at $z=1$, we cannot apply the theory of stationary branching processes directly, hence a consideration of non-stationarity (N-dependences of $\phi_{N-1,N}(z)$) becomes non-trivial.

We introduce some notation. Put $\vec{1} \stackrel{d}{=} {}^t(1,\cdots,1)$. For a $\mathbf{C}^d$-valued differentiable function $F = {}^t(F_1, F_2, \cdots, F_d)$ in d variables $z=(z_1,\cdots,z_d)$, we define a $d\times d$ matrix $\nabla F(z)$ by $\nabla F(z)_{ij} = \dfrac{\partial F_i(z)}{\partial z_j}$. For a $d\times d$ matrix M, $\| M \| \stackrel{d}{=} \sup_{v=(v_1,\cdots,v_d),\ \|v\|=1} \| M\, v \|$ is the usual operator norm, with $\| v \| = \sqrt{\sum_{i=1}^{d} |v_i|^2}$ for a d-component vector v. We also use the notation $\exp(t) \stackrel{d}{=} (\exp(t_1),\cdots,\exp(t_d))$ for $t=(t_1,\cdots,t_d)\in\mathbf{C}^d$.

Definition 1. Let $\{D_N\}$, $N=0,1,2,\cdots$, be a series of d-dimensional diagonal matrices, $D_N = \mathrm{diag}(D_{N,1}, D_{N,2}, \cdots, D_{N,d})$, with positive diagonal elements; $D_{N,i} > 0$. For a constant $\delta > 1$, we say that $\{D_N\}$ is of order δ if $\sup_{N\in\mathbf{Z}_+, i=1,\cdots,d} D_{N,i} < \infty$ and $\sup_{N\in\mathbf{Z}_+, i=1,\cdots,d} \delta^{-N} D_{N,i}^{-1} < \infty$ are satisfied.

Definition 2. Let $\{D_N\}$, $N=0,1,2,\cdots$, be a series of d-dimensional diagonal matrices of order δ (>1) and let ℓ be a constant satisfying $\ell > \delta$. We say that a $\mathbf{C}^d$-valued function F in d variables defined on an open set containing $\bigcup_{N\in\mathbf{Z}_+} \{D_N\,\vec{1}\}$ is $(\{D_N\},\ell)$-regular if the following are satisfied.

1. The function F has the expression
$$F(D_N\vec{1}+t) = D_{N-1}\vec{1} + A_N\, t + \tilde{F}_N(t)\,, \quad N\in\mathbf{Z}_+\,,\ \| t \| < C_0\,\delta^{-N},$$
where A_N is a $d\times d$ matrix independent of t, and $\tilde{F}_N$ is analytic in $\| t \| < C_0\,\delta^{-N}$ and satisfies
$$\| \tilde{F}_N(t) \| \le C_1\,\delta^N \| t \|^2, \quad \| t \| < C_0\,\delta^{-N},\ N\in\mathbf{Z}_+\,,$$
for some positive constants C_0 and C_1 independent of N and t.
2. There exists a matrix A such that
$$\| A - A_N \| < C_2\,\delta^{-N},$$

where C_2 is a constant independent of N. The matrix A has an eigenvalue ℓ, and the absolute values of eigenvalues of A other than ℓ are less than ℓ. Furthermore, each Jordan cell with the eigenvalue ℓ is one dimensional.

The condition 1 implies that

$$F(D_N\vec{1}) = D_{N-1}\vec{1}\,,\ N = 1,2,3,\cdots,$$
$$A_N = \nabla F(D_N\,\vec{1})\,,\ N = 1,2,3,\cdots,$$

and the condition 2 implies that the limit

$$B_n \stackrel{d}{=} \lim_{N\to\infty} \ell^{-N+n} A_{n+1}A_{n+2}\cdots A_N \tag{4}$$

exists (see Appendix).

The conditions 1 and 2 roughly state that for each $N \in \mathbf{Z}_+$, the function F is analytic in a circle of center $D_N\vec{1}$ and radius $O(\delta^{-N})$ and that the first order term of F in z dominates the higher order terms.

Note that the conditions do not exclude the possibility that the function F is singular at the point $z_0 \stackrel{d}{=} \lim_{N\to\infty} D_N\vec{1}$.

Remark. We can extend our results to the case where the function F has N-dependences; i.e. with the replacement $F \to F_N$ in the above definition and in eq. (3), Theorem 1 and 2 below still hold.

Definition 3. Let $\{D_N\}$, $N = 0,1,2,\cdots$, be a series of d-dimensional diagonal matrices of order $\delta\,(>1)$. We call X_N a branching process of type $(d,\{D_N\},\ell)$, if X_N is a discrete-time branching process with d types, and if the generating functions defined by eq. (1) satisfy the recursion equations eq. (2) and eq. (3) for some $(\{D_N\},\ell)$-regular function F.

The conditions in Definition 1 and Definition 2.2 together with $\ell > \delta$ are an extension of the notion of supercriticality. Definition 2.1, on the other hand, is a rather restrictive condition implying the existence of moments of X_N.

In this paper we prove the following.

Theorem 1. *Let X_N be a branching process of type $(d,\{D_N\},\ell)$. Then $\ell^{-N}\,D_N^{-1}\,X_N$ converges weakly to a random variable $Y = (Y_1,\cdots,Y_d)$ as $N\to\infty$.*

$$E[Y_j \mid X_n = e_i] = \ell^{-n}\,\left(D_n^{-1}\,B_n\right)_{ij}\,,$$
$$i = 1,2,\cdots,d\,,\ j = 1,2,\cdots,d\,,\ n \in \mathbf{Z}_+\,.$$

Example. Let X_N be a branching process with 2 types such that the transition probability

$$P_N(\alpha,\beta) = \mathbf{Prob}\{X_{N+1}=\beta \mid X_N=\alpha\}\,,\quad \alpha\in\mathbf{Z}_+{}^2\,,\ \beta\in\mathbf{Z}_+{}^2\,,\ N\in\mathbf{Z}_+\,,$$

is determined by

$$\begin{aligned}
P_N((1,0),(2,0)) &= \frac{(1-\epsilon'_{N+1})^2}{1-\epsilon'_N}\,,\\
P_N((1,0),(k,2)) &= \frac{3(1-\epsilon'_{N+1})^k\epsilon_{N+1}^2}{25(1-\epsilon'_N)}\,,\ k=0,1,2,\cdots,\\
P_N((1,0),\beta) &= 0\,,\ \text{otherwise},\\
P_N((0,1),(k,2)) &= \frac{(1-\epsilon'_{N+1})^k\epsilon_{N+1}^2}{\epsilon_N}\,,\ k=0,1,2,\cdots,\\
P_N((0,1),\beta) &= 0\,,\ \text{otherwise}.
\end{aligned}$$

Here, $\epsilon'_N,\ \epsilon_N,\ N=0,1,\cdots$, are positive constants, defined recursively by $F(1-\epsilon'_{N+1},\epsilon_{N+1})=(1-\epsilon'_N,\epsilon_N)\,,\ N\in\mathbf{Z}_+$, with given positive constants ϵ'_0 and ϵ_0, satisfying $\epsilon'_0<\frac{1}{2}$ and $\epsilon_0<\frac{1}{2}$, and F is defined by

$$F(x,y) = (x^2+\frac{3}{25}\frac{y^2}{1-x}\,,\ \frac{y^2}{1-x})\,.$$

It is easy to see that X_N is a branching process of type $(2,\{D_N\},4)$ with $D_N=\mathrm{diag}(1-\epsilon'_N,\epsilon_N)$. The conditions in the definitions are satisfied with $\delta=\frac{5}{3}$. Also it follows that $\epsilon_N = C\,\delta^{-N}\,(1+O(\delta^{-N}))$ and $\epsilon'_N = C\,\delta^{-N-1}\,(1+O(\delta^{-N}))$, where C is a constant. Theorem 1 therefore implies that $4^{-N}\,((1-\epsilon'_N)^{-1}\,X_{N,1},\epsilon_N^{-1}\,X_{N,2})$ converges weakly as $N\to\infty$. Using the asymptotic form of ϵ'_N and ϵ_N , it further follows that $4^{-N}\,X_{N,1}$ and $(12/5)^{-N}\,X_{N,2}$ converges, or in other words, the average number of type 1 will grow at the rate of 4 per generation asymptotically, while that of type 2 will grow at the rate of 12/5 per generation.

Note that though the birth rate of type 2 from type 1 disappears in the limit $N\to\infty$, the birth affects the average growth rate significantly in the limit, or otherwise the average growth rate for type 1 would have been 2 instead of $\ell=4$. This is because a birth of type 2 will in turn give birth of order $\frac{1}{\epsilon'_N}$ of type 1 in average, which diverges as N is increased.

For the one-type case $d=1$, the branching process of type $(d=1,\{D_N\},\ell)$ falls into the class considered by Biggins and D'Souza [3]. In this case, the conditions in Definition 1 and Definition 2.2 together with $\ell>\delta$ imply the "uniform supercriticality", while Definition 2.1 implies the "dominance condition". (We learned this from Prof. Biggins.) A result in [3] then implies that (for $d = 1$) the process has a single rate of growth.

As is seen in the above example, the conditions (for $d > 1$) do *not* imply the same rate of growth *among the different types.* The gap between the upper and lower bound of $D_{N,i}$ in Definition 1 is crucial for such a possibility.

We are particularly interested in the situation where growth rate of different types differ, and where the birth of less increasing types affects the growth rates of dominant types. Non-stationarity and multi-typedness both enter in a non-trivial way in this phenomena.

The present study in the particular class of branching processes grew out of our attempt in constructing a non-self-similar diffusion on the Sierpinski gasket. We briefly discuss the application in Section 3 (the details, however, will be left for the future publication).

Acknowledgement. The authors would like to thank Prof. M. T. Barlow, Prof. K. Hattori, Prof. J. Kigami, Prof. T. Kumagai, Prof. S. Kusuoka, and Prof. Y. Takahashi for valuable discussions concerning the present work. They would also like to thank Prof. J. D. Biggins who kindly informed them of the connection between the present work and his results on one-type varying environment Galton-Watson processes.

1. Main results.

Let $\{\phi_{n,N}(z)\}$ be a sequence of $\mathbf{C}^d$-valued functions defined by the recursion relation eq. (2) and eq. (3) for a $\mathbf{C}^d$-valued $(\{D_N\}, \ell)$-regular function F. Put

$$\begin{aligned} f_{n,N}(s) &= D_n \phi_{n,N}(\exp(-\ell^{-N} D_N^{-1} s)) \\ &= F^{N-n}(D_N \exp(-\ell^{-N} D_N^{-1} s)) \,. \end{aligned}$$

Theorem 2. *There exist positive constants ϵ and C such that, for every $n \in \mathbf{Z}_+$:*

(i) The function $f_{n,N}(s)$, $N \geq n$, is analytic on $\parallel s \parallel < \epsilon\, \ell^n \delta^{-n}$ and converges uniformly to an analytic function $f_n(s)$ on $\parallel s \parallel < \epsilon\, \ell^n \delta^{-n}$ as $N \to \infty$.

(ii) The function $f_n(s)$ has the expression

$$f_n(s) = D_n \vec{1} - \ell^{-n} B_n\, s + \tilde{f}_n(s)\,, \quad \parallel s \parallel < \epsilon\, \ell^n \delta^{-n}\,,$$

with the bound

$$\parallel \tilde{f}_n(s) \parallel \leq C\, \delta^n \ell^{-2n} \parallel s \parallel^2, \quad \parallel s \parallel < \epsilon\, \ell^n \delta^{-n}\,.$$

Theorem 1 is obtained from Theorem 2 by a standard argument.

Proof of Theorem 1 assuming Theorem 2. Put $\mathbf{C}_+ \stackrel{d}{=} \{s \in \mathbf{C} \mid \Re(s) \geq 0\}$, For each $n \in \mathbf{Z}_+$ and $N \in \mathbf{Z}_+$, satisfying $n \leq N$, and for each $i \in \{1, 2, \cdots, d\}$, let $Q_{n,N,i}$ be the distribution of $\ell^{-N} D_N^{-1} X_N$ conditioned by $X_n = e_i$, and let

$$\begin{aligned}\Phi_{n,N,i}(s) &= \int_{\mathbf{R}^d} \exp(-s \cdot w)\, Q_{n,N,i}(dw) \\ &= \int_{[0,\infty)^d} \exp(-s \cdot w)\, Q_{n,N,i}(dw)\,, \quad s \in {\mathbf{C}_+}^d\,,\end{aligned}$$

be the Laplace transform (generating function) of $Q_{n,N,i}$. Put $\Phi_{n,N} = (\Phi_{n,N,1}, \cdots, \Phi_{n,N,d})$. With eq. (1) we have

$$\Phi_{n,N}(s) = \phi_{n,N}(\exp(-\ell^{-N} D_N^{-1} s))\,.$$

Theorem 2 then implies that there exists a positive constant ϵ such that $\{\Phi_{n,N}\}_{N=n,n+1,n+2,\cdots}$ converges uniformly to an analytic function on $\| s \| < \epsilon\ (\leq \epsilon\, \ell^n\, \delta^{-n})$. Therefore we have for each $i \in \{1, 2, \cdots, d\}$ and each $n \in \mathbf{Z}_+$,

$$M \stackrel{d}{=} \sup_N \{ \int_{[0,\infty)^d} \exp(\epsilon \vec{1} \cdot w/(2\sqrt{d})) Q_{n,N,i}(dw) \} < \infty\,.$$

Fix $R > 0$ and put $E_j \stackrel{d}{=} \{w \in [0,\infty)^d \mid w_j \geq R\}$, $j = 1, 2, \cdots, d$. Then we have

$$\begin{aligned}M &\geq \sup_N \{ \int_{E_j} \exp(\epsilon \vec{1} \cdot w/(2\sqrt{d})) Q_{n,N,i}(dw) \} \\ &\geq \exp(\epsilon R/(2\sqrt{d})) \sup_N Q_{n,N,i}(E_j)\,.\end{aligned}$$

Hence

$$\begin{aligned}\sup_N Q_{n,N,i}([0,R]^d) &\geq 1 - \sum_j \sup_N Q_{n,N,i}(E_j) \\ &\geq 1 - M\, d\, \exp(-\epsilon R/(2\sqrt{d}))\,.\end{aligned}$$

The right hand side converges to 1 as $R \to \infty$. Therefore the sequence of measures $\{Q_{n,N,i}\}_{N=n,n+1,\cdots}$ is tight.

Since $\{Q_{n,N,i}\}$ is tight, there exists a subsequence that converges weakly to a probability measure. For each convergent subsequence $\{Q_{n,k_N,i}\}_{N=1,2,\cdots}$, $\{\Phi_{n,k_N,i}(s)\}$ converges uniformly on compact sets in

${\mathbf{C}_+}^d$. Therefore $\{\Phi_{n,k_N,i}(s)\}$ converges to a function that is analytic in the interior of ${\mathbf{C}_+}^d$ and continuous in ${\mathbf{C}_+}^d$.

On the other hand, Theorem 2 implies that $\{\Phi_{n,N,i}(s)\}$ converges to an analytic function in a neighborhood of $s=0$. Therefore, $\{\Phi_{n,k_N,i}\}$ converges to an analytic function independent of the choice of subsequences in the interior of ${\mathbf{C}_+}^d$. Since the limit of $\{\Phi_{n,k_N,i}\}$ is continuous in ${\mathbf{C}_+}^d$, this further implies that $\{\Phi_{n,k_N,i}\}$ converges to a continuous function independent of the choice of subsequences in ${\mathbf{C}_+}^d$. We have shown that for any convergent subsequence $\{Q_{n,k_N,i}\}$ of $\{Q_{n,N,i}\}$ the corresponding sequence of the characteristic functions $\{\Phi_{n,k_N,i}(\sqrt{-1}x)\}$ converges to a function independent of the choice of subsequences for $x\in\mathbf{R}^d$, which implies that the limit measure $Q_{n,i}$ of the sequence $\{Q_{n,k_N,i}\}$ is independent of the subsequence. Since $\{Q_{n,N,i}\}$ is tight, this implies that $\{Q_{n,N,i}\}$ converges weakly to a probability measure $Q_{n,i}$. The proof of the statement on the expectation values is straightforward. This completes the proof. ∎

2. Proof of Theorem 2.

We put, for $n_1,n_2\in\mathbf{Z}_+$ with $n_1\le n_2$,

$$B_{n_1,n_2} \stackrel{d}{=} \ell^{-n_2+n_1}A_{n_1+1}A_{n_1+2}\cdots A_{n_2}, \tag{5}$$

$$B_{n_1,n_1} \stackrel{d}{=} I. \tag{6}$$

Theorem 2 follows from Propositions 1 and 2 below.

Proposition 1. *There exist positive constants ϵ and C_3 independent of $n,N\in\mathbf{Z}_+$ such that the function $f_{n,N}(s)$, $n\le N$, has the expression*

$$f_{n,N}(s) = D_n\vec{1} - \ell^{-n}B_{n,N}s + \tilde{f}_{n,N}(s), \quad \| s \| < 2\epsilon\,\ell^n\,\delta^{-n}, \tag{7}$$

with the bound

$$\| \tilde{f}_{n,N}(s) \| \le C_3\delta^n\ell^{-2n} \| s \|^2, \quad \| s \| < 2\epsilon\,\ell^n\,\delta^{-n}. \tag{8}$$

Proposition 2. *For $n,N,N'\in\mathbf{Z}_+$ with $n\le N\le N'$, it holds that*

$$\| \tilde{f}_{n,N}(s) - \tilde{f}_{n,N'}(s) \| \le C_4 r^{-N} \| s \|^2, \quad \| s \| < \epsilon\,\ell^n\,\delta^{-n}, \tag{9}$$

where $C_4>0$ and $r>1$ are constants independent of n,N,N' and s.

Proof of Theorem 2 assuming Propositions 1 and 2. Proposition 2 implies that the family $\{\tilde{f}_{n,N}\}_{N=n,n+1,\ldots}$ constitutes a Cauchy series. Then, taking the limit $N\to\infty$ in (7) and (8), we obtain Theorem 2, where

$$\tilde{f}_n(s) = \lim_{N\to\infty}\tilde{f}_{n,N}(s). \quad ∎$$

Proof of Proposition 1. Let us prove the proposition by induction on $n = N, N-1, \cdots, 0$. Since

$$f_{N,N}(s) = D_N \exp(-\ell^{-N} D_N^{-1} s) \tag{10}$$

and since there exists a constant $C_5 > 0$ such that

$$\| \ell^{-N} D_N^{-1} s \| < C_5, \quad \| s \| < 2\epsilon \ell^N \delta^{-N},$$

we have

$$\| \tilde{f}_{N,N}(s) \| \leq \alpha \delta^N \ell^{-2N} \| s \|^2, \quad \| s \| < 2\epsilon \ell^N \delta^{-N}, \tag{11}$$

for some constant α independent of N and s.

Suppose that

$$\| \tilde{f}_{n,N}(s) \| \leq \alpha_n \delta^n \ell^{-2n} \| s \|^2, \quad \| s \| < 2\epsilon \ell^n \delta^{-n}, \tag{12}$$

holds for an $n \leq N$, where α_n is a constant independent of N and s. Let us estimate the function $f_{n-1,N}(s)$ given by

$$f_{n-1,N}(s) = F(f_{n,N}(s)) = F(D_n \vec{1} - \ell^{-n} B_{n,N} s + \tilde{f}_{n,N}(s)). \tag{13}$$

Firstly, the function $f_{n-1,N}(s)$ is well-defined by (13) for s such that

$$\| \ell^{-n} B_{n,N} s - \tilde{f}_{n,N}(s) \| < C_0 \delta^{-n}. \tag{14}$$

As is shown in Appendix, there exists a constant $C_6 > 0$ independent of n such that

$$\| B_{n,N} \| < C_6. \tag{15}$$

Then the condition (14) holds for s satisfying

$$\| s \| < 2\epsilon \ell^{n-1} \delta^{-n+1}, \tag{16}$$

if

$$C_6 \frac{\delta}{\ell} 2\epsilon + 4\epsilon^2 (\frac{\delta}{\ell})^2 \alpha_n < C_0. \tag{17}$$

Secondly, by using

$$\tilde{f}_{n-1,N}(s) = A_n \tilde{f}_{n,N}(s) + \tilde{F}_n(-\ell^{-n} B_{n,N} s + \tilde{f}_{n,N}(s))$$

and

$$\| \ell^{-n} B_{n,N} s - \tilde{f}_{n,N}(s) \| \leq (C_6 + 2\epsilon\alpha_n)\ell^{-n} \| s \|,$$

we can estimate $\tilde{f}_{n-1,N}(s)$ and reproduce (12) with n replaced by $n-1$, where we put

$$\alpha_{n-1} = \| A_n \| \delta\ell^{-2}\alpha_n + C_1\delta\ell^{-2}(C_6 + 2\epsilon\alpha_n)^2. \tag{18}$$

Since $\delta < \ell$ and

$$\| A_n \| < \ell + C_2\delta^{-n},$$

the solution to (18) with $\alpha_N = \alpha$ satisfies (17), if ϵ is sufficiently small. ∎

The proof of Proposition 2 is based on the recursion relation

$$f_{m-1,N}(s) - f_{m-1,N'}(s) = F(f_{m,N}(s)) - F(f_{m,N'}(s)) \tag{19}$$

and on the following difference property for F.

Lemma . *For $m \in \mathbf{Z}_+$ and $t, t' \in \mathbf{C}^d$ such that $\| t \|, \| t' \| < \frac{C_0}{2}\delta^{-m}$,*

$$F(D_m\vec{1} + t) - F(D_m\vec{1} + t') = (A_m + G_m)(t - t'), \tag{20}$$

where $G_m = G_m(t, t')$ obeys the bound

$$\| G_m \| \leq C_7\delta^m \max(\| t \|, \| t' \|). \tag{21}$$

Proof. The matrix element of $A_m + G_m$ is given by the corresponding element of ∇F at some point on the line segment connecting $D_m\vec{1} + t$ and $D_m\vec{1} + t'$. Note that

$$\nabla F(D_m\vec{1} + t'') = A_m + \nabla\tilde{F}_m(D_m\vec{1} + t'')$$

holds and that $\tilde{F}_m(D_m\vec{1} + t'')$ is analytic on $\{t'' \in \mathbf{C}^d | \ \| t'' \| < C_0\delta^{-m}\}$ and bounded by $C_1\delta^m \| t'' \|^2$. Now, we fix t'' such that $\| t'' \| < \frac{C_0}{2}\delta^{-m}$ and define the function $v(\tau) = \tilde{F}_m(D_m\vec{1} + t'' + \tau e_i)$ of τ analytic on $\{\tau \in \mathbf{C} | \ |\tau| < \frac{C_0}{2}\delta^{-m}\}$, where

$$\begin{aligned} e_i &= {}^t(e_{i,1}, e_{i,2}, \cdots, e_{i,d}), \\ e_{i,j} &= \delta_{i,j}. \end{aligned}$$

Then, integrating the function $v(\tau)\tau^{-2}$ along the contour $|\tau| =\| t'' \|$, we can bound $v'(0) = \frac{\partial}{\partial t''_i}\tilde{F}_m(D_m\vec{1}+t'')$. As a result, we have

$$\begin{aligned}\| \frac{\partial}{\partial t''_i}\tilde{F}_m(D_m\vec{1}+t'') \| \leq & \sup_{|\tau|=\|t''\|} \frac{\| v(\tau) \|}{\| t'' \|} \\ & \leq 4C_1\delta^m \| t'' \|, \quad \| t'' \| < \frac{C_0}{2}\delta^{-m}.\end{aligned}$$

This proves the lemma. ∎

Proof of Proposition 2. Fixing n, N, N' and s such that $n \leq N \leq N'$ and $\| s \| < \epsilon\delta^{-n}\ell^n$, we estimate $\| \tilde{f}_{m,N}(s) - \tilde{f}_{m,N'}(s) \|, m = N, N-1, \cdots, n$.

As is shown in Appendix, there exist positive constants $C_8, r' > 0$ with $1 < r' < \ell$ such that

$$\| B_{m,N} - B_{m,N'} \| < C_8 r'^{-N+m}, \quad m \leq N < N'. \tag{22}$$

Now, (7), (19), and (20) imply

$$\begin{aligned}&\tilde{f}_{m-1,N}(s) - \tilde{f}_{m-1,N'}(s) \\ &=F(f_{m,N}(s)) - F(f_{m,N'}(s)) + \ell^{-m+1}(B_{m-1,N} - B_{m-1,N'})s \\ &=(A_m + G_m)(f_{m,N}(s) - f_{m,N'}(s)) + \ell^{-m+1}(B_{m-1,N} - B_{m-1,N'})s \\ &=(A_m + G_m)(\tilde{f}_{m,N}(s) - \tilde{f}_{m,N'}(s)) + G_m\ell^{-m}(B_{m,N} - B_{m,N'})s,\end{aligned} \tag{23}$$

where G_m depends on s. By the argument obtaining (14) from (15), (16), and (17) with the ϵ replaced by $\frac{\epsilon}{2}$, we see that t and t' defined by

$$\begin{aligned}t =& - D_m\vec{1} + f_{m,N}(s) = \ell^{-m}B_{m,N}s - \tilde{f}_{m,N}(s) \\ t' =& - D_m\vec{1} + f_{m,N'}(s) = \ell^{-m}B_{m,N'}s - \tilde{f}_{m,N}(s)\end{aligned}$$

satisfy the assumption of Lemma. Then we have the bound

$$\| G_m \| \leq C_7(C_6 + C_3\epsilon)\delta^m\ell^{-m} \| s \|,$$

since (8) and (15) imply

$$\| t \|, \| t' \| \leq (C_6 + C_3\epsilon)\ell^{-m} \| s \| .$$

Then, from (23), (15), and (22), we obtain

$$\begin{aligned}\| \tilde{f}_{m-1,N}(s) - & \tilde{f}_{m-1,N'}(s) \| \\ & \leq \| A_m + G_m \| \| \tilde{f}_{m,N}(s) - \tilde{f}_{m,N'}(s) \| \\ & \quad + C_7C_8(C_6 + C_3\epsilon)\delta^m\ell^{-2m}r'^{-N+m} \| s \|^2 .\end{aligned} \tag{24}$$

On the other hand, (8) implies

$$\| \tilde{f}_{N,N}(s) - \tilde{f}_{N,N'}(s) \| \leq \| \tilde{f}_{N,N}(s) \| + \| \tilde{f}_{N,N'}(s) \| \leq 2C_3\delta^N \ell^{-2N} \| s \|^2 . \tag{25}$$

The inequalities (24) and (25) together with the bound

$$\begin{aligned} \| A_m + G_m \| &< \ell + C_2\delta^{-m} + C_7(C_6 + C_3\epsilon)(\frac{\delta}{\ell})^m \| s \| \\ &< \ell + C_2\delta^{-m} + C_7\epsilon(C_6 + C_3\epsilon)(\frac{\delta}{\ell})^{m-n} \end{aligned}$$

prove (9) with $r = \min(r', \ell/\delta)$. ∎

3. An application.

In this section, we briefly explain an application of our theorems to the construction of "asymptotically one-dimensional" diffusion processes on fractals. (The idea is announced in [5]. Details are in preparation.) Unlike the other part of this paper, we assume in this section that the reader is familiar with the works on diffusion processes on finitely ramified fractals (for example, [1][4][7][8][9]).

In the construction of diffusion processes on the finitely ramified fractals, one starts with a set of random walks on pre-fractals and obtains the diffusion process on the fractal as a weak limit of the (time-scaled) random walks. This approach was first established for the Sierpinski gasket [4][8][1]. One of the keys to this construction is decimation invariance, which means that the transition probability of the random walk is a fixed point of the corresponding renormalization group transformations (in the terminology of [6;section4.9]) in the space of transition probabilities. In general, one may assign different transition probabilities to different kinds of jumps of the random walk, in which case the space of transition probabilities become multi-dimensional. The condition that the obtained diffusion process spans the whole fractal implies that every component of the fixed point is positive (non-degenerate fixed point). In fact, Lindstrøm defined a class of fractals (nested fractals) in such a way that the corresponding renormalization group has a non-degenerate fixed point [9;section IV, V], and succeeded in constructing diffusion processes on nested fractals.

In [6;section 5.4] we introduced *abc*-gaskets as examples of the fractals where non-degenerate fixed points of the corresponding renormalization groups are absent (if the parameters a, b, c satisfy certain conditions), while they always have unstable degenerate fixed points which correspond to the random walks on (one-dimensional) chains. The problem then arises; can we construct diffusion processes on finitely ramified fractals whose renormalization groups have only degenerate fixed points?

The idea of the solution is to use the renormalization group trajectories that “emerge” from a neighborhood of the unstable degenerate fixed points. We take the 111-gasket (which is just the Sierpinski gasket), as an example to explain briefly our idea of the construction, with emphasis on how the problem is related to non-stationary branching processes.

Let G_n denote the set of vertices of the pre-Sierpinski gasket with the smallest unit being the equilateral triangle of side length 2^{-n} (see, for example, [1] for the notation), and let $G = \overline{\bigcup_{n=0}^{\infty} G_n}$ denote the Sierpinski gasket. For a process X taking values in G we define a stopping time $T_i^n(X)$, $n = 0, 1, 2, \cdots$, by $T_0^n(X) = \inf\{t \geq 0 \mid X(t) \in G_n\}$, and $T_{i+1}^n(X) = \inf\{t > T_i^n(X) \mid X(t) \in G_n \setminus \{X(T_i^n(X))\}\}$ for $i = 0, 1, 2, \cdots$. $T_i^n(X)$ is the time that X hits G_n for the i-th time, counting only once if it hits the same point more than once in a row. Denote by $W^n(X) = T_1^n(X) - T_0^n(X)$ the time interval to hit two points in G_n. For an integer n and a process X on G or on G_N for some $N > n$, decimation is an operation that assigns a walk X' on G_n defined by $X'(i) = X(T_i^n(X))$.

The constructions in [4][8][1] use the sequence of simple random walks; random walks with same transition probability in every direction. The sequence of simple random walks $\{Y_N\}$ on G_N $(N = 1, 2, 3 \cdots)$ has the property of decimation invariance; namely, the random walk Y' on G_n defined by $Y'(i) = Y_N(T_i^n(Y_N))$ has the same law as Y_n. This implies that $\{W^n(Y_N)\}$, $N = n+1, n+2, n+3, \cdots$, is a (one-type) *stationary* branching process (see, for example, [1;Lemma2.5(b)]). A limit theorem for a stationary branching process gives a necessary estimate, which, together with other ingredients, finally leads to the theorem that the sequence of processes X_N, $N = 1, 2, 3, \cdots$, defined by $X_N(t) = Y_N([\lambda^{-N} t])$, where $\lambda = E[W^0(Y_1)]$, converges weakly to a diffusion process as $N \to \infty$. The situation is similar for the case of multi-dimensional parameter space, which appears in nested fractals [8], where the limit theorems for multi-type stationary branching processes can be employed for necessary estimates [7].

As a generalization of the simple random walk let us consider a random walk $Z = Z_{N,\vec{x}}$, $\vec{x} = {}^t(x, y, z)$, on G_N defined as follows. Z is a Markov chain taking values on G_N, and the transition probability is defined in such a way that at each integer time Z jumps to one of the four neighboring points and the relative rate of the jump is $x : y : z$ for {a horizontal jump} : {a jump in $60°$ (or $-120°$) direction} : {a jump in $120°$ (or $-60°$) direction}. For a simple random walk, $x = y = z$. The parameter space P can be defined as $P = \{(x, y, z) \mid x + y + z = 1,\ x \geq 0,\ y \geq 0,\ z \geq 0\}$. There is a one to one correspondence between a point in P and a random walk.

$Z_{N,\vec{x}}$ does not have the property of decimation invariance. Instead, the decimated walk Z' defined by $Z'(i) = Z_{N,\vec{x}}(T_i^{N-1}(Z_{N,\vec{x}}))$ has the same law as $Z_{N-1,\vec{x}'}$, where $\vec{x}' = T\vec{x} = {}^t(C\,(x + yz/3),\ C\,(y + zx/3),\ C\,(z +$

$xy/3))$, $1/C = 1 + (xy + yz + zx)/3$. The map T maps P into P . The dynamical system on the parameter space P defined by T is the renormalization group. The simple random walk corresponds to $(1/3, 1/3, 1/3) \in P$ which is a non-degenerate fixed point of T . $(0, 0, 1) \in P$ is a degenerate fixed point corresponding to the random walk along a one-dimensional chain.

We choose a sequence $\{Z_{N,\vec{x}_N}\}$ as follows. Let $0 < w_0 < 1$ and define w_N , $N = 1, 2, 3, \cdots$, inductively by $w_{N+1} = (6 - w_N)^{-1} \left(-2 + 3w_N + \left(4 + 6w_N + 6w_N^2\right)^{1/2}\right)$. Note that $w_0 > w_1 > w_2 > \cdots \to 0$. Put $\vec{x}_N = (1 + 2w_N)^{-1}\ (w_N, w_N, 1)$ and consider $Z_N = Z_{N,\vec{x}_N}$. One then sees that if $n < N$, then Z_n is a decimation of Z_N . This property of $\{Z_N\}$ corresponds to the decimation invariance of simple random walks $\{Y_N\}$. The special choice of the parameters simply means $\vec{x}_{N-1} = T\vec{x}_N$. Letting $N \to \infty$, $\vec{x}_N$ approaches the degenerate unstable fixed point $(0, 0, 1)$. Put $W_1^n(Z_N) = W^n(Z_N)$, and let $W_2^n(Z_N)$ be the number of diagonal (off-horizontal) jumps in the time interval $(T_0^n(Z_N)\ ,\ T_1^n(Z_N)]$, and $W_3^n(Z_N)$ be the number of visit in the same time interval to the points from which two horizontal lines emerges. Let D_N , $N = 0, 1, 2, \cdots$, be a sequence of diagonal matrices defined by $D_N = diag\left(\dfrac{1}{1 + 3w_N}\ ,\ w_N\ ,\ \dfrac{1 + 3w_N}{2 + 2w_N}\right)$. Then by explicit calculations one obtains the following Proposition.

Proposition 3. $W_N = {}^t(W_1^n(Z_{n+N})\ ,\ W_2^n(Z_{n+N})\ ,\ W_3^n(Z_{n+N}))$, $N = 0, 1, 2, \cdots$, *is a branching process of type* $(d = 3, \{D_N\}, \ell = 6)$.

With this proposition, Theorems 1 and 2 can be applied, which take a part of the role of the limit theorems for stationary branching processes for the fixed point theories. Together with other ingredients we finally obtain the theorem that the sequence of processes X_N , $N = 1, 2, 3, \cdots$, defined by $X_N(t) = Z_N([\lambda^{-N}t])$, where $\lambda = \ell$, converges weakly to a diffusion process as $N \to \infty$. (The resulting process is different from the already known diffusion process on the Sierpinski gasket.) A Proposition similar to the Proposition 3 holds for general *abc*-gaskets, hence the Theorems 1 and 2 are generally applicable.

In contrast to the decimation invariant (fixed point) theories, where the existing limit theorems for the stationary branching processes worked, we needed limit theorems for the non-stationary multi-type branching processes. The present study grew out of such problems.

Appendix.

In what follows, we show that the matrix $B_{n,N}$ defined by (5) and (6) satisfies (15) and (22) for some constants $C_6, C_8, r' > 0$ with $1 < r' < \ell$ under the assumption 2 in Definition 2. The matrix B_n defined by (4) exists if (22) holds.

Proof of (15). Since

$$\begin{aligned}\| \ell^{-1}A_k \| &\leq \ell^{-1} \| A \| + \ell^{-1} \| A - A_k \| \\ &< 1 + \ell^{-1}C_2\delta^{-k},\end{aligned}$$

we have

$$\| B_{n,N} \| < \prod_{k=n+1}^{N} (1 + \ell^{-1}C_2\delta^{-k}) < C_6. \quad \blacksquare$$

Proof of (22). Since

$$B_{n_1,n_2} - (\ell^{-1}A)^{n_2-n_1} = \sum_{k=n_1}^{n_2-1} B_{n_1,k}\ell^{-1}(A_{k+1} - A)(\ell^{-1}A)^{n_2-k-1},$$

we have

$$\begin{aligned}\| B_{n_1,n_2} - (\ell^{-1}A)^{n_2-n_1} \| &< \sum_{k=n_1}^{n_2-1} C_6\ell^{-1}C_2\delta^{-k-1} \\ &< C_9\delta^{-n_1}.\end{aligned}$$

Then

$$\begin{aligned}&\| B_{k,N} - B_{k,N'} \| \\ &\quad \leq \| B_{k,N} - (\ell^{-1}A)^{N-k} \| + \| B_{k,N'} - (\ell^{-1}A)^{N'-k} \| \\ &\qquad + \| (\ell^{-1}A)^{N-k}(1 - (\ell^{-1}A)^{N'-N}) \| \\ &\quad < 2C_9\delta^{-k} + C_{10}(\frac{\ell_1}{\ell})^{N-k},\end{aligned}$$

where $\ell_1 < \ell$ is a positive constant such that the absolute values of eigenvalues of the matrix A except ℓ are less than ℓ_1. Therefore we have, for $m \leq k \leq N$,

$$\begin{aligned}\| B_{m,N} - B_{m,N'} \| &\leq \| B_{m,k} \| \| B_{k,N} - B_{k,N'} \| \\ &< C_6(2C_9\delta^{-k} + C_{10}(\frac{\ell_1}{\ell})^{N-k}).\end{aligned}$$

We here assume that $N - m$ is even without loss of generality and put $k = (N+m)/2$. Then the above estimate turns out to be

$$\| B_{m,N} - B_{m,N'} \| < C_6(2C_9\delta^{-\frac{N+m}{2}} + C_{10}(\frac{\ell_1}{\ell})^{\frac{N-m}{2}}).$$

Choosing $r' = \min(\delta^{1/2}, (\ell/\ell_1)^{1/2})$, we obtain (22). Obviously, $1 < r' < \ell$ holds. ∎

REFERENCES

[1] M. T. Barlow, E. A. Perkins, *Brownian Motion on the Sierpinski Gasket,* Probab. Th. Rel. Fields **79** (1988) 543-623.

[2] B. A. Cevaschanov, *Branching Processes,* Nauka, Moscow, 1971.

[3] J. C. D'Souza, J .D. Biggins, *The Supercritical Galton-Watson Process in Varying Environments on one-type varying environment Galton-Watson process,* Stoc. Proc. Appl. **42** (1992).

[4] S. Goldstein, *Random walks and diffusion on fractals,* IMA Math. Appl. **8** (1987) 121-129.

[5] T. Hattori, *Construction of asymptotically one-dimensional continuous Markov process on Sierpinski gasket,* RIMS Kokyuroku **783** (1992) 46-64 *(in Japanese).*

[6] K. Hattori, T. Hattori, H. Watanabe, *Gaussian field theories on general networks and the spectral dimensions,* Progr. Theor. Phys. Supplement **92** (1987) 108-143.

[7] T. Kumagai, *Estimates of the transition densities for Brownian motion on nested fractals,* To appear in Probab. Th. Rel. Fields.

[8] S. Kusuoka, *A diffusion process on a fractal,* Proc. Taniguchi Symp. (1987) 251-274.

[9] T. Lindstrøm, *Brownian motion on nested fractals,* Mem. Amer. Math. Soc. **420** (1990) 1-128.

Tetsuya Hattori
Department of Information Sciences
Faculty of Engineering
Utsunomiya University
Ishii-cho
Utsunomiya 321
JAPAN

Hiroshi Watanabe
Department of Mathematics
Nippon Medical School
Nakahara-ku
Kawasaki 211
JAPAN