エネルギー準位統計の数学的基礎の研究

（課題番号 13640100 ）

平成 13 年度一平成 14 年度科学研究費補助金（基盤研究 (C)(2)）

研究成果報告書

平成１５年５月

研究代表者　南 就将
（筑波大学数学系助教授）

はしがき

本研究の目的は、ランダム行列およびランダムなシュレーディンガー作用素のスペクトルの統計的性質を数学的に解明することであったが、成果としては準備的なものにとどまった感がある。この報告書には主に代表者（南）が行なった研究の成果で、論文として未出版のものを収める。全体は二部から成り、それぞれの内容は下記の通りである。なお各分担者の研究成果は既に他の媒体に発表済み、あるいは発表予定である。

第一部：特異なポテンシャルをもつ一次元シュレーディンガー作用素の定義とそのランダム系への応用について. 1994年にH.P.McKeanはホワイトノイズ（ブラウン運動の見本関数の形式的な導関数）をポテンシャル項に持つ一次元のシュレーディンガー作用素を有限領域においてディリクレおよびノイマン境界条件の下に考察し、その第一固有値の確率分布を調べた。その結果自体興味深く、より深い研究に値するが、それ以前の基礎的な問題として、超関数であるホワイトノイズをポテンシャルとするシュレーディンガー作用素の定義についてはMcKeanは全く触れていない。一般に連続関数の形式的な導関数をポテンシャルとするシュレーディンガー作用素の定式化は1977年に福島と中尾が対称形式の方法により、また1986年に南がより初等的な方法で行なっているが、福島—中尾の方法はディリクレおよびノイマン境界条件を同等には扱えず、また南の方法にはMcKeanの理論を正当化するために必要となるSturmの振動定理を証明するのに不適当な点があった。これらの問題を解決するためにスペクトル理論の最近の研究を調べたところ、Savchuk と Shkali kov (1999) による簡明な方法が有用であることがわかった。特に彼等が導入した quasi -derivative を用いると、ホワイトノイズに見られるようなポテンシャルの特異性を表面に出さずにシュレーディンガー作用素を扱うことができ、その結果一般的な境界条件の下での自己共役性、スペクトルの離散性、振動定理等を容易に示すことができ、それらを用いて McKean の理論を正当化することができた。(Sav-chuk と Shkalikov の証明には細部に不備な点があるので、それらを修正しつつ用いた。)なお、この研究は大学院生の永井克己氏との共同研究である。

第二部: Galton-Watson tree **の頂点数の分布について**. 例えば A. Khourun-zhy による最近の研究（Sparse random matrices: spectral edge and statistics of rooted trees. Adv. in Appl. Probab. 33 (2001), no.1, 124–140 ）が示すように、ランダム行列のスペクトルの揺らぎを調べるためにランダムな樹形図の頂点の個数を数える必要に迫られることがある。一方ランダムな樹形図はいわゆる Galton-Watson 過程（離散時間の分枝過程）の軌跡として生成される。（これを Galton-Watson tree とよぶことにする。）筆者は Galton-Watson tree の頂点数に関する R.Otter (1949) の先駆的な研究が何故か閑却されていたことに気づき、Otter の研究を発展させることによりいくつかの新しい結果を得た。報告書の第２部は主としてこの成果の説明に充てる。また Otter による Galton-Watson tree の数学的構成が 1986 年の J. Neveu によるものと数学的に同等であることもわかったので、付録としてその報告を行う。なお、この部分は大学院生の倉田健司氏との共同研究である。

しかしながら、以上の成果をランダム行列の計算に具体的に結びつける結果はまだ得られていない。

研究組織

研究代表者：南 就将（筑波大学数学系助教授）
研究分担者：神田 護（筑波大学数学系教授）
研究分担者：赤平 昌文（筑波大学数学系教授）
研究分担者：森田　純（筑波大学数学系教授）
研究分担者：青嶋　誠（筑波大学数学系助教授）
研究分担者：筧 知之（筑波大学数学系助教授）
研究分担者：田崎　博之（筑波大学数学系助教授）
研究分担者：土居　伸一（筑波大学数学系助教授（平成１３年度のみ。その後転出のため分担者より削除。））
研究分担者：三河　寛（筑波大学数学系講師）
研究分担者：籠屋　恵嗣（筑波大学数学系助手）

研究経費

平成１３年度：	2,100 千円
平成１４年度：	1,300 千円
計：	3,400 千円

研究発表

N. Minami (with T. Hiratsuka): Derivation of Wigner's semi-circle law for a class of matrix ensemble via Brownian motion. Tsukuba J. Math., Vol.25, No.2 (2001) 442–464

M. Akahira(with K.Takeuchi): Information inequalities in a family of uniform distributions. Ann. Inst. Stat. Math. Vol.53, No.3 (2001) 427–435

M. Akahira: Confidence interval for the difference of means; application to Behrens-Fisher type problem. Statist. Papers Vol.43, No.2 (2002) 273– 284

J. Morita (with S. Berman and Y. Yishii): Some factorizations in universal enveloping algebras of three dimensional Lie albebras and generalizaations. Canad. Math. Bull. Vol.45, No.4 (2002) 525–536

M. Aoshima (with Y. Takada and M.S. Srivastava): A two-stage procedure for estimating a linear function of k multinomial mean vectors when covariance matrices are known. Journal of Statistical Planning and Infernce Vol.100 (2002) 109–119

田崎博之：等質空間の部分多様体の積分幾何学　数学５４巻３号　(2002) 280–291

S.Doi (A.Iwatsuka and T.Mine): the uniqueness of the integarted density of states for the Schrödinger operators with magnetic fields. Math. Z. Vol.237, No.2 (2001) 335–371

H. Mikawa: On primes in arithmetic progressions. Tsukuba J. Math. Vol.25, No.1 (2001) 121–153

第一部

特異なポテンシャルをもつ一次元シュレーディンガー作用素の定義とそのランダム系への応用について

The Schrödinger operator with the white noise potential –a remark on its definition–

0. Introduction

In 1994, H.P.McKean proved that the first eigenvalue $\lambda_1(L)$ of the Schrödinger operatort with white noise potential

$$H_L(\omega) = -\frac{d^2}{dt^2} + B'_\omega(t) \quad , \quad 0 \leq t \leq L \, ,$$

under Dirichlet and Neumann boundary conditions obeys the following limiting probability law:

$$\lim_{L\to\infty} P(L\psi(-\lambda_1(L) > x) = e^{-x} \quad x \geq 0 \, ,$$

where

$$\psi(L) = (0 \vee \lambda)^{1/2} \exp(-\frac{8}{3}(0 \vee \lambda)^{3/2} \, .$$

Namely for large L, the random variable $L\psi(-\lambda_1(L))$ is approximately obeys the standard exponential law.

On the other hand, as is well known, the sample function $B_\omega(t)$ of the Brownian motion is not differential, so that the definition of the operator $H_L(\omega)$ needs some justification.

1. Definition of the operator $H_L(\omega)$ according to Savchuk and Shkalikov

Let $Q \in L^2([0, L])$ be real valued. Savchuk and Shkalikov (Math. Notes vol. 66 (1999)) gave a meaning to the expression

$$H_Q u = -\frac{d^2}{dt^2} + Q'(t)u$$

in the following way: for a differentiable function u, let

$$u^{[1]}(t) := u'(t) - Q(t)u(t)$$

be the "quasi-derivative "of u. Using this notation, we can formally rewrite $H_Q u$ in the form not involving Q' :

$$H_Q u = -\frac{d}{dt}u^{[1]} - Qu^{[1]} - Q^2 u \, .$$

Now let us define the operator H_M in $L^2([0,L])$ by letting

$$D(H_M) := \{u \in L^2([0,L])\ ;\ u \text{ and } u^{[1]} \text{ are absolutely continuous },\ H_Q u \in L^2([0,L])\}$$

$$H_M u = H_Q u \quad \text{for} \quad u \in D(H_M)\ .$$

For $\alpha \in [0,\pi)$ and $\beta \in (0,\pi]$, we define $D(H_{\alpha,\beta})$ to be the totality of those $u \in D(H_M)$ which satisfies the two boundary conditions:

$$u(0)\cos\alpha - u^{[1]}(0)\sin\alpha = 0\ ;\ u(L)\cos\beta - u^{[1]}(L)\sin\beta = 0\ ,$$

and define $H_{\alpha,\beta}$ to be the restriction of H_M to the domain $D(H_{\alpha,\beta})$. Then $H_{\alpha,\beta}$ is self-adjoint and has purely discrete spectrum. If, moreover, Q is bounded, then $H_{\alpha,\beta}$ is lower semibounded. Let $\lambda < \lambda_2 < \cdots$ be its eigenvalues. We have the following oscillation theorem: the eigenfunction $u(\cdot,\lambda_n)$ corresponding to the eigenvalue λ_n has exactly $n-1$ zeros in the interval $(0,L)$.

To prove this, we introduce new variables r and θ

$$u(t) = r(t)\sin\theta(t)\ ;\ u^{[1]}(t) = r(t)\cos\theta\ .$$

(Prüfer's transformation.) Then u has a zero at $t = t_0$ if and only if $\theta(t_0) = 0$ (mod π). $\theta(t)$ satisfies the ordinary differential equation

$$\theta'(t) = 1 + (Q(t)^2 + \lambda - 1)\sin^2\theta(t) + Q(t)\sin 2\theta$$

which does not include Q'. The rest of the argument is standard.

We obtain the definition of $H_L(\omega)$ in the introduction by letting $Q = B_\omega(\cdot)$, the sample function of the Brownian motion.

2. Other approaches.

2-2. Already in 1977, Fukushima and Nakao gave definition of the Schrödinger operator

$$H_L = -\frac{d^2}{dt^2} + Q'(t) \quad , \quad 0 \le t \le L$$

under the Dirichlet boundary condition, where $Q(L)$ is bounded and Borel. According to Fukushima and Nakao, H_L is the self-adjoint operator corresponding to the quadratic form

$$\mathcal{E}_Q(u,v) = \int_0^L u'(t)\bar{v}'(t)dt - \int_0^L \{u'(t)\bar{v}(t) + u(t)\bar{v}'(t)\}Q(t)dt$$

with domain

$$H_0^1(0,L) = \{u \in L^2(0,L);\ u \text{ is a. c. },\ u' \in L^2\ ,\ u(0+) = u(L-) = 0\}\ .$$

2-2. Let $Q(\cdot)$ be cadlag. In 1986, Minami, the present author, gave a definition of $H_Q = -\frac{d^2}{dt^2} + Q'(t)$ in $L^2(\mathbf{R})$ in essentially the same manner as Savchuk and Shkalikov. Namely I defined the function space $D(Q)$ to be the totality of those $u \in L^2(\mathbf{R})$ which are absolutely continuous, differentiable from the right, and are such that there exist $v \in L^2(\mathbf{R})$ satisfying

$$u^+(t) - u^+(0) = Q(t)u(t) - Q(0)u(0) - \int_0^t \{Q(y)u^+(y) + v(y)\}dy\ .$$

This v is uniquely determined for a given $u \in D(Q)$, and we let $H_Q u := v$.

restricting $\mathbf{R}$ to $[0, L]$, we would obtain an equivalent definition of the Schrödinger operator as that of Savchuk and Shkalikov. However, in my formulation it is sometimes very combersome to convert differential equation to integral equation. In this sense, the notion of "quasi-derivative "of Savchuk and Shkalikov is a prety good idea.

3. Some comments on McKean's results.

3-1 On the proof of McKean's limit theorem.

Consider the case of Dirichlet boundary condition. Suppose we are given a $Q \in L^2_{loc}([0,\infty))$ and consider the equation

$$H_Q u = \lambda u\ ,\ u(0) = 0\ ,\ u'(0) = u^{[1]}(0) = 1\ .$$

Let $\tau_\lambda(Q)$ be the left most zero in $(0, \infty)$ of the solution u of this equation. By the oscillation theorem, we have the following equivalence:

$$\lambda < \lambda_1(L; Q) \ \Leftrightarrow\ \tau_\lambda(Q) > L\ .$$

On the other hand, if we let $Q(\cdot) = B_\omega(\cdot)$, and if we consider the random function $z_\lambda(t,\omega) := u'(t)/u(t)$, then $\{z_\lambda(t)\}_t$ is a diffusion process with the state space $[-\infty, +\infty]$. $\tau_\lambda(B_\omega)$ can be interpreted to be the first hitting time to $-\infty$ of this diffusion process. hence the proof of McKean's theorem reduces to the analysis of this special stochastic process.

3-2 some questions.

Before closing this note, let us point out some questions.

(i) Obviously, we should try various stochastic process $\{Q_\omega(t)\}_t$ other than Brownian motion. In order to employ the technique of markov processes, we would let $\{Q_\omega(t)\}_t$ a Levy process.

(ii) Even in the white noise case, where $\{Q_\omega(t)\}_t$ is the Brownian motion, the limiting joint distribution of $(\lambda_1(L), \ldots, \lambda_k(L))$ as $L \to \infty$ deserve to be investigated. In this connection, it should be noticed that Grenkova, Molchanov and Sudarev (C.M.P. vol.90 (1983)) obtained a result of this kind when $\{Q_\omega(t)\}_t$ is the Poisson process.

(iii) Let $\psi_L(t,\omega)$ be the ground state of $H_L(\omega)$, namely the normalized eigenfunction corresponding to the first eigenvalue $\lambda_1(L)$. By scaling, we obtain random probability measure

$$\mu_\omega^L(dt) = L\psi_L(t,\omega)^2 dt$$

on the unit interval $[0,1]$. What is then the limiting behavior of μ_ω^L as $L \to \infty$? it is of interest to consider the same question for random operators of the type

$$L_\omega = -\frac{d}{dt}a(t,\omega)\frac{d}{dt}$$

instead of H_ω.

第二部

Galton-Watson tree の頂点数の分布について

1 Introduction

Let $\Pi = \{p_n\}_{n=0}^{\infty}$ be a probability distribution on non-negative integers and let $\{X_n\}_{n=0}^{\infty}$ ($X_0 = 0$) be the Galton-Watson process (or the discrete time branching process) with the offspring distribution Π. Various long time behaviors of the integer valued Markov chain $\{X_n\}$ form the main subject matter of classical treatises of Galton-Watson processes (e.g. [7], [1], [9], [14]). There, the random tree structure which is obvious in the intuitive description of the process is rather implicit, or even entirely omitted from the exposition as in [14]. In this paper, we are interested in the typical shape of the random tree, which we shall call the Galton-Watson tree, obtained as the "trajectory "of the branching mechanism giving rise to the process $\{X_n\}$. Namely a Galton-Watson tree is the random graph each of whose vertices gives birth to a random number of children according to the probability distribution Π , and independently of each other. Two vertices are adjacent if and only if one is the parent of the other. Let Z be the total number of vertices of the Galton-Watson tree and let Y_k be the number of vertices with k children. Motivated by the classical pioneering work by Otter ([13]), which does not seem to have received full appreciation, we shall investigate in some deatil the asymptotic behavior of the probability distribution of Z and $\mathcal{Y}_k := \sum_{j=0}^{k} Y_j$, $k = 0, 1, 2, \ldots$. It will be seen that $\mathcal{Y}_k$ is the total number of vertices of a Galton-Watson tree which is hidden in the original tree. We shall also prove a central limit theorem for the distribution of $(Y_0, Y_1, \ldots)$ conditioned on Z. As a corollary of this theorem, the central limit theorem due to Mahmoud for vertices of uniform binary tree ([11]) will be reproduced. Our proofs are based on the analysis of generating functions, and we rely entirely upon the Lagrange inversion formula in obtaining explicit formulas for various probabilities and upon the classical saddle point method in obtaining their asymptotic expressions.

In order to prepare a solid basis for our work, it is necessary to give a precise description of "trees", and to introduce a suitable σ-field $\mathcal{F}$ and a probability measure P on the space Ω of such trees. In this work we follow the convenient and elegant construction of the probability space $(\Omega, \mathcal{F}, P)$ due to Neveu ([12]). (It should be remaked, however, that an alternative construction of the probability space of a Galton-Watson tree had already be given by Otter [13], which turns out to be equivalent to that of Neveu (see [10]), and which even has some advantage when one considers Ω as a topological space.) Let U be the totality of finite sequence $u = j_1 j_2 \cdots j_n$ of strictly positive integers. We can write $U = \sum_{n \geq 0} \mathbf{N}^n$, where $\mathbf{N} = \{1, 2, \ldots\}$ and $\mathbf{N}^0 = \{\phi\}$, ϕ standing for the empty sequence. A tree ω is by definition a subset of U satisfying the following three conditions: (a) $\phi \in \omega$; (b) if $uj \in \omega$, then $u \in \omega$; (c) for each $u \in \omega$, there is a non-negative integer $\nu_u(\omega)$ such that $uj \in \omega \Leftrightarrow 1 \leq j \leq \nu_u(\omega)$. Here, for two sequences $u = j_1 j_2 \cdots j_n$ and $v = k_1 k_2 \cdots k_p$, we write uv for their conjunction: $uv = j_1 j_2 \cdots j_n k_1 k_2 \cdots k_p$. In particular, $uj = j_1 j_2 \cdots j_n j$. Thus ω is a graph with vertices $u \in \omega$ and edges of the form (u, uj). The first condition (a) says that ω always contain the special vertex ϕ, the root of ω. The second condition (b) expresses the essential feature of a tree. The third condition (c) says that our ω is what is called the ordered tree in combinatorics. $\nu_u(\omega)$ is the number of children of the vertex $u \in \omega$, and we follow Otter in calling it the "type" of the vertex u.

For each $u \in U$, let $\Omega_u = \{\omega \in U \mid \omega \ni u\}$ be the set of trees containing u as its vertex, and let $\mathcal{F}$ be the σ-field generated by the family $\{\Omega_u \mid u \in U\}$. For $n = 1, 2, \ldots$, we let $\mathcal{F}_n$ be the σ-field generated by Ω_u's with $|u| \leq n$, $|u|$ being the length of the sequence u. For each $u \in U$, we also define the translation $T_u : \Omega_u \mapsto \Omega$ by $T_u(\omega) = \{v \in U | uv \in \omega\}$.

Now let $(\Omega^*, \mathcal{F}^*, P^*) = \prod_{u \in U} (\mathbf{Z}_+, \Pi)$ be the product over U of the discrete probability space $(\mathbf{Z}_+, \Pi)$, $\mathcal{F}^*$ being generated by the coordinate maps ν_u^*, $u \in U$, and let the measurable map $\psi : (\Omega^*, \mathcal{F}^*) \mapsto (\Omega, \mathcal{F})$ be defined by

$$\psi(\omega^*) = \{u = j_1 \cdots j_p \mid j_{k+1} \leq \nu^*_{j_1 \cdots j_p} \ , \ 0 \leq k < p\} \ .$$

If we define $P = P^* \circ \psi^{-1}$, then for each n, $\{T_u(\omega) \ ; \ u \in \omega, |u| = n\}$ are independent under $P(\cdot | \mathcal{F}_n)$, and $X_n(\omega) := \sharp\{u \in \omega; \ |u| = n\}$, $n = 1, 2, \ldots$, form the Galton-Watson process with the offspring distribution Π. See [N] for the precise statement of this proposition.

In this setting, our Z and Y_k can be expressed as

$$Z(\omega) = \sum_{u \in U} 1_{\Omega_u} \ ; \ Y_k(\omega) = \sharp\{u \in \omega \mid \nu_u(\omega) = k\} = \sum_{u \in U} 1_{\Omega \setminus \Omega_{u(k+1)}}(\omega) \ ,$$

so that Z and Y_k are certainly $\mathcal{F}$-measurable.

Since we are interested in the probability distribution of Z and Y_k, we shall assume throughout this work that $p_0 > 0$, in which case we have $Z < \infty$ with positive probability. We also assume $p_0 + p_1 < 1$ in order to avoid the trivial case, which can be treated separately if necessary.

The outline of the present paper is as follows. In Section 2, we consider the asymptotic behavior of $P(Z = n)$ as $n \to \infty$. Under a condition on the generating function of Π, we shall show that $P(Z = n)$ has an asymptotic expansion as $n \to \infty$. If we take its main term, we can reproduce Otter's result. We also give some partial results when the above mentioned condition fails to hold. In Section 3, we treat the asymptotic behavior of $P(\mathcal{Y}_k = n)$ as $n \to \infty$, for $k = 0, 1, 2, \ldots$. In Section 4, we give some explicit formulas for the joint distribution of Y_k, $k \geq 0$ conditioned on the event $\{Z = n\}$. We then consider its limit as $n \to \infty$ and briefly give an alternative proof of Otter's law of large numbers (Thorem 6 of [13]). We shall also prove the correspondin central limit theorem. Finally some miscellaneous applications are collected in section 5.

2 Asymptotic results for $P(Z = n)$

2.1 Preliminaries

Let $f(z) = \sum_{n=0}^{\infty} p_n z^n$ be the generating function of the distribution $\Pi = \{p_n\}_{n=0}^{\infty}$, and let $\rho \geq 1$ be its radius of convergence. As anounced in the introduction, we assume $p_0 > 0$ and $p_0 + p_1 < 1$, so that $f(0) > 0$ and $f(z)$ is strictly convex on $[0, \rho)$. Let further $\mathcal{P}(z) = \sum_{n=1}^{\infty} P(Z = n) z^n = \mathbf{E}[z^Z]$ be the generating function of Z. Since a tree ω always contains the root ϕ, we have $Z \geq 1$. From the relation

$$Z(\omega) = 1 + \sum_{j=1}^{\nu_\phi(\omega)} \sum_{u \in U} 1_{\Omega_{ju}}(\omega) = 1 + \sum_{j=1}^{\nu_\phi(\omega)} Z \circ T_j(\omega) \tag{1}$$

and the conditional independence given $\mathcal{F}_1$ of $Z \circ T_j$, $j = 1, \ldots, \nu_\phi(\omega)$, we obtain

$$\begin{aligned} \mathcal{P} &= \sum_{k=0}^{\infty} P(\nu_\phi = k) \mathbf{E}\left[z^{1+Z\circ T_1+\cdots+Z\circ T_k} \mid \nu_\phi = k\right] \\ &= z \sum_{k=0}^{\infty} p_k (\mathbf{E}[z^Z])^k = z f(\mathcal{P}(z)) \ . \end{aligned}$$

Thus $w = \mathcal{P}(z)$ satisfies the functional equation $w = zf(w)$. Since we are assuming $f(0) \neq 0$, we can apply Lagrange inversion formula (see e.g. [8]), to show that $P(Z = n)$, which is the n-th coefficient of $\mathcal{P}(z)$, is given by

$$P(Z = n) = \frac{1}{n} \mathrm{Res}\left[\left(\frac{f(z)}{z}\right)^n\right] = \frac{1}{2\pi n} \oint \left(\frac{f(z)}{z}\right)^n dz \ , \tag{2}$$

where $\oint$ denotes the contour integral along a circle surounding the origin. (2) shows that $P(Z = n)$ is the coefficient of z^{n-1} of the power series $f(z)^n$ divided by n. But $f(z)^n$ is the generating function of Π^{*n}, hence $P(Z = n)$ is equal to the probability that the 1st generation of our galton-Watson process consists of $n-1$ individuals when there are n individuals at the 0-th generation. In other words, if we let $\{q(n,m)\}_{n,m\geq 0}$ the transition probability of the Galton-Watson process with the offspring distribution Π, then we have

$$P(Z = n) = \frac{1}{n} q(n, n-1) \ . \tag{3}$$

By purely probabilistic argument, Dwass ([5]) obtained this seemingly non-trivial relation. The basic functional equation $\mathcal{P}(z) = zf(\mathcal{P}(z))$ seems to date back to the work of Hawkins and Ulam (see the footnote on page 32 of [7]), and was independently re-discovered by Good ([6]) and Otter ([13]), the latter having noted the formula (2) at the same time. This was also re-discovered by Boyd ([2]), who used it to give an alternative proof of the relation (3).

Now let α ($\alpha \geq 1$) be the radious of convergence of the power series $\mathcal{P}(z)$ and let $a = \mathcal{P}(\alpha-)$.

Proposition 1 *Under the condition* $p_0 > 0$ *and* $p_0 + p_1 < 1$, *we have*

(i) $a \leq \rho$;

(ii) α *and* a *are finite;*

(iii) $f(a)/a = 1/\alpha$;

(iv) $f'(a) \leq 1/\alpha$;

(v) $f'(a) = 1/\alpha$ *if* $a < \rho$;

(vi) *if, conversely,* $f'(\zeta) = f(\zeta)/\zeta$ *for some* $0 < \zeta < \rho$, *then* $a < \rho$ *and* $a = \zeta$.

Proof. All the statements except (vi) are proved in [13]. To show (vi), suppose $a = \rho$. Then as z increases from 0 to α, $\mathcal{P}(z)$ strictly increases from 0 to ρ. Hence there exists a $z_0 \in (0, \alpha)$ such that $\zeta = \mathcal{P}(z_0)$. If we put $z = z_0$ in the formula $zf(\mathcal{P}(z)) = \mathcal{P}(z)$ and in $zf'(\mathcal{P}(z))\mathcal{P}'(z) + f(\mathcal{P}(z)) = \mathcal{P}'(z)$ which is obtained by differentiating the former with respect to z, then by $f'(\zeta) = f(\zeta)/\zeta$, we get $\mathcal{P}'(z) + f(\zeta) = \mathcal{P}'(z)$, whence $f(\zeta) = 0$. But this is a contradiction because $f(z)$ is increasing and $f(0) = p_0 > 0$. Thus we must have $a < \rho$, and by (iii) and (v), $f'(z) = f(z)/z$ is satisfied also at $z = a$. But since $(zf'(z) - f(z))' = zf''(z) > 0$ on $(0, \rho)$, such z is unique, so that $\zeta = a$.

We shall refer the condition of the statement (vi) as *Condition A*. Since $(zf'(z) - f(z)|_{z=0} = -p_0 < 0$, it is obviously equivalent to

$$\lim_{z \uparrow \rho}(zf'(z) - f(z)) > 0 \,. \tag{4}$$

It is also easy to see that Condition A holds if and only if $f(z)/z$ attains a unique minimum in $(0, \rho)$, which is $f(a)/a$.

When, on the other hand, Condition A fails, we have $\rho = a < \infty$, and $f(z)/z$ is non-increasing on $(0, a]$. Hence the number a is always characterized by

$$\frac{f(a)}{a} = \inf_{0<z<\rho} \frac{f(z)}{z} \,. \tag{5}$$

As a trivial a priori estimate for $P(Z = n)$, we have the following

Proposition 2 *(i) Let d be the greatest common diviser of all $n \geq 0$ such that $p_n > 0$. Then $P(Z = n) = 0$ if $n \not\equiv 1 \pmod d$.*
(ii) $\limsup_{n\to\infty} P(Z = n)^{1/n} = f(a)/a$, where a is given by (5).

Proof. (i) This is a part of Theorem 4 of [13]. Here we give an alternative proof.

In the expansion

$$\left(\frac{f(z)}{z}\right)^n = \sum_{j_1,j_2,\ldots,j_n\geq 0} p_{j_1}p_{j_2}\cdots p_{j_n} z^{j_1+j_2+\cdots+j_n-n} \,,$$

we have $p_{j_1}p_{j_2}\cdots p_{j_n} > 0$ only when all $j_1, j_2, \ldots, j_n$ are integer multiple of d. Hence if $n \not\equiv 1 \pmod d$, then we have

$$j_1 + j_2 + \cdots + j_n - n \neq -1$$

for such $j_1, j_2, \ldots, j_n$, so that the residue at the origin of $(f(z)/z)^n$ vanishes and so does $P(Z = n)$ by (2). This proves (i). The assertion (ii) is obvious from the Cauchy-Hadamard formula for the radius of convergence of power series and from (iii) of Proposition 1. Note that wehave $0 < f(a)/a \leq 1$ by $1 \leq \alpha < \infty$.

2.2 Asymptotic expansion of $P(Z = n)$ as $n \to \infty$

The folowing is an improvement of Theorem 4 of [13].

Theorem 1 *Under Condition A, $P(Z = n)$ has an asymptotic expansion of the form*

$$P(Z = n) \sim \sum_{k=1}^{\infty} c_k \left(\frac{f(a)}{a}\right)^n n^{-k-1/2} \tag{6}$$

as $n \to \infty$ keeping $n \equiv 1 \pmod d$. Here a and d are the numbers given in Propositions 1 and 2 respectively, and

$$c_1 = d\sqrt{\frac{f(a)}{2\pi f''(a)}} \ .$$

Proof. Throughout the proof, we assume $n \equiv 1 \pmod d$. Condition A means that we can take the contour $|z| = a$ in the integral of (2). Noting that $f(z)$ depends only on z^d, we can rewrite (2) as

$$P(Z = n) = \frac{ad}{2\pi n} \int_{\pi/d}^{\pi/d} \exp\{n\psi(\theta)\} e^{i\theta} d\theta \ , \tag{7}$$

where we let $\psi(\theta) = \log[f(ae^{i\theta})/ae^{i\theta}]$. From $f'(a) = f(a)/a$, it is easy to verify $\psi(0) = \log f(a)/a$, $\psi'(0) = 0$ and $\psi''(0) = -a^2 f''(a)/f(a) < 0$. Thus we are in a typical situation in which the Laplace method yields the asymptotic expansion (see [4] and [3]). Since the technique is standard, we shall only sketch the outline, omitting the details.

Pick a $\delta \in (0, \pi/d)$ sufficiently small so that for $|\theta| \le \delta$, we have the inequality $\mathbf{Re}\psi(\theta) \le -\eta\theta^2$ with an $\eta > 0$ and the expansion

$$\psi(\theta) = \log\frac{f(a)}{a} - K\theta^2 + \sum_{k=3}^{\infty} B_k \theta^n \ ,$$

where we have let $\psi''(0) = -2K$. It is easy to show $|f(ae^{i\theta})| < f(a)$ for $\theta \in [-\pi/d, \pi/d] \setminus \{0\}$, so that

$$\Delta := \sup_{\delta \le |\theta| \le \pi/d} \left|\frac{f(ae^{i\theta})}{ae^{i\theta}}\right| < \frac{f(a)}{a} \ .$$

Hence

$$\begin{aligned} P(Z = n) &= \frac{ad}{2\pi n} \int_{-\delta}^{\delta} \exp\{n\psi(\theta)\} e^{i\theta} d\theta + \mathcal{O}(n^{-1}\Delta^n) \\ &= \frac{ad}{2\pi n} \left(\frac{f(a)}{a}\right)^n \int_{-\delta}^{\delta} \exp(-nK\theta^2) \exp(i\theta + n\theta^3\beta(\theta)) d\theta + \mathcal{O}(n^{-1}\Delta^n) \end{aligned} \tag{8}$$

for large n, where $\beta(\theta) := \sum_{k\geq 3} B_k \theta^{k-3}$.

Following [4], introduce the power series

$$P(\omega,\theta) = \exp(i\theta + \omega\beta(\theta)) = \sum_{\ell+m\geq 0} c_{\ell m}\omega^\ell \theta^m$$

and the polynomial

$$P_A(\omega,\theta) = \sum_{\ell+m\leq A} c_{\ell m}\omega^\ell \theta^m .$$

Then we have

$$P(\omega,\theta) - P_A(\omega,\theta) = \mathcal{O}(\omega^{A+1}) + \mathcal{O}(\theta^{A+1})$$

as $|\omega| + |\theta| \to 0$.

We now replace δ by $\tau_n := n^{-1/3}$ in the integral of (8). The resulting error can be estimated to be $\mathcal{O}(n^{-2/3}e^{-\eta n^{1/3}})$. Then we put $P_A(n\theta^3,\theta)$ in place of $\exp(i\theta + n\theta^3\beta(\theta)) = P(n\theta^3,\theta)$. The bound of the error committed by this procedure is given by

$$\begin{aligned}
&\frac{ad}{2\pi n}\left(\frac{f(a)}{a}\right)^n \times \mathcal{O}\left(\int_0^{\tau_n} \exp(-nK\theta^2)\{(n\theta^3)^{A+1} + \theta^{A+1}\}d\theta\right) \\
= \; &\mathcal{O}\left(\left(\frac{f(a)}{a}\right)^n n^{-2-A/2}\right) ,
\end{aligned}$$

which absorbs the other smaller error terms. Thus we have for any $A > 0$,

$$P(Z=n) = \frac{ad}{2\pi n}\left(\frac{f(a)}{a}\right)^n \left[\int_{\tau_n}^{\tau_n} \exp(-nK\theta^2) P_A(n\theta^3,\theta)d\theta + \mathcal{O}(n^{-1-A/2})\right] . \tag{9}$$

Finally we extend the integral to the whole real line with the error estimated as $\mathcal{O}(e^{-\alpha n^{1/3}})$ for some $\alpha > 0$. Thus we arrive at

$$\begin{aligned}
P(Z=n) &= \frac{ad}{2\pi n}\left(\frac{f(a)}{a}\right)^n \left[\sum_{\ell+m\leq A} c_{\ell m} n^\ell \int_{-\infty}^{\infty} \exp(-nK\theta^2)\theta^{3\ell+m}d\theta + \mathcal{O}(n^{-1-A/2})\right] \\
&= \frac{ad}{2\pi}\left(\frac{f(a)}{a}\right)^n \left[\sum_{\ell+m\leq A} \gamma_{\ell m} K^{-\frac{3\ell+m+1}{2}} n^{-1-\frac{\ell+m+1}{2}} + \mathcal{O}(n^{-1-\frac{A}{2}})\right] ,
\end{aligned} \tag{10}$$

where $\gamma_{\ell m} = 0$ if $\ell + m$ is odd and $= \Gamma((3\ell + m + 1)/2)$ if $\ell + m$ is even. If we take $A = 2k - 1$, $k = 1, 2, \ldots$, then the error term becomes $O(n^{-k-3/2})$, while the final term in the sum is const.$(f(a)/a)^n n^{-k-1/2}$. The validity of (10) for all $A > 0$ thus implies the asymptotic expansion

$$P(Z = n) \sim \frac{ad}{2\pi} \sum_{k=1}^{\infty} \left(\sum_{\ell+m=2(k-1)} \gamma_{\ell m} K^{-\frac{3\ell+m+1}{2}} \right) \left(\frac{f(a)}{a} \right)^n n^{-k-\frac{1}{2}}$$

as $n \to \infty$. The coefficient for $k = 1$ is

$$\frac{ad}{2\pi} \gamma_{00} K^{-1/2} = \frac{ad}{2\pi} c_{00} \Gamma\left(\frac{1}{2}\right) \left(\frac{a^2 f''(a)}{2f(a)} \right)^{-\frac{1}{2}} = d\sqrt{\frac{f(a)}{2\pi f''(a)}} \ .$$

This completes the proof.

2.3 Some partial results when Condition A fails

When Condition A fails to hold, we have $\rho = a < \infty$, and $f(\rho) = f(a) = a/\alpha < \infty$. Also we have $f'(\rho) = f'(a) \leq 1/\alpha < \infty$, namely $f'(\rho) \leq f(\rho)/\rho$. On the other hand, $f(\rho)/\rho = f(a)/a \leq 1/\alpha \leq 1$ because of $\alpha \geq 1$. In particular, $f(z) = f(x + iy)$ is of C^1 with respect to (x, y) on the closed disk $\{|z| \leq \rho\}$, hence we can take the circle $\{|z| = \rho\}$ as the contour in (2), so that

$$P(Z = n) = \frac{\rho d}{2\pi n} \int_{-\pi/d}^{\pi/d} \exp(n\psi(\theta)) e^{i\theta} d\theta \ ,$$

where we let $\psi(\theta) = \log\{f(\rho e^{-i\theta}/\rho e^{i\theta}\}$. We shall consider two cases separately.

2.3.1 The case $f'(\rho) = f(\rho)/\rho$.

In this case, we shall assume $f^{(4)}(\rho) < \infty$. As before, we have $\psi'(0) = 0$ and

$$-K := \psi''(0) = \rho^2 f''(\rho)/f(\rho) < 0 \ ,$$

so that

$$\psi(\theta) = \psi(0) - \frac{K}{2}\theta^2 + i\beta\theta^3 + \phi(\theta) , \tag{11}$$

where $i\beta = \psi^{(3)}(0)$ with real β and $\phi(\theta) = \mathcal{O}(\theta^4)$ as $\theta \to 0$. Moreover, $\psi(0) = \log(f(\rho)/\rho)$ is the maximum value of $\Re\psi(\theta)$. Now we can procede as in §2.2, to obtain

$$\begin{aligned} P(Z=n) &= \frac{\rho d}{2\pi n}\left\{\int_{-\delta}^{\delta} \exp(n\psi(\theta))e^{i\theta}d\theta + \mathcal{O}(\Delta^n)\right\} \\ &= \frac{\rho d}{2\pi n}\left\{\int_{-\tau_n}^{\tau_n} \exp(n\psi(\theta))e^{i\theta}d\theta + \mathcal{O}\left(\left(\frac{f(\rho)}{\rho}\right)^n \int_{\tau_n}^{\delta} e^{-(\eta/2)n\theta^2}d\theta\right)\right. \\ &\quad \left. +\mathcal{O}(\Delta^n)\right\} , \end{aligned}$$

where δ, Δ, τ_n and η have the same meaning as before. Since

$$\int_{\tau_n}^{\delta} e^{-(\eta/2)n\theta^2}d\theta = \mathcal{O}(n^{-2/3}e^{-(\eta/2)n^{1/3}}) ,$$

it absorbs the error term $\mathcal{O}(\Delta^n)$. Hence taking (11) and $\tau_n = n^{-1/3}$into account, we can write

$$\begin{aligned} P(Z=n) &= \frac{\rho d}{2\pi n}\left(\frac{f(\rho)}{\rho}\right)^n \left\{\int_{-\tau_n}^{\tau_n} \exp\left(-\frac{K}{2}n\theta^2 + i\beta n\theta^3 + n\phi(\theta)\right) e^{i\theta}d\theta+\right. \\ &\quad \left. +\mathcal{O}(n^{-2/3}e^{-(\eta/2)n^{1/3}})\right\} \\ &= \frac{\rho d}{2\pi n}\left(\frac{f(\rho)}{\rho}\right)^n \left\{\int_{-\tau_n}^{\tau_n} \exp\left(-\frac{K}{2}n\theta^2 + i\beta n\theta^3\right) e^{i\theta}d\theta\right. \\ &\quad \left. +\mathcal{O}\left(\int_0^{\tau_n} e^{-\frac{K}{2}n\theta^2} n\phi(\theta)d\theta\right) + \mathcal{O}\left(n^{-2/3}e^{-(\eta/2)n^{1/3}}\right)\right\} . \end{aligned}$$

Noting $\phi(\theta) = \mathcal{O}(\theta^4)$, we can estimate

$$\int_0^{\tau_n} e^{-\frac{K}{2}n\theta^2} n\phi(\theta)d\theta = \mathcal{O}(n^{-3/2}) .$$

On the other hand,

$$
\begin{aligned}
&\int_{-\tau_n}^{-\tau_n} \exp\left(-\frac{K}{2}n\theta^2 + i\beta n\theta^3 + i\theta\right) d\theta \\
= &\int_{-\tau_n}^{-\tau_n} e^{-\frac{K}{2}n\theta^2} \cos(\beta n\theta^3 + \theta) d\theta \\
= &\int_{-\tau_n}^{-\tau_n} e^{-\frac{K}{2}n\theta^2} \left\{\cos(\beta n\theta^3)\cos\theta - \sin(\beta n\theta^3)\sin\theta\right\} d\theta \\
= &\int_{-\tau_n}^{-\tau_n} e^{-\frac{K}{2}n\theta^2} d\theta + \mathcal{O}\left(\int_{-\tau_n}^{-\tau_n} e^{-\frac{K}{2}n\theta^2}(n^2\theta^6 + \theta^2 + n\theta^3) d\theta\right) \\
= &\frac{1}{\sqrt{n}}\int_{-\infty}^{\infty} e^{-\frac{K}{2}z^2} dz + \mathcal{O}(n^{-3/2}) ,
\end{aligned}
$$

thus

$$
P(Z = n) = \sqrt{\frac{f(\rho)}{2\pi f''(\rho)}}\left(\frac{f(\rho)}{\rho}\right)^n n^{-3/2}(1 + \mathcal{O}(n^{-1})) .
$$

2.3.2 The case $f'(\rho) < f(\rho)/\rho$

In this case, we shall assume $f^{(k)}(\rho) < \infty$ for some $k \geq 3$. As before, $\psi(0) = \log(f(\rho)/\rho)$ is the maximum value of $\Re\psi(\theta)$. Moreover we have

$$
\psi'(0) = i\frac{\rho f'(\rho) - f(\rho)}{f(\rho)} =: -i\beta
$$

with $\beta > 0$ and

$$
\psi''(0) = \frac{\rho f'(\rho)}{f(\rho)^2}(\rho f'(\rho) - f(\rho)) - \rho^2\frac{f''(\rho)}{f(\rho)} < 0 .
$$

Thus letting $K = -\psi''(0)$, we can write

$$
\psi(\theta) = \psi(0) - i\beta\theta - \frac{1}{2}K\theta^2 + \psi(\theta) , \tag{12}
$$

with a $\phi \in C^k$ satisfying $\phi(\theta) = \mathcal{O}(\theta^3)$ and $\phi(0) = \phi'(0) = \phi''(0) = 0$.

We now procede, with the same notation as before,

$$
P(Z = n) = \frac{\rho d}{2\pi n}\left\{\int_{-\tau_n}^{\tau_n} \exp(n\psi(\theta))e^{i\theta} d\theta + \mathcal{O}\left(\left(\frac{f(\rho)}{\rho}\right)^n n^{-2/3}e^{-\eta n^{1/3}}\right)\right\} .
$$

If we let

$$J_n := \int_{-\tau_n}^{\tau_n} \exp(n\psi(\theta))e^{i\theta}d\theta \ ,$$

then

$$J_n = \left(\frac{f(\rho)}{\rho}\right)^n \frac{1}{\sqrt{n}} \int_{-n^{1/6}}^{n^{1/6}} e^{-i\beta\sqrt{n}}z\{e^{-\frac{K}{2}z^2} e^{n\phi(z/\sqrt{n})+iz/\sqrt{n}}\}dz \ .$$

It is easy to see that

$$\left| \frac{d^i}{dz^i} \left\{ e^{-\frac{K}{2}z^2} e^{n\phi(z/\sqrt{n})+iz/\sqrt{n}} \right\} \right| = \mathcal{O}(e^{-cz^2})$$

for some $c > 0$, so that integrating by parts k times, we obtain

$$J_n = \left(\frac{f(\rho)}{\rho}\right)^n \frac{1}{\sqrt{n}} \times \mathcal{O}(n^{-k/2}) \ ,$$

and consequently

$$P(Z = n) = \mathcal{O}\left(\left(\frac{f(\rho)}{\rho}\right)^n n^{-\frac{1}{2}-\frac{k}{2}}\right) \ ,$$

as $n \to \infty$.

For illustration, consider the case $\rho = 1$, so that $f(\rho)/\rho = 1$. The condition $f'(\rho) < f(\rho)/\rho$ means that our galton-Watson tree is subcritical. What has just been proved shows that

$$P(Z = n) = \mathcal{O}(n^{-\frac{1}{2}-\frac{k}{2}})$$

if $f^{(k)}(1) < \infty$ $(k \geq 3)$. If, for example, $p_n \sim e^{-\delta n^\gamma}$ with $\delta > 0$ and $0 < \gamma < 1$, then $f^{(k)}(1) < \infty$ for all $k \geq 3$, and consequently $P(Z = n)$ decays faster than any power of n. But still one has

$$\limsup_{n\to\infty} P(Z = n)^{1/n} = 1$$

by Proposition 2.

3 The distribution of $\mathcal{Y}_k$.

In this section, we consider $\mathcal{Y}_k(\omega) := \sum_{j=0}^{k} Y_j(\omega)$, the number of those vertices of the Galton-Watson tree ω having at most k children, $\mathcal{Y}_0(\omega)$ being in particular the number of "leaves "of ω.

Let

$$Q(z) = Q_k(z) := \mathbf{E}[1_{\{\mathcal{Y}_k<\infty\}} z^{\mathcal{Y}_k}] = \sum_{n=1}^{\infty} P(\mathcal{Y}_k = n) z^n$$

be the generating function of the distribution of $\mathcal{Y}_k$. (Note that by our assumption $p_0 > 0$, we have $P(Y_0 > 0) = 1$. Thus $P(\mathcal{Y}_k = 0) = 0$ for all $k \geq 0$.)

Since we have

$$\mathcal{Y}_k(\omega) = 1_{\{\nu\leq k\}}(\omega) + \sum_{j=1}^{\nu_\phi(\omega)} (\mathcal{Y}_k \circ T_j)(\omega) \ ,$$

we can proceed, just as in section 2.1,

$$\begin{aligned} Q(z) &= \mathbf{E}\left[z^{1_{\{\nu\leq k\}}} \prod_{j=1}^{\nu_\phi} \left(1_{\{\mathcal{Y}_k<\infty\}} \circ T_j\right) \left(z^{\mathcal{Y}_k} \circ T_j\right) \right] \\ &= \mathbf{E}\left[z^{1_{\{\nu\leq k\}}} \mathbf{E}\left[1_{\{\mathcal{Y}_k<\infty\}} \circ T_j \prod_{j=1}^{\nu_\phi} z^{\mathcal{Y}_k} \circ T_j \,\middle|\, \mathcal{F}_1 \right]\right] \\ &= \mathbf{E}\left[z^{1_{\{\nu\leq k\}}} \prod_{j=1}^{\nu_\phi} \mathbf{E}\left[1_{\{\mathcal{Y}_k<\infty\}} z^{\mathcal{Y}_k}\right] \right] \\ &= z\mathbf{E}\left[1_{\{\nu\leq k\}} Q(z)^{\nu_\phi}\right] + \mathbf{E}\left[1_{\{\nu>k\}} Q(z)^{\nu_\phi}\right] \\ &= z\sum_{j=0}^{k} p_j Q(z)^j + f(Q(z)) - \sum_{j=0}^{k} p_j Q(z)^j \ . \end{aligned}$$

Thus $w = Q(z)$ solves the equation

$$z\sum_{j=0}^{k} p_j w^j = w - f(w) + \sum_{j=0}^{k} p_j w^j \ ,$$

or

$$z = \frac{w}{g_k(w)} ,$$

where we have set

$$g_k(w) = \frac{w \sum_{j=0}^{k} p_j w^j}{w - f(w) + \sum_{j=0}^{k} p_j w^j} .$$

For later use, let us further define

$$h_k(w) = \sum_{j=0}^{k} p_j w^j \quad \text{and} \quad \varphi_k(w) = \sum_{j>k} p_j w^{j-1} . \tag{13}$$

Then

$$g_k(w) = \frac{h_k(w)}{1 - \varphi_k(w)} . \tag{14}$$

We shall now prove that $g_k(w)$ is the generating function of a probability distribution $\Pi^{(k)} := \{p_n^{(k)}\}_{n=0}^{\infty}$ which satisfies our basic assumption $p_0^{(k)} > 0$ and $p_0^{(k)} + p_1^{(k)} < 1$.

To this end, let us write

$$\begin{aligned} g_k(w) &= h_k(w) \sum_{n=0}^{\infty} \varphi_k(w)^n \\ &= \left(\frac{h_k(w)}{h_k(1)} \right) h_k(1) \sum_{n=0}^{\infty} \varphi_k(1)^n \left(\frac{\varphi_k(w)}{\varphi_k(1)} \right)^n . \end{aligned}$$

If we set

$$H(w) := \frac{h_k(w)}{h_k(1)} ;$$

$$G(\zeta) := h_k(1) \sum_{n=0}^{\infty} \varphi_k(1)^n \zeta^n ;$$

$$\psi(w) := \frac{\varphi_k(w)}{\varphi_k(1)} ,$$

then H, G, ψ are generating functions respectively of

$$P(\nu_\phi \in \bullet \mid \nu_\phi \leq k) ;$$

the geometric distribution with parameter $\varphi_k(1)$;

and

$$P(\nu_\phi - 1 \in \bullet \mid \nu_\phi > k) .$$

Consequently, if X, N, $S_1, S_2, \ldots$ are independent integer-valued random variables such that X [resp. N, S_j] H [resp. G, ψ] as its generating function, then $g_k(w) = G(\psi(w))H(w)$ is the generating function of the random variable

$$\sum_{j=1}^{N} S_j + X .$$

Obviously $g_k(0) = p_0$ for $k \geq 1$ and $= p_0/(1-p_1)$ for $k = 0$, whence $p_0^{(k)} = g_k(0) > 0$. By

$$g_k'(w) = \frac{h_k'(w)(1-\varphi_k(w)) + h_k(w)\varphi_k'(w)}{(1-\varphi_k(w))^2}$$

and by $\varphi_k(w) + h_k(w) = 1$, we get

$$g_k'(1) = \frac{h_k'(1) + \varphi_k'(1)}{h_k(1)} = 1 + \frac{f'(1) - 1}{h_k(1)} .$$

This shows that $g_k'(1) < 1$ [resp. $= 1$ or > 1] if and only if $f'(1) < 1$ [resp. $= 1$ or > 1]. On the other hand, we have

$$g_k'(0) = \begin{cases} p_0p_2/(1-p_1)^2 & (k=0) \\ p_1 + p_0p_2 & (k=1) \\ p_1 & (k>1) \end{cases} .$$

Hence for $k > 1$, $p_0^{(k)} + p_1^{(k)} = g_k(0) + g_k'(0) = p_0 + p_1 < 1$ by the assumption. Similarly for $k = 1$,

$$g_1(0) + g_1'(0) = p_0 + p_1 + p_0p_2 < p_0 + p_1 + p_2 \leq 1 .$$

To consider the case $k = 0$, let $q_2 = \sum_{n \geq 2} p_n > 0$. Then

$$\begin{aligned} p_0(1-p_1+p_2) &= (1-p_1-q_2)(1-p_1+p_2) \\ &\leq (1-p_1-q_2)(1-p_1+q_2) \\ &= (1-p_1)^2 - q_2^2 < (1-p_1)^2 , \end{aligned}$$

and hence

$$g_0(0) + g_0'(0) = \frac{p_0}{1-p_1} + \frac{p_0 p_1}{(1-p_1)^2} = \frac{p_0(1-p_1+p_2)}{(1-p_1)^2} < 1 \ .$$

Finally, if $t > 0$ satisfies $f(t) = t$, then $h_k(t) + t\varphi_k(t) = t$ for any $k \geq 0$, namely

$$g_k(t) = \frac{h_k(t)}{1-\varphi_k(t)} = t \ .$$

These considerations can be summerized in the following theorem:

Theorem 2 *For any $k \geq 0$, the distribution of $\mathcal{Y}_k$ is equal to that of the total progeny of the Galton-watson tree with the offspring distribution $\Pi^{(k)} = \{p_n^{(k)}\}_{n=0}^{\infty}$. As the original $\Pi = \{p_n\}_{n=0}^{\infty}$, $\Pi^{(k)}$ satisfies the condition $p_0^{(k)} > 0$, $p_0^{(k)} + p_1^{(k)} < 1$. The new Galton-Watson tree is sub-critical [resp. critical or super-critical] if and only if the roiginal one is so. Moreover the extinction probability of both Galton-watson trees coincide.*

Corollary 1 $P(Z = \infty \ , \ \mathcal{Y}_{<\infty}) = 0$.

Proof. Obviously $\{Z < \infty\} \subset \{\mathcal{Y}_k < \infty\}$. But the probability of both events, being the extinction probability of the old and the new Galton-Watson tree respectively, are equal.

Remark 1 *The new Galton-Watson tree, the total progeny of which has the same distribution as $\mathcal{Y}_k$, can actually constructed as a functional on Ω, the probability space of the original Galton-Watson tree. In other words, there hides another Galton-Watson tree in the original one which has $\mathcal{Y}_k$ as its total progeny. the discussion of the construction of this hidden tree will be postponed to other opportunity.*

By Theorem 2, the analysis of the asymptotic behavior of $P(\mathcal{Y}_k = n$ as $n \to \infty$ is essentially reduced to that of $P(Z = n)$ with a proper choice of offspring distribution. Namely if we let σ_k be the radius of convergence of the power series $g_k(w)$ and if we define $b_k \in (0, \sigma_k]$ following (??), that is by the relation

$$\inf_{0<w<\sigma_k} \frac{g_k(w)}{w} = \frac{g_k(b_k)}{b_k} ,$$

then we have

Proposition 3 *(i) Let $d(\Pi^{(k)})$ be the greatest common divisor of all $n \geq 0$ such that $p_n^{(k)} > 0$. Then $P(\mathcal{Y}_k = n) > 0$ only for those n such that $n \equiv 1$ (mod $d(\Pi^{(k)})$).*

(ii) $\limsup_{n\to\infty} P(\mathcal{Y}_k = n)^{1/n} = g_k(b_k)/b_k$.

(iii) If $0 < b_k < \sigma_k$, then the following asymptotic expansion holds as $n \to \infty$ keeping $n \equiv 1$ (mod $d(\Pi^{(k)})$):

$$P(\mathcal{Y}_k = n) \sim \sum_{j=1}^{\infty} c_j^{(k)} \left(\frac{g_k(b_k)}{b_k}\right)^n n^{-j-1/2},$$

where

$$c_1^{(k)} = d(\Pi^{(k)}) \sqrt{\frac{g_k(b_k)}{2\pi g''(b_k)}} .$$

There is no simple relation between $d(\Pi)$ and $d(\Pi^{(k)})$, but we can prove the following proposition:

Proposition 4 *$d(\Pi^{(k)})$ is giben by*

$$d(\Pi^{(k)}) = g.c.d. \left[\{0 \leq j \leq k \mid p_j > 0\} \cup \{j \geq k \mid p_{j+1} > 0\}\right] .$$

Proof. Note that

$$d(\Pi^{(k)}) = \max\{\ell \geq 1 \mid g_k(e^{\frac{2\pi}{\ell} i}) = 1\} .$$

On the other hand, $g_k(w) = 1$ is equivalent to $h_k(w) = 1 - \varphi_k(w)$, namely to

$$\Gamma_k(w) := \sum_{j=0}^{k-1} p_j w^j + (p_k + p_{k+1}) w^k + \sum_{j=k+1}^{\infty} p_{j+1} w^{j+1} = 1 ,$$

hence

$$
\begin{aligned}
& d(\Pi^{(k)}) \\
= \; & \max\{\ell \geq 1 \mid \Gamma_k(e^{\frac{2\pi}{\ell}i}) = 1\} \\
= \; & \begin{cases} g.c.d.\,[\{0 \leq j \leq k \mid p_j > 0\} \cup \{k\} \cup \{j \geq k \mid p_{j+1} > 0\}] & \text{if } p_k + p_{k+1} > 0 \\ g.c.d.\,[\{0 \leq j \leq k \mid p_j > 0\} \cup \{j \geq k \mid p_{j+1} > 0\}] & \text{if } p_k + p_{k+1} >= 0 \;, \end{cases}
\end{aligned}
$$

which gives the desired result.

Now we shall give a condition for $0 < b_k < \sigma_k$ in terms of $\Pi = \{p_n\}_{n=0}^{\infty}$. By (14), our consideration is divided into 4 cases.

case 1. $\varphi_k(\rho-) > 1$. In this case, there is a $\sigma_k \in (0, \rho)$ such that $\varphi_k \sigma_k) = 1$. This σ_k is the radius of convergence of $g_k(w)$ and we have $g_k(\sigma_k-) = \infty$. Consequently $b_k \in (0, \sigma_k)$.

case 2. If $\varphi_k(\rho-) = 1$, then $\sigma_k = \rho < \infty$ and $g_k(\sigma_k) = \infty$. In this case also, one has $b_k \in (0, \sigma_k)$.

case 3. When $0 < \varphi_k(\rho-) < 1$, one has $\sigma = \rho < \infty$ and $g_k(\sigma_k-) < \infty$.

case 4. Finally when $\varphi_k(\rho-) = 0$, then $g_k(w) = f(w)$ and we trivially have $b_k = a$.

In the last two cases, both $b_k < \sigma_k$ and $b_k = \sigma_k$ are possible.

Proposition 5 $0 < b_k < \sigma_k$ *if and only if*

$$\lim_{z \uparrow \rho}\{(z - f(z))h_k'(z) - (1 - f'(z))h_k(z)\} > 0 \;.$$

Proof. Define

$$\Delta_k(z) := (z - f(z))h_k'(z) - (1 - f'(z))h_k(z) \;.$$

Then

$$z g_k'(z) - g_k(z) = \frac{\Delta_k(z)}{(1 - \varphi_k(z))^2} \;, \tag{15}$$

and after some computation

$$\begin{aligned}\Delta_k(z) &= \sum_{\ell=0}^{k}(\ell-1)p_\ell z^\ell + \sum_{j=k+1}^{\infty}\sum_{\ell=0}^{k}(j-\ell)p_j p_\ell z^{j+\ell-1} \\ &= \sum_{\ell=1}^{k}(\ell-1)p_\ell z^\ell + \sum_{\ell=1}^{k}\sum_{j=k+1}^{\infty}(j-\ell)p_j p_\ell z^{j+\ell-1} + p_0\left(\sum_{j=k+1}^{\infty} j p_j z^{j-1} - 1\right) \quad (16)\end{aligned}$$

From the above preliminary consideration and the condition (??) (see section 2) applied to $g_k(z)$, we see that $0 < b_k < \sigma_k$ if and only if $\varphi_k(\rho-) \geq 1$ or $\varphi_k(\rho-) < 1$, $\lim_{z\uparrow\sigma_k}\Delta_k(\sigma_k) > 0$. But in case $\varphi_k(\rho-) \geq 1$, we have $\varphi_k(\sigma_k) = 1$, namely

$$\sum_{j=k+1}^{\infty} p_j \sigma_k^{j-1} = 1 \ .$$

Since $p_0 + p_1 < 1$, this implies

$$\sum_{j=k+1}^{\infty} j p_j \sigma_k^{j-1} > 1 \ ,$$

so that $\Delta_k(\sigma_k) > 0$ by (16). Thus $0 < b_k < \sigma_k$ implies $\Delta_k(\rho-) \geq \Delta_k(\sigma_k) > 0$.

Conversely suppose $\Delta_k(\rho =) > 0$. If, in addition, $\varphi_k(\rho-) \geq 1$, then one has $0 < b_k < \sigma_k$ and if $\varphi_k(\rho-) < 1$, then since $\sigma_k = \rho$,

$$\lim_{z\uparrow\sigma_k}\{zg_k'(z) - g_k(z)\} = \frac{1}{(1-\varphi_k(\rho-))^2}\Delta_k(\rho-) > 0 \ ,$$

namely $0 < b_k < \sigma_k$.

Since $\mathcal{Y}_k \uparrow Z$ as $k \to \infty$, the number $g_k(\sigma_k)/b_k$ governing the exponential behavior of $P(\mathcal{Y}_k = n)$ is expected to increase with k. This is actually the case as the following proposition shows.

Proposition 6 *For each $k \geq 0$, we have*

$$\frac{g_k(b_k)}{b_k} \leq \frac{g_{k+1}(b_{k+1})}{b_{k+1}} \leq \frac{f(a)}{a} \ . \qquad (17)$$

If $f(z)$ is critical (i.e. $f'(1) = 1$) or sub-critical ($f'(1) < 1$) and $\rho = 1$, then each member equals to 1. If $p_{k+1} = 0$, then $g_k(z) = g_{k+1}(z)$ and the first inequality is actually an equality. If, further, $p_\ell = 0$ for all $\ell > k$, then $g_k(z) = f(z)$ and we have equality also in the second inequalty. In all the other cases, the inequality is strict.

Finally it holds that

$$\lim_{k\to\infty} \frac{g_k(b_k)}{b_k} = \frac{f(a)}{a} . \tag{18}$$

Proof. First note that

$$g_{k+1}(z) - g_k(z) = \frac{z(z - f(z))(h_{k+1}(z) - h_k(z))}{(z - f(z) + h_k(z))(z - f(z) + h_k(z))} . \tag{19}$$

For the moment, we suppose that there exists q_1 and q_2 such that $f(q_i) = q_i$, $i = 1, 2$, and $0 < q_1 < q_2 \leq \rho$. Since

$$\varphi_k(q_2) = 1 - \frac{h_k(q_2)}{q_2} < 1 ,$$

we have $q_2 \leq \sigma_k$, thus $g_k(z)$ is well defined on $[q_1, q_2]$. But for $z \in [q_1, q_2]$, we have $f(z) \leq z$, and hence $g_{k+1}(z) \geq g_k(z)$ holds on $[q_1, q_2]$ by (19). By the convexity of $f(z)/z$ and $g_k(z)/z$, and by $f(q_i)/q_i = g_k(q_i)/q_i = 1$, $i = 1, 2$, we see that a and b_k belong to $[q_1, q_2]$.

Now since we have $h_k(z) \uparrow f(z)$, we also have $g_k(z) \uparrow f(z)$ on $[q_1, q_2]$ as $k \to \infty$. By Dini's theorem, this convergence is uniform on the compact interval $[q_1, q_2]$, that is for any $\epsilon > 0$, $g_k(z)/z > (f(z)/z) - \epsilon$ holds for all $z \in [q_1, q_2]$ and for all sufficiently large k. Taking $z = b_k$ and $z = a$, we obtain

$$\frac{f(a)}{a} \geq \frac{g_k(a)}{a} \geq \frac{g_k(b_k)}{b_k} > \frac{f(b_k)}{b_k} - \epsilon \geq \frac{f(a)}{a} - \epsilon$$

for large k. This shows $\lim_{k\to\infty} g_k(b_k)/b_k = f(a)/a$.

Now suppose $f(z)$ is supercritical and let q be the extinction probability: $f(q) = q$. Then we can take $q_1 = q$, $q_2 = 1$ and (17) and (19) are proved in this case. When $f(z)$ is critical, or subcritical with $\rho = 1$, we have

$$a = f(a) = b_k = g_k(b_k) = 1$$

and there is nothing to be proved. Finally consider the case where $f(z)$ is sub-critical (i.e. $f'(1) < 1$), $\rho > 1$, but $f(z) < z$ for $1 < z < \rho$. In this case, we have in particular $\rho < \infty$ and $f(\rho) < \infty$, and $g_k(z) \uparrow f(z)$ holds uniformly on $[1, \rho]$. Thus the above argument can be applied to prove $g_k(b_k)/b_k \uparrow f(a)/a$.

Finally suppose that $f(z)$ is super-critical or sub-critical with $\rho > 1$ and that $p_{k+1} > 0$. Then $h_{k+1} > h_k(z)$ for $z > 0$ and hence we have $g_{k+1}(z) > g_k(z)$ on (q_1, q_2) or on $(1, \rho)$. Thus

$$\frac{g_{k+1}(b_{k+1})}{b_{k+1}} > \frac{g_k(b_{k+1})}{b_{k+1}} \geq \frac{g_k(b_k)}{b_k}$$

and the inequalty is strict.

4 The joint distribution of Z and $\mathcal{Y}_k$.

For complex numbers u, v_j $(j = 0, 1, \ldots)$ with $|u|, |v_j| \leq 1$, let

$$w := G(u; v_0, v_1, \ldots) := \mathbf{E}\left[Z < \infty;\ u^Z \prod_{j=0}^{\infty} v_j^{Y_j}\right] \tag{20}$$

be the joint probability generating function of $Z < Y_0 Y_1, \ldots$. Since $\sum_{j=0}^{\infty} Y_j = Z$, $\prod_{j=0}^{\infty} v_j^{Y_j}$ is actually a finite product. Due to the relations (??) and

$$Y_k(\omega) = 1_{\{\nu_\phi(\omega)=k\}} + \sum_{j=1}^{\nu_\phi(\omega)} Y_k(T_j(\omega))\ , \tag{21}$$

we can proceed, using the formula of Neveu (page 202 of [[12]]),

$$\begin{aligned} w &= \mathbf{E}\left[Z < \infty;\ u^{1+\sum_{j=1}^{\nu_\phi} Z \circ T_j} \prod_{k=0}^{\infty} v_k^{1_{\nu_\phi = k} + \sum_{j=1}^{\nu_\phi} Y_k \circ T_j}\right] \\ &= \mathbf{E}\left[\left(\prod_{j=1}^{\nu_\phi} 1_{\{Z<\infty\}} \circ T_j\right) u \left(\prod_{k=0}^{\infty} v_k^{1_{\nu_\phi=k}}\right) \left(\prod_{j=1}^{\nu_\phi} u^Z \prod_{k=0}^{\infty} v_k^{Y_k}\right) \circ T_j\right] \\ &= u\mathbf{E}\left[\prod_{k=0}^{\infty} v_k^{1_{\{\nu_\phi=k\}}} \mathbf{E}\left[\left(\prod_{j=1}^{\nu_\phi} 1_{\{Z<\infty\}} u^Z \prod_{k=0}^{\infty} v_k^{Y_k}\right) \circ T_j \middle| \mathcal{F}_1\right]\right] \end{aligned}$$

$$
\begin{aligned}
&= u\mathbf{E}\left[\prod_{k=0}^{\infty} v_k^{1_{\{\nu_\phi=k\}}} \prod_{j=1}^{\nu_\phi} \mathbf{E}\left[Z<\infty;\ u^Z \prod_{k=0}^{\infty} v_k^{Y_k}\right]\right] \\
&= u\mathbf{E}\left[\prod_{k=0}^{\infty} v_k^{1_{\{\nu_\phi=k\}}} w^{\nu_\phi}\right] \\
&= u\sum_{k=0}^{\infty} p_k v_k w^k .
\end{aligned}
$$

Thus if we define

$$
\Phi(w) = \Phi(w; v_0, v_1, \ldots) := \sum_{k=0}^{\infty} p_k v_k w^k ,
$$

then w is a solution of the equation $w = u\Phi(w)$. If we assume $v_0 \neq 0$, then $\Phi(0) = p_0 v_0 \neq 0$, and we can again apply Lagrange inversion, to obtain the formula

$$
\mathbf{E}\left[Z=n;\ \prod_{k=0}^{\infty} v_k^{Y_k}\right] = \frac{1}{2\pi i n} \oint \left\{\frac{\Phi(w)}{w}\right\}^n dw \tag{22}
$$

for the coefficients of the power series

$$
w = \sum_{n=1}^{\infty} \mathbf{E}\left[Z=n;\ \prod_{k=0}^{\infty} v_k^{Y_k}\right] u^n .
$$

Now if $Z(\omega) = n$, then we have $Y_k(\omega) = 0$ for $k \geq n$. Moreover $\sum_{k=0}^{n-1} Y_k(\omega) = 1 + \sum_{k=1}^{n-1} kY_k(\omega) = n$. Hence the probability $P(Z=n, Y_k = n_k,\ k=0,1,\ldots)$ is non-zero only for those values of n_k such that

$$
n_k = 0 \text{ for } k \geq n\ ;\ \sum_{k=0}^{n-1} n_k = 1 + \sum_{k=1}^{n-1} k n_k = n\ , \tag{23}
$$

and for such (n_k), we obtain from (23)

$$
\begin{aligned}
&P(Z=n\ ,\ Y_k = n_k\ ,\ k=0,1,\ldots) \\
&= \left(\prod_{k\geq 0} \frac{1}{n_k!}\right) \frac{\partial^n}{\partial v_0^{n_0} \partial v_1^{n_1} \cdots \partial v_{n-1}^{n_{n-1}}} \mathbf{E}\left[Z=n\ ;\ \prod_{k=0}^{n-1} v_k^{Y_k}\right]\Bigg|_{v_0=v_1=\cdots=0} \\
&= \frac{1}{2\pi i n} \left(\prod_{k\geq 0} \frac{1}{n_k!}\right) \frac{\partial^n}{\partial v_0^{n_0} \partial v_1^{n_1} \cdots \partial v_{n-1}^{n_{n-1}}} \oint \left\{\frac{\Phi(w; v_0, v_1, \ldots)}{w}\right\}^n dw\Bigg|_{v_0=v_1=\cdots=0} .
\end{aligned}
$$

On the other hand, it is easily seen that

$$
\begin{aligned}
&\frac{\partial^n}{\partial v_0^{n_0}\partial v_1^{n_1}\cdots\partial v_{n-1}^{n-1}}\,\Phi(z;v_0,v_1,\ldots)^n|_{v_0=v_1=\cdots=0}\\
=\;& n!\left(\prod_{k=0}^{n-1}p_k^{n_k}\right)w^{n-1}\,.
\end{aligned}
$$

Thus, interchanging the contour integral and the differentiation, we finally obtain the following result.

Proposition 7

$$
\begin{aligned}
&P(Z=n\ ,\ Y_k=n_k\ ,\ k=0,1,\ldots)\\
=\;&\begin{cases}(n-1)!\left(\prod_{k=0}^{n-1}\frac{p_k^{n_k}}{n_k!}\right) & \text{if } (n_k) \text{ satisfies } (23)\\ & 0 \quad \text{otherwise.}\end{cases}
\end{aligned}
$$

Let us discuss several examples.

4.1

Suppose $\Pi=\{p_k\}_{k=0}^{\infty}$ is the geometric distribution with parameter p: $p_k=(1-p)p^k$, $k=0,1,2,\ldots$. In this case $f(z)=(1-p)/(1-pz)$, and we obtain

$$
\begin{aligned}
P(Z=n) &= \frac{1}{2\pi i n}\oint\frac{1}{z^n}\left(\frac{1-p}{1-pz}\right)^n dz\\
&= \frac{1}{n!}\frac{d^{n-1}}{z^{n-1}}\left(\frac{1-p}{1-pz}\right)^n\Bigg|_{z=0}\\
&= (1-p)^n p^{n-1}\frac{1}{n}\binom{2n-2}{n-1}\,. \qquad (24)
\end{aligned}
$$

For a finite tree $\omega\in\Omega$ with $Z(\omega)=n$, we thus obtain

$$
\begin{aligned}
P(\{\omega\}|\,Z=n) &= \frac{1}{P(Z=n)}\prod_{u\in\omega}p_{\nu_u(\omega)}\\
&= \frac{1}{\frac{1}{n}\binom{2n-2}{n-1}}\,. \qquad (25)
\end{aligned}
$$

The right hand side being independent of the detailed data $\{\nu_u(\omega)\}_{u\in\omega}$ of ω, this shows that conditioned the event $\{Z = n\}$, all trees of size n are equiprobable. In other words, the total number of trees with size n is $\frac{1}{n}\binom{2n-2}{n-1}$, which is a well known result in combinatorics.

For a sequence (n_k) of non-negative integers satisfying (23), we have, by Proposition 7 and (??),

$$
\begin{aligned}
P(Y_k = n_k,\ k \geq 0 \mid Z = n) &= \frac{(n-1)!\left(\prod_{k\geq 0} \frac{(1-p)p^{kn_k}}{n_k!}\right)}{\frac{1}{n}\binom{2n-2}{n-1}(1-p)^n p^{n-1}} \\
&= \frac{\frac{1}{n}\binom{n}{n_0\ n_1\ \cdots}}{\frac{1}{n}\binom{2n-2}{n-1}},
\end{aligned}
$$

where

$$
\binom{n}{n_0 n_1 \cdots} := \frac{n!}{n_0! n_! \cdots} \quad (\sum_{j\geq 0} n_j = n)
$$

is the multinomial coefficient. By the equi-probability of the trees under the condition $\{Z = n\}$, this shows that the total number of trees ω such that $Z(\omega) = n$, $Y_k(\omega) = n_k$, $k \geq 0$ is given by

$$
\frac{1}{n}\binom{n}{n_0\ n_1\ \cdots},
$$

where (n_k) satissfies (23). This result is also well known and is attributed to Tutte [15], but it had already been stated by Otter (page 213 of [13]), though with somewhat implicit proof.

4.2

Next consider the case where Π is the binomial distribution $B(d, p)$, so that $f(z) = (1 - p + pz)^d$. Applying (2), we can easily compute

$$
P(Z = n) = \frac{1}{n}\binom{dn}{n-1}(1-p)^{dn-n+1}p^{n-1}.
$$

On the other hand, for a tree ω of size n,

$$\begin{aligned} P(\{\omega\}) &= \prod_{u\in\omega} p_{\nu_u(\omega)} = \prod_{u\in\omega} \binom{d}{\nu_u(\omega)} p^{\nu_u(\omega)}(1-p)^{d-\nu_u(\omega)} \\ &= p^{n-1}(1-p)^{dn-n+1} \prod_{u\in\omega} \binom{d}{\nu_u(\omega)}, \end{aligned}$$

and hence

$$P(\{\omega\} \mid Z = n) = \frac{1}{\frac{1}{n}\binom{dn}{n-1}} \prod_{u\in\omega} \binom{d}{\nu_u(\omega)}. \tag{26}$$

These results have the following interpretation. Consider the bond percolation on the rooted d-ary tree. Here we mean by the rooted d-ary tree an element $T \in \Omega$ such that $\nu_u(T) = d$ for all $u \in T$, and we suppose that each bond, or edge, of the graph T is "open"with a certain probability $0 < p < 1$ and is "closed"with probability $1 - p$, independently from each other. We then call a percolation cluster on T a subset γ of T consisting of those vertices $u \in T$ which are connected to the root ϕ through open bonds. A percolation cluster γ thus obtained is a "tree"which is different from our galton-watson tree defined in section 1 in the sense that each vertex $u \in \gamma$ can have k children, among d possible ones, in $\binom{d}{k}$ different ways. (When $d = 2$, the percolation clusters thus obtained is often called "uniform binary trees"(see e.g. [?]).) If we renumber the children of each vertices of a percolation cluster γ, then we shall obtain a Galton-Watson tree ω. Denote this correspondence by Ψ. When ω is finite, then there are exactly $\prod_{u\in\omega}\binom{d}{\nu_u(\omega)}$ percolation clusters which correspond to the same ω.

Now let $\mathcal{C}_n$ be the totality of percolation clusters of size n. From what we have seen above,

$$\sharp\mathcal{C}_n = \sum_{\omega\in\Omega; Z(\omega)=n} \sharp\Psi^{-1}(\{\omega\}) = \sum_{\omega\in\Omega; Z(\omega)=n} \prod_{u\in\omega} \binom{d}{\nu_u(\omega)}.$$

On the other hand, we have

$$1 = \sum_{\omega\in\Omega; Z(\omega)=n} P(\{\omega\} \mid Z = n) = \frac{1}{\frac{1}{n}\binom{dn}{n-1}} \sum_{\omega\in\Omega; Z(\omega)=n} \prod_{u\in\omega} \binom{d}{\nu_u(\omega)},$$

thus

$$\sharp\mathcal{C}_n = \frac{1}{n}\binom{dn}{n-1} .$$

If we let $E_u(\gamma)$ be the number of outward edges issued from the vertex u of a percolation cluster $\gamma \in \mathcal{C}_n$, then it is clear that

$$\sharp\{u \in \gamma |\ E_u(\gamma) = k\} = Y_k(\Psi(\gamma)) .$$

Thus if we are given $(n_k)_{k\geq 0}$ which is subject to the condition (23), then it holds that

$$\begin{aligned}
& \{\gamma \in \mathcal{C}_n \mid Y_k(\Psi(\gamma)) = n_k \ , \ k \geq 0\} \\
= & \sum_{\omega\in\Omega;\ Z(\omega)=n \ , \ Y_k(\omega)=n_k \ , \ k\geq 0} \sharp\Psi^{-1}(\{\omega\}) \\
= & \sharp\{\omega \in \Omega \mid Z(\omega) = n,\ Y_k(\omega) = n_k,\ k \geq 0\} \times \prod_{\nu=0}^{d} \binom{d}{\nu}^{n_\nu} \\
= & \frac{1}{n}\binom{n}{n_0 n_1 \cdots} \prod_{\nu=0}^{d} \binom{d}{\nu}^{n_\nu} .
\end{aligned}$$

This shows that if Q_n is the uniform distribution on $\mathcal{C}_n$, then from (26),

$$\begin{aligned}
Q_n(\sharp\{u \in \gamma |\ E_u(\gamma) = k\} = n_k,\ k \geq 0) &= \frac{\frac{1}{n}\binom{n}{n_0 n_1 \cdots}\prod_{\nu=0}^{d}\binom{d}{\nu}^{n_\nu}}{\frac{1}{n}\binom{nd}{n-1}} \\
&= P(Y_k = n_k \ , \ k \geq 0 \mid Z = n) \quad (27)
\end{aligned}$$

Later we shall apply this result to the case of uniform binary tree.

4.3

In case Π is the Poisson distribution with parameter $\lambda > 0$ so that $f(z) = e^{\lambda(z-1)}$, we have

$$P(Z = n) = e^{-\lambda n}\frac{(\lambda n)^{n-1}}{n!} ;$$

$$P(\{\omega\} \mid Z = n) = \frac{n!}{n^{n-1}} \prod_{u\in\omega} \frac{1}{\nu_u(\omega)} ;$$

$$P(Y_k = n_k \ , \ k \geq 0 \mid Z = n) = \frac{n!(n-1)!}{n^{n-1}} \prod_{k \geq 0} \frac{1}{n_k!} \left(\frac{1}{k!} \right)^{n_k} \ .$$

Similarly, when Π is the negative binomial distribution with $f(z) = \{(1-p)/(1-pz)\}^{\alpha}$, $\alpha > 0$, $0 < p < 1$, then

$$P(Z = n) = (1-p)^{\alpha n} p^{n-1} \binom{\alpha n + n - 1}{n} \ ;$$

$$P(\{\omega\} \mid Z = n) = \frac{1}{\binom{\alpha n + n - 1}{n}} \prod_{u \in \omega} \binom{\alpha + \nu_u(\omega) - 1}{\nu_u(\omega)} \ ;$$

$$P(Y_k = n_k \ , \ k \geq 0 \mid Z = n) = \frac{1}{\binom{\alpha n + n - 1}{n}} \prod_{k \geq 0} \frac{1}{n_k} \binom{\alpha + k - 1}{k}^{n_k} \ .$$

Here of course, ω has size n and (n_k) is subject to (23).

4.4

As is seen in the above examples, the conditional distribution $P(\{\omega\} \mid Z = n)$ does not depend on the parameter p or λ. This is generally true when the offspring distribution is a power series distribution. Namely, let $B(\theta) = \sum_{k \geq 0} b_k \theta^k$ be a power series with non-negative coefficients and positive radius of convergence ρ. Assuming $b_0 > 0$ and $b_k > 0$ for some $k \geq 2$, we let

$$p_k(\theta) = \frac{b_k \theta^k}{B(\theta)} \ , \ k = 0, 1, 2, \ldots$$

for $0 < \theta < \rho$. Then we have the following

Proposition 8 *Let P_θ be the probability measure for the Galton-Watson tree with offspring distribution $\Pi(\theta) = \{p_k(\theta)\}_{k=0}^{\infty}$. Then for any $n \geq 1$, the conditional distribution $P_\theta(\bullet | Z = n)$ does not depend on θ.*

Proof. The generating function of $\Pi(\theta)$ is given by $B(\theta z)/B(\theta)$. Hence by (2),

$$P(Z = n) = \frac{1}{2\pi i n} \oint \frac{1}{z^n} \left(\frac{B(\theta z)}{B(\theta)} \right)^n dz = \frac{1}{2\pi i n} \frac{\theta^{n-1}}{B(\theta)^n} \oint \frac{B(w)^n}{w^n} dw \ .$$

On the other hand, for a tree ω with $Z(\omega) = n$,

$$P(\{\omega\}) = \prod_{u\in\omega} p_{\nu_u(\omega)}8\theta) = \frac{\theta^{n-1}}{B(\theta)^n} \prod_{u\in\omega} b_{\nu_u(\omega)} ,$$

so that by dividing, the facters involving θ cancel out in $P(\{\omega\}|Z = n)$.

5 The law of large numbers and the central limit theorem for Y_k's conditioned on Z

Throughout this section, we assume that Condition A holds. The following theorem is due to Otter ([13]):

Theorem 3 *For $k = 0, 1, 2, \ldots$, we have*

$$P\left(\frac{Y_k}{n} \in \bullet \mid Z = n\right) \longrightarrow \delta_{a^k p_k / f(a)}(\bullet) ,$$

where we let $n \to \infty$ keeping $n \equiv 1$ (mod $d(\Pi)$).

Corresponding to this, it is natural to consider the central limit theorem. To state the result, let us ddefine the infinite symmetric matrix $V = (v_{jk})_{j.k=0}^{\infty}$ by

$$\begin{aligned}(\vec{t}, V\vec{t}) &= \frac{1}{f(a)} \sum_{k\geq 0} p_k t_k^2 - \frac{1}{f(a) f''(a)} \{\sum_{k\geq 1} k p_k t_k a^{k-1} - \sum_{k\geq 0} p_k t_k a^{k-1}\}^2 \\ &\quad - \frac{1}{f(a)^2} \left(\sum_{k\geq 0} p_k t_k a^k\right)^2 ,\end{aligned}$$

where $\vec{t} = (t_k)_{k=0}^{\infty}$ and $t_k = 0$ except for finite number of k's. for each $k \geq 0$, we denote $V^K = (v_{jk})_{j,k=0}^{K}$ the restriction of V. Then we obtain the following result.

Theorem 4 *For each $K \geq 0$,*

$$P\left(\left\{\sqrt{n}\left(\frac{Y_k}{n} - \frac{p_k a^k}{f(a)}\right)\right\}_{k=0}^{K} \in \bullet \mid Z = n\right) \longrightarrow N(0, V^K)$$

as $n \to \infty$ keeping $n \equiv 1$ (mod $d(\Pi)$), where $N(0, V^K)$ denotes the multi-dimensional Gaussian distribution with mean 0 and covariance matrix V^K.

As a special case, let us consider the case where $\Pi = B(2,p)$ with $0 < p < 1$, and denote by P_p the cprresponding probability measure of Galton-Watson tree. On the other hand, if Q_n is the uniform distribution on the set $\mathcal{C}_n$ of size n percolation clusters on the rooted binary tree T, then as was shown in 4.2,

$$Q_n(X_k = n_k \ , \ i = 0,1,2) = P_p(Y_k = n_k \ , \ k = 0,1,2 \mid Z = n) \ .$$

Here X_k ($k = 0,1,2$) is the number of the vertices of the percolation cluster ξ having k children. The conditional probability on the right hand side being independent of p, we set $p = 1/2$ without loss of generality. Then $a = f(a) = 1$, $p_0 = p_2 = 1/4$, $p_1 = 1/2$, and $p_k = 0$ for $k \geq 3$. With this choice of p, we now apply Theorems 3 and 4, to conclude

$$Q_n\left(\frac{X_k}{n} \in \bullet\right) \longrightarrow \delta_{p_i} \ , \ n \to \infty$$

and

$$Q_n\left\{\left(\sqrt{n}(\frac{X_0}{n} - p_0), \ \sqrt{n}(\frac{X_1}{n} - p_1), \ \sqrt{n}(\frac{X_2}{n} - p_2)\right) \in \bullet\right\} \longrightarrow N(0, C) \ ,$$

with

$$C = \frac{1}{16}\begin{pmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{pmatrix} .$$

The latter result is due to Mahmoud ([11]).

参考文献

[1] K.B. Athreya, P.E. Ney: Branching Processes

[2] A.V. Boyd: Formal power series and the total progeny in a branching proces, J.Math.Anal.Appl. Vol.34, 565–566 (1971)

[3] E.T. Copson: Asymptotic Expansions

[4] N.G. de Bruijn: Asymptotic methods in analysis

[5] M. Dwass: The total progeny in a branching process and a related random walk. J.Appl.Prob. Vol.6, 682–686 (1969)

[6] I.J. Good: The number of individuals in a cascade process. Proc.Camb.Phil.Soc. Vol.45, 360–363 (1949)

[7] Th. Harris: The theory of branching processes.

[8] J.Hofbauer: Lagrange inversion. Seminaire Lotharingian de Combinatoire, I.R.M.A. no.191/S-05, Strasbourg (1982)

[9] P. Jagers: Branching processes with biological applications

[10] K. Kurata, N. Minami: The equivalence of two constructions of Galton-Watson processes. preprint

[11] H.M. Mahmoud: The joint distribution of the three types of nodes in uniform binary trees. Algorithmica Vol.13, 313–323 (1995)

[12] J. Neveu: Arbres et processus de Galton-Watson. Ann.I.H.P. Vol.22, 199–207 (1986)

[13] R. Otter: The multiplicative processes. Ann. Math. Stat. Vol.20, 206–224 (1949)

[14] B.A. Sevast'yanov: Branching processes

[15] W. T. Tutte: The number of planted plane trees with a given partition. Amer. Math. Monthly Vol.71, 272–277 (1964)

第二部付録

Galton-Watson tree の２つの構成の同値

The Equivalence of Two Constructions of Galton-Watson processes

6 Introduction

In his elegant paper, J. Neveu[(2)] gave a convenient definition of a σ-field $\mathcal{F}$ and a probability measure P on the space Ω of trees, thus obtaining a probability space which describes every detail of a Galton-Watson branching process. Neveu was motivated by the lack of fundamental theory of this kind in the existing works on branching processes, but apparently he was not aware that already in 1949, R. Otter[(3)] had given a similar definition which turns out to be equivalent to Neveu's definition. The purpose of this note is to show this equivalence.

Otter's construction looks more honest to our intuition of branching processes, if not as concise as Neveu's, and is closely related to the topological structure of the space Ω of trees, as will be discussed in the final section of this note. Also, it is remarkable that Otter already emphasizes the conceptual advantage of realizing Galton-Watson processes on the space of trees, instead of treating them merely as Markov chains with a special kind of transition probabilities.

7 The construction by Neveu[(2)]

Let U be the totality of finite sequences of strictly positive integers, of which the empty sequence is denoted by ϕ. We can write $U = \sum_{n\geq 0}(\mathrm{N}^*)^n$, where $\mathrm{N}^* = \{1, 2, \ldots\}$, $(\mathrm{N}^*)^0 = \{\phi\}$, and $(\mathrm{N}^*)^n$ is the totality of sequences $u = j_1 j_2 \cdots j_n$ of length n. By defintion, a tree is a subset $\omega \subset U$ which satisfies the following three conditions: a) $\phi \in \omega$; b) $uj \in \omega$ implies $u \in \omega$ whenever $u \in U$ and $j \in \mathrm{N}^*$, where we write $uj = j_1 j_2 \cdots j_n j$ if $u = j_1 j_2 \cdots j_n$; c) for each $u \in \omega$, there is an integer $\nu_u(\omega) \geq 0$ such that for any $j \in \mathrm{N}^*$, $uj \in \omega$ if and only if $1 \leq j \leq \nu_u(\omega)$. Let Ω be the totality of trees in this sense. Then $\mathcal{F}$ is defined to be the smallest σ-field containing all subsets $\Omega_u := \{\omega \in \Omega \mid \omega \ni u\}$, $u \in U$.

Given a probability distribution $\Pi = \{p_n\}_{n=0}^{\infty}$ on non-negative integers N, let $(\Omega^*, \mathcal{F}^*, P^*)$ be the product space $(\mathrm{N}, \Pi)^{\otimes U}$, $\mathcal{F}^*$ being the σ- field generated by the coordinate maps $\nu_v^* : \omega^* = (\omega_u^*)_{u \in U} \mapsto \omega_v^*$, $v \in U$. If we define the map $\psi : \Omega^* \mapsto \Omega$ by

$$\psi(\omega^*) = \{u = j_1 \dots j_p \in U \mid j_{k+1} \le \nu_{j_1 \cdots j_k}(\omega^*) \ , \ \text{for } 0 \le k < p\} \ ,$$

then ψ is $\mathcal{F}^*/\mathcal{F}$-measurable, and we can consider the induced measure $P = P^* \circ \psi^{-1}$ on $(\Omega, \mathcal{F})$. The resulting probability space $(\Omega, \mathcal{F}, P)$ then describes every detail of the Galton-Watson process with the offspring distribution Π. In particular, if $Z_n(\omega)$ is the cardinality of $\omega \cap (\mathrm{N}^*)^n$, then under this probability measure P, $\{Z_n(\omega)\}_{n=0}^{\infty}$ is an N-valued Markov chain with transition probability $P(x, y) = (\Pi^{*x})(\{y\})$, where " $*$"deotes the convolution.

8 The construction by Otter[3]

Since the definition of trees by Otter[3] is obviously seen to be equivalent to Neveu's definition, we shall work on the same space Ω introduced in the previous section.

For each tree ω, we define the set $\mathcal{I}(\omega)$ of its inner points and the set $\mathcal{E}(\omega)$ of its endpoints as follows:

$$\mathcal{I}(\omega) := \{u \in \omega \mid \nu_u(\omega) > 0\} \quad ; \quad \mathcal{E}(\omega) := \{u \in \omega \mid \nu_u(\omega) = 0\} \ .$$

For two trees $\omega, \omega' \in \Omega$, we say that ω extends ω' and write $\omega \ge \omega'$, if the following two conditions hold: i) $\omega \supset \omega'$; ii) $\nu_u(\omega) = \nu_u(\omega')$ for each $u \in \mathcal{I}(\omega')$.

Next let Ω^f be the totality of finite trees. For a $T \in \Omega^f$ and $\Lambda = (\lambda_e; e \in \mathcal{E}(T)) \in \mathrm{N}^{\mathcal{E}(T)}$, define

$$[T; \Lambda] = \{\omega \in \Omega \mid \omega \ge T \ , \ \nu_e(\omega) = \lambda_e \text{ for } e \in \mathcal{E}(T)\} \ ,$$

which Otter called a "neighborhood ".

Given a probability distribution $\Pi = \{p_n\}_{n=0}^{\infty}$ on $\mathbf{N}$, Otter constructs a σ-field $\mathcal{B}$ and a probability measure Q on $(\Omega, \mathcal{B})$ in the following way.

First he let $\mathcal{S}$ be the class of subsets $S \subset \Omega$ of the form $S = \sum_{\Lambda \in A}[T; \Lambda]$ with $T \in \Omega^f$ and $A = \prod_{e \in \mathcal{E}(T)} A_e \subset \mathbf{N}^{\mathcal{E}(T)}$. Then $\mathcal{S}$ is shown to be a semi-algebra. Next he defines a set function $\tilde{Q}$ on $\mathcal{S}$ by

$$\tilde{Q}([T;\Lambda]) = \left(\prod_{u \in \mathcal{I}(T)} p_{\nu_u(T)}\right)\left(\prod_{e \in \mathcal{E}(T)} p_{\lambda_e}\right) \quad ; \quad \tilde{Q}(S) = \sum_{\Lambda \in A} \tilde{Q}([T;\Lambda]) ,$$

where the value of $\tilde{Q}(S)$ turns out to be independent of the expression of S. He then shows that $\tilde{Q}$ is countably additive on $\mathcal{S}$, and hence is extended to a probability measure Q on the σ-field $\mathcal{B}$ generated by $\mathcal{S}$ (or equivalently generated by $\mathcal{N}$).

Otter's discussion is actually quite sketchy, and in order to fill its details, it is necessary to prepare several lemmas concerning the sets $[T; \Lambda]$, which are not as trivial as regarded by Otter. Among those lemmas, the following one will be useful in the rest of this note:

Lemma 1. The class $\mathcal{N}$ defined by

$$\mathcal{N} := \{[T;\Lambda] \mid T \in \Omega^f \ , \ \Lambda \in \mathbf{N}^{\mathcal{E}(T)}\} \cup \{\emptyset\}$$

is a π-system, namely is closed under the formation of finite intersection.

Proof. We shall show that if $[T_i; \Lambda_i] \in \mathcal{N}$, $i = 1, 2$, have non-empty intersection, then we can construct a $T_3 \in \Omega^f$ and $\Lambda_3 \in \mathbf{N}^{\mathcal{E}(T_3)}$ such that

$$[T_1;\Lambda_1] \cap [T_2;\Lambda_2] = [T_3;\Lambda_3] .$$

For this purpose, let $T_3 = T_1 \cup T_2$. Then obviously $T_3 \in \Omega^f$. It is also easy to see that $\mathcal{I}(T_3) = \mathcal{I}(T_1) \cup \mathcal{I}(T_2)$, and that

$$\mathcal{E}(T_3) = [\mathcal{E}(T_1) \setminus T_2] + [\mathcal{E}(T_2) \setminus T_1] + [\mathcal{E}(T_1) \cap \mathcal{E}(T_2)] .$$

Now suppose $\omega \in [T_1; \Lambda_1] \cap [T_2; \Lambda_2]$. If $u \in \mathcal{I}(T_1) \cap \mathcal{I}(T_2)$, then since $\omega \geq T_i$, $i = 1, 2$, we have $\nu_u(T_1) = \nu_u(\omega) = \nu_u(T_2)$. In this case, we also have

$\nu_u(T_3) = \nu_u(\omega)$. If on the other hand, $u \in \mathcal{I}(T_1) \setminus \mathcal{I}(T_2)$, we see from $\omega \geq T_1$ that $\nu_u(T_1) = \nu_u(T_3) = \nu_u(\omega)$, and if $u \in \mathcal{I}(T_2) \setminus \mathcal{I}(T_1)$, we see from $\omega \geq T_2$ that $\nu_u(T_2) = \nu_u(T_3) = \nu_u(\omega)$. Thus $T_3 \geq T_i$, $i = 1, 2$, and whenever $u \in \mathcal{I}(T_3)$, one has $\nu_u(T_3) = \nu_u(\omega)$, showing $\omega \geq T_3$.

Let us write $\Lambda_i = (\lambda_e^i; e \in \mathcal{E}(T_i))$ for $i = 1, 2$. If $e \in \mathcal{E}(T_1) \setminus T_2$, then $\omega \in [T_1; \Lambda_1]$ implies $\nu_e(\omega) = \lambda_e^1$, and if $e \in \mathcal{E}(T_2) \setminus T_1$, then $\omega \in [T_2; \Lambda_2]$ implies $\nu_e(\omega) = \lambda_e^2$. Finally if $e \in \mathcal{E}(T_1) \cap \mathcal{E}(T_2)$, then we must have $\lambda_e^1 = \nu_e(\omega) = \lambda_e^2$ by $\omega \in [T_1; \Lambda_1] \cap [T_2; \Lambda_2]$. Thus if we define $\Lambda_3 = (\lambda_e^3; e \in \mathcal{E}(T_3))$ by

$$\lambda_e^3 = \begin{cases} \lambda_e^1, & \text{for } e \in \mathcal{E}(T_1) \setminus T_2 \\ \lambda_e^2, & \text{for } e \in \mathcal{E}(T_2) \setminus T_1 \\ \lambda_e^1 = \lambda_e^2, & \text{for } e \in \mathcal{E}(T_1) \cap \mathcal{E}(T_2) \ , \end{cases}$$

then we must have $\omega \in [T_3; \Lambda_3]$. Since ω was arbitrarily chosen from $[T_1; \Lambda_1] \cap [T_2; \Lambda_2]$, we conclude $[T_1; \Lambda_1] \cap [T_2; \Lambda_2] \subset [T_3; \Lambda_3]$.

To prove the converse inclusion, let $w \in [T_3; \Lambda_3]$ be arbitrary. Since $T_3 \geq T_i$, $i = 1, 2$, we have $w \geq T_i$, $i = 1, 2$. If $e \in \mathcal{E}(T_1) \setminus \mathcal{I}(T_2)$, then $e \in \mathcal{E}(T_3)$ and $\nu_e(w) = \lambda_e^3 = \lambda_e^1$. If on the other hand, $e \in \mathcal{E}(T_1) \cap \mathcal{I}(T_2)$, then $\nu_e(w) = \nu_e(T_2)$. But we are assuming $[T_1; \Lambda_1] \cap [T_2; \Lambda_2] \neq \emptyset$, so that there exists an ω such that $\omega \geq T_2$ and that $\nu_e(\omega) = \lambda_e^1$. Hence $\nu_e(T_2) = \nu_e(\omega) = \lambda_e^1$. In any case, $e \in \mathcal{E}(T_1)$ implies $\nu_e(w) = \lambda_e^1$. Thus $w \in [T_1; \Lambda_1]$. Similarly one can show $w \in [T_2; \Lambda_2]$, completing the proof of the lemma.

9 The equivalence of two constructions

We are now ready to prove our main assertion.

Proposition 1. $\mathcal{B} = \mathcal{F}$ and $Q = P$, hence the probability space constructed by Otter and Neveu coincide.

Proof. Since we can write, with the convention $u0 = u$,

$$[T; \Lambda] = \left(\bigcap_{u \in T} \Omega_u \right) \cap \left(\bigcap_{u \in \mathcal{I}(T)} \Omega^c_{u(\nu_u(T)+1)} \right) \cap \left(\bigcap_{u \in \mathcal{E}(T)} \Omega_{u\lambda_u} \setminus \Omega_{u(\lambda_u+1)} \right) ,$$

we see $[T; \Lambda] \in \mathcal{F}$ and hence $\mathcal{B} \subset \mathcal{F}$.

Conversely we can also write

$$\Omega_u = \bigcup\{[T;\Lambda] \mid u \in T \in \Omega^f \ , \ \Lambda \in \mathbf{N}^{\mathcal{E}(T)}\} \ ,$$

hence $\Omega_u \in \mathcal{B}$ for any $u \in U$, proving $\mathcal{F} \subset \mathcal{B}$.

On the other hand, since

$$\psi^{-1}([T;\Lambda]) = \left(\bigcap_{u\in\mathcal{I}(T)} \{\omega^* \mid \nu_u^*(\omega^*) = \nu_u(T)\}\right) \cap \left(\bigcap_{u\in\mathcal{E}(T)} \{\omega^* \mid \nu_u^*(\omega^*) = \lambda_u\}\right) \ ,$$

we have

$$P([T;\Lambda]) = P^*(\psi^{-1}([T;\Lambda])) = \left(\prod_{u\in\mathcal{I}(T)} p_{\nu_u(T)}\right)\left(\prod_{e\in\mathcal{E}(T)} p_{\lambda_e}\right) = Q([T;\Lambda]) \ .$$

Thus P and Q coincide on the π-system $\mathcal{N}$, hence on the λ-system $\mathcal{F} = \mathcal{B}$ which is generated by $\mathcal{N}$.

10 Ω as a metric space

As was mentioned by Otter(3), and as is obvious from Lemma 1, we can make Ω into a topological space by calling "open"those subsets $G \subset \Omega$ which are written as unions of $[T;\Lambda]$'s. In particular, our σ-field $\mathcal{F}$ is the Borel σ-field corresponding to this topology. Let us briefly show that this topology is generated by a metric on Ω.

For $n \geq 0$, let $z_n(\omega) = \omega \cap (\mathbf{N}^*)^n$. Given $\omega \ , \ \omega' \in \Omega$, let us define

$$\mu(\omega,\omega') = \sup\{n \geq 0 \mid z_n(\omega) = z_n(\omega') \ \} \ ,$$

where we let $\mu(\omega,\omega') = \infty$ if $z_n(\omega) = z_n(\omega')$ for all $n \geq 1$, namely if $\omega = \omega'$. Since we have

$$\mu(\omega,\omega'') \geq \min\{\mu(\omega,\omega') \ , \ \mu(\omega',\omega'') \ \} \ ,$$

we can define a metric d on Ω by letting $d(\omega,\omega') := \exp\{-\mu(\omega,\omega')\}$.

Now let $G \subset \Omega$ be open in the sense already defined. Then for each $\omega \in G$, we can choose a $[T;\Lambda]$ such that $\omega \in [T;\Lambda] \subset G$. Let $h(T) = \max\{n \geq$

$0|T \cap (\mathrm{N}^*)^n \neq \emptyset\}$ be the "height "of the finite tree T. Then $\omega' \in [T;\Lambda]$ as soon as $z_n(\omega') = z_n(\omega)$ for $0 \leq n \leq h(T)+1$, or equivalently as soon as $d(\omega',\omega) \leq \exp\{-(h(T)+1)\}$. This shows that G is open with respect to the metric d. Conversely, suppose G is open in the metric d. Then for each $\omega \in G$, there is an integer $m \geq 1$ such that $\omega' \in G$ whenever $z_n(\omega') = z_n(\omega)$, $0 \leq n \leq m$. If we define $T = \cup_{n=0}^{m-1} z_n(\omega)$ and $\Lambda = (\lambda_e \; ; \; e \in \mathcal{E}(T))$ with $\lambda_e = \nu_e(\omega)$, then the last condition is equivalent to $\omega' \in [T;\Lambda]$. Thus $\omega \in [T;\Lambda] \subset G$ and G is open in the original sense.

Since $\mathcal{N} = \{[T;\Lambda]\} \cup \{\emptyset\}$ is a countable basis of topology, our metric space (Ω, d) thus obtained is separable. It is also complete. To see this, let $\{\omega^{(k)}\}_{k=1}^{\infty}$ be a Cauchy sequence with respect to the metric d. Then we can choose an increasing sequence $N_n \nearrow \infty$ such that $d(\omega^{(k)},\omega^{(\ell)}) \leq e^{-n}$ for all $k,\ell \geq N_n$. In other words, we have $z_j(\omega^{(k)}) = z_j(\omega^{(\ell)})$, $1 \leq j \leq n$ for all $k,\ell \geq N_n$. If we let $\omega^{(\infty)} = \{\phi\} \cup \bigcup_{n=1}^{\infty} z_n(\omega^{(N_n)})$, then $\omega^{(\infty)}$ is a tree, and it is clear that $d(\omega^{(k)},\omega^{(\infty)}) \to 0$ as $k \to \infty$.

Noting that $\mathcal{N}$ is also a π-system, we can apply Theorem 2.2 of Billingsley[(1)], to obtain the following criterion for the weak convergence of probability measures on Ω.

Proposition 2. Let P and P_n, $n = 1,2,\dots$ be probability measures on Ω. If $P_n([T;\Lambda]) \to P([T;\Lambda])$ for each $[T;\Lambda] \in \mathcal{N}$, then one has $P_n \Rightarrow P$.

As an immediate corollary of this proposition, we see that if P_n and P are the probability measures for Galton-Watson processes with offspring distributions Π_n and Π respectively, and if $\Pi_n \Rightarrow \Pi$ on N, the $P_n \Rightarrow P$ on Ω.

Acknowledgement The authors are grateful to Professor M. Kanda for his continual encouragement. The second author is also grateful to Professors K. Fukuyama and Y. Higuchi for valuable discussion and for giving him opportunity to lecture on this subject.

References

1. Billingsley, P. (1999). *Convergence of Probability Measures.*, 2nd edn. John Wiley, New York.
2. Neveu, J. (1986). Arbres et processus de Galton-Watson. *Ann. Inst. H. Poincaré, Probabilités et Statistique* **22**, 199–207.
3. Otter, R. (1949). The Multiplicative Process. *Ann. Math. Statist.* **20**, 206–224.