ADVANCES IN MATHEMATICS 10, 401–416 (1973)

Random Variables, Trees, and Branching Random Walks*

A. JOFFE

Département de Mathématiques, Université de Montréal, Montréal, Québec, Canada

AND

A. R. MONCAYO

CIMASS, Universidad Nacional Autónoma de México, México 20

1. INTRODUCTION

The study of sums of independent random variables defined on a tree has not been systematically treated in the literature, except for the case of the random tree generated by a Galton–Watson process. Harris [2, p. 75] conjectured that if a supercritical Galton–Watson process performs a random walk, then conditional on nonextinction, the proportion of particles of the nth generation lying in the interval $(-\infty, n\mu + \sigma x n^{1/2})$ converges in probability to $\phi(x)$, the normalized normal distribution. Here, μ and σ^2 denote, respectively, the mean and the variance of the random walk. Ney proved, for a slightly different model, that this is indeed the case. In [8] he dealt with Kolmogorov [7] binary cascade problem, and in [9] he extended his results substantially. Kharlamov [6] has used similar techniques to obtain limit theorems in the theory of multitype age dependent branching processes.

Joffe conjectured in [3] that the aforementioned convergence holds with probability one. In [4] we proved that conjecture for the binary tree. The present paper is the beginning of a systematic study of the sums of independent identically distributed random variables taken along each branch of a given tree and among other things proofs are given of results announced in [5].

In Section 2, we introduce the notation to be used in the description of trees. A crucial role is played by the function $\alpha(n, k)$ which counts the

* This research was supported by CIMASS (UNAM) and the National Research Council of Canada.

number of couples of paths of length n, having in common some subpath of precise length k. In Section 3, we describe the model. Section 4 contains nine lemmas of a computational nature from which the theorems of Section 5 immediately follow. In Section 6 we apply one of those theorems to branching random walks to prove the conjecture stated in [3], assuming for the sake of simplicity that nonextinction occurs with probability one.

The random characteristic function of the random point distribution is used systematically; this leads to substantial simplifications on the second order methods used in [6, 8, and 9]. The decomposition in (3.9) is essential to our method. Theorems 1 and 3 of Section 5 give very general conditions for the convergence with probability one of the random point distribution to the normal distribution. These conditions may be difficult to apply; for this reason we derive by the same techniques the more specialized Theorems 2 and 4. Theorems 3 and 4 require fewer assumptions on the tree and more assumptions on the tail of the random variables than Theorems 1 and 2. It would be interesting to have an example of a tree for which the conclusion of Theorem 3 holds for random variables having moments of sufficiently high order but fails for some random variable having first and second moments only. Theorem 5 gives a necessary and sufficient condition on the tree for mean square convergence. Theorem 5 indicates limitations to the extension of Theorems 1 and 3. It would also be interesting to find an example of a tree for which mean square convergence holds but not almost everywhere convergence.

In Section 6 it is shown that for the random tree of a Galton–Watson process the $\alpha(n, k)$, properly normalized, form a martingale; this easy observation seems to have many applications which we hope to investigate in the future.

2. Notations and Definitions

For the purpose of this paper a tree $(Y, \leqslant)$ is a set Y with a partial order $\leqslant$ satisfying the following conditions:

(i) Each nonempty subset of Y has an infimum in Y. Let $0 = \inf Y$.

(ii) For each $\alpha \in Y$, the set $P_\alpha = \{\beta \mid 0 < \beta \leqslant \alpha\}$ is finite and totally ordered by $\leqslant$.

(iii) For each positive integer n the set $Y_n = \{\alpha \mid |P_\alpha| = n\}$ is finite and $|Y_n| \geq 1$. (For any set A, $|A|$ denotes the cardinality of A.). Let $\xi_n = |Y_n|$.

Elements of Y_n will be referred to as paths of length n or members of the nth generation. Define $\alpha(n, k)$ by

$$\alpha(n, k) = |\{(\tau, \tau') \in Y_n \times Y_n\,;\, \tau \wedge \tau' \in Y_k\}|, \tag{2.1}$$

where as usual $\wedge$ denotes the inf. Clearly we have

$$\sum_{k=0}^{n} \alpha(n, k) = \xi_n^2. \tag{2.2}$$

For any $\tau \in Y$, we define the tree $(Y^\tau, \leqslant)$ where

$$Y^\tau = \{\alpha \mid \tau \leqslant \alpha\}, \tag{2.3}$$

clearly Y^τ inherits the structure of a tree where τ plays the role of 0. We denote by $\alpha_\tau(n, k)$ the corresponding quantities to (2.1) for Y^τ, similarly let

$$Y_n^\tau = Y^\tau \cap Y_n \qquad \text{and} \qquad \xi_\tau(n) = |Y_n^\tau|. \tag{2.4}$$

We have the following relations, where k denotes the length of τ:

$$\sum_{l=0}^{n-k} \alpha_\tau(n-k, l) = \xi_\tau^2(n), \tag{2.5}$$

$$\sum_{\tau \in Y_k} \alpha_\tau(n-k, l) = \alpha(n, k+l). \tag{2.6}$$

Let $p_{n,k} = (\alpha(n, k)/\xi^2(n))$. We say that the tree Y is regular if:

$$\lim_{n\to\infty} p_{n,k} = p_k \qquad \text{with} \quad \sum_{k=0}^{\infty} p_k = 1. \tag{2.7}$$

Let g be a nonnegative nondecreasing function defined on the integers:

(a) We say that the tree Y is g-regular if it is regular and if:

$$\lim_{n\to\infty} \sum_{k=0}^{n} g(k)\, p(n, k) = \sum_{k=0}^{\infty} g(k)\, p_k < \infty. \tag{2.8}$$

(b) We say that the tree Y is weakly g-regular it

$$\sup_n \sum_{k=0}^{n} g(k)\, p(n, k) < \infty. \tag{2.9}$$

A g-regular tree is obviously weakly g-regular.

For any integer $m \geqslant 2$, we define the m-adic tree to be the one for which all $\xi_\tau(k+1) = m$ where k denotes the length of τ. For the m-adic tree we have

$$\alpha(n, k) = (m-1)\, m^{2n-k-1} \qquad \text{if} \quad 0 \leqslant k < n, \tag{2.10.1}$$

$$\alpha(n, n) = m^n, \tag{2.10.2}$$

$$p(n, k) = p_k = (m-1)\, m^{-k-1} \qquad \text{if} \quad k < n, \tag{2.10.3}$$

$$p(n, n) = m^{-n}. \tag{2.10.4}$$

In particular the m-adic tree is, therefore, g-regular for any function $g(k) \leqslant Mc^k$ with $c < m$.

3. Description of the model

We consider a family of independent identically distributed random variables X_τ indexed by $Y - 0$. To simplify notations and statement of the theorems, we assume X's to have mean 0 and variance 1. We denote by $\Phi(t)$ the characteristic function of X. From the central limit theorem we know that

$$\lim_{n\to\infty} \varphi^n(tn^{-1/2}) = e^{-t^2/2}.$$

For each fixed t and n sufficiently large, $\varphi(tn^{-1/2})$ will be bounded away from 0 and this will be always understood even if not mentioned later.

At each path τ of length n we associate the random variables

$$S_\tau = \sum_{0<\beta\leqslant\tau} X_\beta, \tag{3.2}$$

$$S_\tau^* = S_\tau n^{-1/2}.$$

The finite set $\{S_\tau^*\}_{\tau\in Y_n}$ defines a random point distribution on the real line. By assigning at each point the weight $(1/\xi_n)$, we define a random probability measure

$$\mu_n(B) = \frac{1}{\xi_n} \sum_{\tau\in Y_n} V_B(S_\tau^*), \tag{3.3}$$

where V_B denotes the characteristic function of the Borel set B.

The random characteristic function $\Psi_n(t)$ of μ_n is given by

$$\Psi_n(t) = \frac{1}{\xi_n} \sum_{\tau \in Y_n} e^{itS_\tau^*}. \tag{3.4}$$

Let $\mathscr{F}_k$ be the σ-field generated by the X_τ, $\tau \in \bigcup_{j=1}^k Y_j$. We define the following useful quantities where $k \leqslant n$:

$$\sigma_n^2(t) = E \mid \Psi_n(t) - \varphi^n(tn^{-1/2})|^2, \tag{3.5}$$

$$A_{n,k}(t) = \Psi_n(t) - E(\Psi_n(t) \mid \mathscr{F}_k), \tag{3.6}$$

$$B_{n,k}(t) = E(\Psi_n \mid \mathscr{F}_k) - \varphi^{n-k}(tn^{-1/2}), \tag{3.7}$$

$$C_{n,k}(t) = \varphi^{n-k}(tn^{-1/2}) - e^{-t^2/2}. \tag{3.8}$$

We want to study the convergence of $\Psi_n(t)$ to the characteristic function $e^{-t^2/2}$ of the normal distribution.

Clearly we have

$$| \Psi_n(t) - e^{-t^2/2} | \leqslant | A_{n,k} | + | B_{n,k} | + | C_{n,k} |. \tag{3.9}$$

4. Some Lemmas

The following lemmas hold.

LEMMA 1. $E\Psi_n(t) = \varphi^n(t/n^{1/2})$.

LEMMA 2. $\sigma_n^2(t) = | \varphi(t/n^{1/2})|^{2n} \sum_{k=0}^n p(n, k)(| \varphi(t/n^{1/2})|^{-2k} - 1)$.

LEMMA 3. $E(\Psi_n(t)| \mathscr{F}_k) = 1/\xi_n \sum_{\tau\in Y_k} e^{itn^{-1/2}S_\tau}\xi_\tau(n)\, \varphi^{n-k}(tn^{-1/2})$.

LEMMA 4. $E \mid A_{n,k}^2 \mid = 1/\xi_n^2 \mid \varphi(tn^{-1/2})|^{2(n-k)} \sum_{l=0}^{n-k} \alpha(n, k + l) \times (| \varphi(tn^{-1/2})|^{-2l} - 1)$.

Proofs. Lemma 1 is obvious. To prove Lemma 2, note that

$$\sigma_n^2(t) = E\Psi_n\overline{\Psi}_n - | \varphi(tn^{-1/2})|^{2n},$$

then

$$E\Psi_n\overline{\Psi}_n = \frac{1}{\xi_n^2} \sum_{(\tau,\tau')\in Y_n\times Y_n} Ee^{it(S_\tau^* - S_{\tau'}^*)}. \tag{4.1}$$

Observe that if k denotes the length of $\tau \wedge \tau'$ the first kX's in $S_\tau{}^* - S_{\tau'}^*$ cancel, and we are left with the difference of two sums of $n-k$ independent random variables. Each term in the sum of (4.1) is going to contribute to $|\varphi(tn^{-1/2})|^{2(n-k)}$. Using the definition of $\alpha(n, k)$ in (2.1) Lemma 2 is easily proved.

To prove Lemmas 3 and 4, note that we can express $\Psi_n(t)$ as:

$$\begin{aligned}\Psi_n(t) &= \frac{1}{\xi_n} \sum_{\tau \in Y_k} e^{itn^{-1/2}S_\tau} \sum_{\tau' \in Y_n^\tau} e^{itn^{-1/2}(S_{\tau'} - S_\tau)} \\ &= \frac{1}{\xi_n} \sum_{\tau \in Y_k} Y_\tau ,\end{aligned} \tag{4.2}$$

where

$$Y_\tau = e^{itn^{-1/2}S_\tau} T_\tau \qquad \text{and} \qquad T_\tau = \sum_{\tau' \in Y_n^\tau} e^{itn^{-1/2}(S_{\tau'} - S_\tau)}. \tag{4.3}$$

Observe that $e^{itn^{-1/2}S_\tau}$ and T_τ are independent random variables and that the T_τ are jointly independent random variables ($\tau \in Y_k$). It follows that the Y_τ are conditionally $\mathscr{F}_k$-independent. Using this remark, Lemma 3 is proved by taking the conditional expectation on $\mathscr{F}_k$ in (4.3) and observing that $\xi_\tau(n)$ is the cardinality of $Y_n{}^\tau$.

To prove Lemma 4 is now quite straightforward. Using (4.2) and proceeding as in the proof of Lemma 2 we have

$$|A_{n,k}| = \frac{1}{\xi_n} \left| \sum_{\tau \in Y_k} e^{itn^{-1/2}S_\tau}(T_\tau - \xi_\tau(n)\, \varphi^{n-k}(tn^{-1/2}) \right|$$

and

$$E\,|A_{n,k}|^2 = \frac{1}{\xi_n^2} \sum_{\tau \in Y_k} \sum_{\tau' \in Y_k} E e^{itn^{-1/2}(S_\tau - S_{\tau'})}(T_\tau - ET_\tau)(\bar{T}_{\tau'} - E\bar{T}_{\tau'}). \tag{4.4}$$

From the previous remark, if we take first the conditional expectation on $\mathscr{F}_k$ in (4.4), all the terms with $\tau \neq \tau'$ vanish in the sum, and we are left with

$$E\,|A_{n,k}|^2 = \frac{1}{\xi_n^2} \sum_{\tau \in Y_k} E\,|T - ET_\tau|^2. \tag{4.5}$$

To compute $E\,|T_\tau - ET_\tau|^2$ we proceed as in the proof of Lemma 2 but working this time with the family $Y_n{}^\tau$. We obtain

$$E\,|T_\tau - ET_\tau|^2 = |\varphi(tn^{-1/2})|^{2(n-k)} \sum_{l=0}^{n-k} \alpha_\tau(n-k, l)(|\varphi(tn^{-1/2})|^{-2l} - 1). \tag{4.6}$$

Substituting (4.6) in (4.5) and using (4.4) we have proved Lemma 4.

LEMMA 5. *Let* $c_k(n) = \sum_{j=k}^{n} (j/n)(\alpha(n,j)/\xi_n^2)$. *Let* $k_n \uparrow \infty$ *in such a way that* $\sum_{n=1}^{\infty} c_{k_n}(n) < \infty$ *then* $A_{n,k_n}(t)$ *converges to zero with probability one.*

Proof. Using the expansion $\varphi(tn^{-1/2}) = 1 - (t^2/2n) + o(1/n)$ we obtain from Lemma 4

$$E\,|\,A_{n,k}\,|^2 = \frac{1}{\xi_n^2}\,|\,\varphi(tn^{-1/2})|^{2(n-k)} \sum_{l=0}^{n-k} \frac{l}{n}\,\alpha(n, k+l)(t^2 + o(1/n))$$

$$= \frac{1}{\xi_n^2}\,|\,\varphi(tn^{-1/2})|^{2(n-k)} \sum_{j=k}^{n} \frac{j-k}{n}\,\alpha(n, j)(t^2 + o(1/n)).$$

The lemma follows by using the Borel–Cantelli lemma.

LEMMA 6. *If the tree is weakly g-regular with* $g(k) = k^{1+\alpha}$, $\alpha > 0$, *then Lemma* 5 *holds if we take* $k_n \geqslant \log^{\beta} n$ *with* $\beta > (1/\alpha)$.

Proof. From $\sum_{j=1}^{\infty} j^{1+\alpha} p(n,j) \geqslant k^{\alpha} \sum_{j=k}^{\infty} jp(n,j)$ we obtain that there is a constant C such that

$$c_k(n) \leqslant \frac{C}{nk^{\alpha}}\,;$$

the lemma follows from the convergence of the series $\sum 1/n \log^{\gamma} n$ with $\gamma > 1$.

LEMMA 7. *Let* $d_{\epsilon}(n,k) = P(|\,X_1 + \cdots + X_k\,| > n^{1/2}\epsilon)$.
For any sequence k_n *increasing to infinity such that* $\sum_n d_{\epsilon}(n, k_n) < \infty$, *for any* $\epsilon > 0$, *we have* $B_{n,k_n} \to 0$ *almost surely.*

Proof. From (3.7) and Lemma 3 we have

$$B_{n,k} = \varphi^{n-k}(tn^{-1/2}) \frac{1}{\xi_n} \sum_{\tau \in Y_k} (e^{itn^{-1/2} S_\tau} - 1)\,\xi_\tau(n). \tag{4.7}$$

To show that B_{n,k_n} goes to zero it is clearly sufficient to prove that

$$\eta_{n,k}(t) = \frac{1}{\xi_n} \sum_{\tau \in Y_k} e^{itn^{-1/2} S_\tau} \xi_\tau(n) \text{ goes to one.}$$

But $\eta_{n,k}(t)$ is the characteristic function of the random measure $D_{n,k}(B) = 1/\xi_n \sum_{\tau \in Y_k} V_B(S_\tau n^{-1/2})\,\xi_\tau(n)$, where V_B is the characteristic function of the Borel set B. The lemma will be established if we can prove that, with probability one, the sequence D_{n,k_n} converges weakly to δ_0 the probability measure which assigns probability one at zero. Next observe that

$$ED_{n,k}(B) = P(X_1 + \cdots + X_k/n^{1/2} \in B). \tag{4.8}$$

Let $B_\epsilon^+ = [\epsilon, \infty)$ and $B_\epsilon^- = (-\infty, -\epsilon]$. We have

$$P(D_{n,k}(B_\epsilon^+) > \alpha) \leqslant \frac{ED_{n,k}(B_\epsilon^+)}{\alpha} \leqslant \frac{d(n, k)}{\alpha}$$

and

$$P(D_{n,k}(B_\epsilon^-) \leqslant \alpha) \leqslant \frac{d(n, k)}{\alpha}.$$

It follows from the Borel–Cantelli lemma that with probability one $D_{n,k_n}(B_\epsilon^+) \to 0$ and $D_{n,k_n}(B_\epsilon^-) \to 0$, the lemma follows easily from elementary properties of weak convergence.

COROLLARY. *If $EX^2 \log^\gamma | X | < \infty$ with $\gamma > 1$ and if $k_n \leqslant C \log n$ where $0 < \beta < (\gamma - 1)/3$ then B_{n,k_n} goes to zero almost surely.*

Proof. Is easily established from the lemma and the following computation

$$\begin{aligned} d(n, k) &= P(| X_1 + \cdots + X_k | > \epsilon n^{1/2}) \leqslant kP(| X | > (\epsilon n^{1/2}/k)) \\ &\leqslant Ck^3[\epsilon^2 n \log^\gamma(\epsilon n^{1/2}/k)]^{-1} \end{aligned}$$

as soon as $\epsilon n^{1/2}/k$ is larger than one.

Remark. The expression in (4.8) does not involve the structure of the tree, so we are losing here information by using the lemma of Borel–Cantelli. This is why we have to require more assumptions on the moments of X in order to use the lemma. The next lemma will avoid this argument at the price of some assumptions on the tree.

LEMMA 8. *If Y satisfies the following two conditions* (i) *Y is weakly g-regular for $g(k) = k$;*

(ii) $(\xi_1 + \cdots + \xi_{k_n}/n) \to 0$, *for some sequence $k_n \uparrow \infty$. Then $B_{n,k_n} \to 0$ almost surely.*

Proof. From (3.7) and the inequality $|e^{it}-1| \leqslant |t|$, we obtain

$$|B_{n,k}| \leqslant \frac{1}{n^{1/2}} \frac{1}{\xi_n} \sum_{\tau \in Y_k} \xi_\tau(n) \, |S_\tau|. \tag{4.9}$$

We majorize $|S_\tau|$ by $\sum_{\alpha \leqslant \tau} |X_\alpha|$, regrouping all the terms in X_α we easily obtain from (4.9)

$$|B_{n,k}| \leqslant \frac{1}{n^{1/2}} \frac{1}{\xi_n} \sum_{\tau \in Y_j, j \leqslant k} \xi_\tau(n) \, |X_\tau|.$$

Using the Cauchy–Schwartz's inequality we obtain

$$|B_{n,k}|^2 \leqslant \frac{1}{n\xi_n^2} \Big[\sum_{j \leqslant k, \tau \in Y_j} \xi^2(n) \Big] \Big[\sum_{\tau \in Y_j, j \leqslant k} |X_\tau^2| \Big],$$

which can be expressed as

$$|B_{n,k}|^2 = L_{n,k} \cdot M_{n,k} \cdot N_{n,k}, \tag{4.10}$$

where $L_{n,k} = (\xi_1 + \cdots + \xi_k)/n$, $M_{n,k} = \sum_{\tau \in Y_j, j \leqslant k} \xi_\tau^2(n)/\xi_n^2$ and

$$N_{n,k} = \frac{\sum_{\tau \in Y_j, j \leqslant k} X_\tau^2}{\xi_1 + \cdots + \xi_k}.$$

Note that by (2.3) and (2.4) we have

$$\begin{aligned} \sum_{\tau \in Y_j, j \leqslant k} \xi_\tau^2 &= \sum_{j=1}^{k} \sum_{l=j}^{n} \alpha(n, l) = \sum_{l=1}^{k} l\alpha(n, l) + k \sum_{l=k+1}^{n} \alpha(n, l) \\ &\leqslant \sum_{l=1}^{n} l\alpha(n, l). \end{aligned}$$

The previous inequality and the assumptions of the lemma imply that M_{n,k_n} is bounded; so is N_{n,k_n} which by the strong law of large numbers converge almost surely to one. Hence, the lemma results from the decomposition in (4.10).

LEMMA 9. *If $\Psi_n(t) \to e^{-t^2/2}$ almost surely for each t then with probability one μ_n converges weakly to the normal distribution.*

Proof. The proof is quite standard, and a brief sketch is given here for the sake of completeness. Let $A = \{(t, \omega) \mid \Psi_n(t, \omega) \nrightarrow e^{-t^2/2}\}$ and

A_t, A_ω be the t and ω sections of A. Let μ be the Lebesgue measure on the real line R. By Fubini theorem

$$0 = \int_R \int_\Omega V_{A_t}\, dP\, d\mu = \int_\Omega \int_R V_{A_\omega}\, d\mu\, dP,$$

where V_{A_t} and V_{A_ω} are the characteristic functions of the sets A_t and A_ω, respectively. Hence, $\mu(A_\omega) = 0$. The lemma results then from elementary property of weak convergence of measure and the Fourier inversion formula.

5. Theorems for Random Walks on Trees

Now we have the machinery to obtain immediately the following theorems.

Theorem 1. *Let Y be a weakly g-regular tree with $g(k) = k$ and $c_k(n)$ be as in Lemma 5 of Section 4. Assume that there is an increasing sequence $k_n \uparrow \infty$ such that*

(a) $\sum c_{k_n}(n) < \infty$,

(b) $\dfrac{k_n}{n} \to 0$,

(c) $\dfrac{\xi_1 + \cdots + \xi_{k_n}}{n} \to 0$,

then μ_n converges weakly to the normal distribution with probability one.

Theorem 2. *If Y is weakly g-regular with $g(k) = k^{1+\alpha}$, $\alpha > 0$ and if $(\xi_1 \cdots \xi_{k_n}/n) \to 0$ with $k_n \geqslant \log^\beta n$ for a β, $\beta > (1/\alpha)$ then μ_n converges weakly to the normal distribution with probability one.*

Theorem 3. *Let $c_k(n)$ be as in Lemma 5 and $d_\epsilon(n, k)$ be as in Lemma 7 of Section 4. If there is a sequence k_n such that*

(a) $\sum c_{k_n}(n) < \infty$,

(b) $\sum d_\epsilon(n, k_n) < \infty$ *for all* $\epsilon > 0$,

(c) $\dfrac{k_n}{n} \to 0$,

then μ_n converges weakly to the normal distribution with probability one.

THEOREM 4. *Let* $EX^2 \log^\gamma |X| < \infty$ *with* $\gamma > 1$, *and assume the tree to be weakly g-regular with* $g(k) = k^{1+\alpha}$ *and* $\alpha > 3/\gamma - 1$, *then* μ_n *converges weakly to the normal distribution with probability one.*

THEOREM 5. *A necessary and sufficient condition for the convergence in mean square of* $\Psi_n(t)$ *to* $e^{-t^2/2}$ *is that as n goes to infinity*

$$\sum_{k=0}^{n} k \frac{\alpha(n, k)}{\xi_n^{\,2}} = o(n).$$

Proofs. To prove the first four theorems, we show first that for each t, $\Psi_n(t)$ converges almost surely to $e^{-t^2/2}$ and then we conclude by Lemma 9 of Section 4. Theorem 1 is then a consequence of Lemmas 5 and 8 of Section 4 and of formula (3.9). Theorem 2 follows from Theorem 1 and Lemma 6 of Section 4. Theorem 3 is immediate from the Lemmas 5 and 7 of Section 4, while Theorem 4 is an easy consequence of Lemma 6 and the corollary of Lemma 7 of Section 4.

To prove Theorem 5 use Lemma 2 of Section 4 and the expansion $\varphi(tn^{-1/2}) = 1 - (t/2n) + \epsilon(n)$ with $\epsilon(n) \to 0$ as n goes to infinity.

6. GALTON–WATSON TREE WITH MEAN LARGER THAN ONE AND FINITE VARIANCE

For a detailed reference of Galton–Watson process see Harris [2] or the forthcoming book of Athreya and Ney [1]. We can view a Galton–Watson process as a random tree in which the number of branches $\xi_\tau(n+1)$, where n is the length of τ, are independent identically distributed random variables which are defined on a probability space $(\Omega, \mathscr{F}, P)$ and take on nonnegative integer values; we will assume here that the mean $m = E\xi_\tau(n+1)$ is finite and larger than one and the variance σ^2 is finite. In order to simplify matter and to avoid to have to condition on nonextinction we will also assume that $P(\xi_\tau(n+1) = 0) = 0$. We call $\mathscr{G}_n$ the σ-field generated by $\xi_\tau(j+1)$ for $j = 1, 2,..., n-1$, where j is the length of τ.

In this case $\alpha(n, k)$ for $k = 1, 2,..., n$, $\xi_\tau(n)$, $\xi(n)$ a.s. o are random variables and from formula (2.1) we easily established the following relations, where $V_{n,k}$ is the characteristic function of the set appearing in (2.1):

$$\text{for } k \leqslant n-1, \quad \alpha(n+1, k) = \sum_{\tau}\sum_{\tau'} \xi_\tau(n+1)\, \xi_{\tau'}(n+1)\, V_{n,k}(\tau, \tau'), \tag{6.1}$$

$$\text{for } k = n, \qquad \alpha(n+1, n) = \sum_{\tau \in Y_n} (\xi_\tau^2(n+1) - \xi_\tau(n+1)), \tag{6.2}$$

and

$$\text{for } k = n+1, \quad \alpha(n+1, n+1) = \xi_{n+1}. \tag{6.3}$$

Taking conditional expectations with respect to $\mathscr{G}_n$ we obtain the following.

LEMMA 1. *For* $k \leqslant n-1$, $E\{\alpha(n+1, k) \mid \mathscr{G}_n\} = m^2\alpha(n, k)$;
for $k = n$, $E\{\alpha(n+1, n) \mid \mathscr{G}_n\} = (\sigma^2 + m^2 - m)\, \alpha(n, n)$;
for $k = n+1$, $E\{\alpha(n+1, n+1) \mid \mathscr{G}_n\} = m\xi_n$ a.s. (ω),

from which the next lemma follows after taking expectations.

LEMMA 2. *For* $k \leqslant n-2$, $E\{\alpha(n, k)\} = m^{2n-k-2}(\sigma^2 + m^2 - m)$;
for $k = n-1$, $E\{\alpha(n, n-1)\} = (\sigma^2 + m^2 - m)\, m^{n-1}$;
for $k = n$, $E\{\alpha(n, n)\} = m^n$.

In particular for k fixed and $n \geqslant k+2$, $(\alpha(n, k)/m^{2n})$ *is a martingale with respect to* $\mathscr{G}_n$.

Proof. By lemma 1: if $k \leqslant n-2$, then

$$\begin{aligned} E\{\alpha(n, k)\} &= m^{2(n-k-2)} E\{\alpha(k+2, k)\} = m^{2(n-k-2)} m^2 E\{\alpha(k+1, k)\} \\ &= m^{2(n-k-2)} m^2 (\sigma^2 + m^2 - m)\, E\{\alpha(k, k)\} \\ &= m^{2(n-k-2)} m^2 (\sigma^2 + m^2 - m)\, m^k, \text{ etc.} \end{aligned}$$

Let $\beta_n = 1/m^{2n} \sum_{k=0}^{n} g(k)\, \alpha(n, k)$ where g is a nondecreasing and nonnegative function and $g(k) \leqslant Mc^k$ for $k = 0, 1, 2, \ldots$ with $c < m$, then we get the following lemma.

LEMMA 3. $\{\beta_n, \mathscr{G}_n\}_{n=1,2,\ldots}$ *is submartingale, and since* $\sup_n E(\beta_n) < \infty$ *it follows by the submartingale convergence theorem that* $\beta_n \to \beta$ a.s. (ω) *where* β *is an* a.s. (ω) *finite random variable.*

Proof. By Lemma 1,

$$E(\beta_n \mid \mathscr{G}_{n-1})$$

$$= \frac{1}{m^{2n}} \left\{ \sum_{k=0}^{n-2} g(k)\, m^2 \alpha(n-1, k) + g(n-1)(\sigma^2 + m^2 - m)\, \xi_{n-1} + g(n)\, m\xi_{n-1} \right\}$$

$$= \frac{1}{m^{2n}} \left\{ \sum_{k=0}^{n-1} g(k)\, m^2 \alpha(n-1, k) + g(n-1)(\sigma^2 - m)\, \xi_{n-1} + g(n)\, m\xi_{n-1} \right\}$$

$\geqslant \beta_{n-1}$ a.s. (ω) since g is nondecreasing.

Also by Lemma 2

$$E\beta_n = \frac{1}{m^{2n}} \sum_{k=0}^{n-2} g(k)\, m^{2n-k-2}(\sigma^2 + m^2 - m) + \frac{m^{n-1}(\sigma^2 + m^2 - m)}{m^{2n}} g(n-1) + g(n) \frac{m^n}{m^{2n}},$$

and by hypothesis we obtain

$$E\beta_n \leqslant M \sum_{k=0}^{n-2} c^k m^{-k-2}(\sigma^2 + m^2 - m) + Mc^{n-1}(\sigma^2 + m^2 - m)\, m^{-n-1} + Mc^n m^{-n}$$
$$\leqslant M(n-1)\, m^{-2}(\sigma^2 + m^2 - m) + M(\sigma^2 + m^2 - m) + M < \infty,$$

from which Lemma 3 follows.

Galton–Watson Random Walk. Let $\{X_{n,m}\}_{n,m=1,2,3\ldots}$ be a sequence of independent identically distributed random variables, defined on a probability space $(\Omega', \mathscr{F}', P')$, having mean zero and variance one. $(\Omega, \mathscr{G}, P)$ is the probability space on which the previous random tree is defined and let $(\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{P})$ be the product probability space of Ω and Ω'. Next we define the random variables $\{X_{\tau_n}\}$ defined on $(\tilde{\Omega}, \tilde{\mathscr{F}}, \tilde{P})$ by $X_\tau((\omega, \omega') = X_{n,\tau(\omega)}(\omega')$, which we call a Galton–Watson random walk and which describes random walk of particles reproducing according to a Galton–Watson process.

Theorem. *For a Galton–Watson random walk we have that* $\{\mu_n\}$ *converges weakly, with probability one, to the standard normal distribution.*

Proof. It is enough to check that the conditions of Theorem 2 of Section 5, namely:

(1). $\sup_n \sum_{k=0}^{n} k^{1+\alpha}(\alpha(n, k)/\xi_n^2) < \infty$,

(2). $(1/n)(\xi_1 + \xi_2 + \cdots + \xi_{k_n}) \to 0$, as n goes to infinity where $k_n = [c(\log n)^\beta] + 1$ with $(1/\alpha) < \beta < 1$ are satisfied a.s. (ω). (1). follows by Lemma 3 by considering $g(k) = k^{1+\alpha}$ because then

$$\sum_{k=0}^{n} k^{1+\alpha} \frac{\alpha(n, k)}{\xi_n^2} = \frac{1}{\xi_n^2} \beta_n < \infty \text{ a.s. } (\omega).$$

(2). follows by considering the inequality

$$\frac{\xi_1 + \xi_2 + \cdots + \xi_{k_n}}{n} \leqslant (k_n m^{k_n}/n)(\xi_{k_n}/m^{k_n}),$$

$\xi_{k_n} m^{-k_n} \to W$ a.s. (ω), where W is a.s. (ω) a finite random variable (see [2, p. 13]) while $k_n m^{k_n}/n \to 0$ as $n \to \infty$, since β can be chosen as small as we please. This proves the theorem.

7. Remarks

1. The method used to prove the theorems of Section 5 is very different that the one used in [4] for the binary tree. In [4], we obtained, to begin with, a weak version of Theorem 5 from which it was possible to conclude that, for $\alpha > 0$, $\Psi_{n^{1+\alpha}} \to e^{-t^2/2}$, as n goes to infinity, almost surely. Then we estimated $| \Psi_{n+p} - \Psi_n |$ for $1 \leqslant p \leqslant n^\alpha$. Such an estimation was easy because of the symmetry of the tree. Now a Galton–Watson tree also behaves in a symmetric manner; thus, the method of [4] could lead to the theorem of Section 6. For a general tree, however, the computations become very involved and do not seem to give more general results than our Theorem 1.

2. For the Galton–Watson random walk (which included the m-adic tree) it has been shown in [3] that

$$W_n(t) = \frac{\Psi_n(t)}{m^n \varphi^n(t)} \tag{7.1}$$

is a martingale. Ney, in a private communication, has informed one of us that Watanabe had made the same observation in the context of branching diffusion (see also [1]). It is easy to see that the $W_n(t)$ converge almost surely to $W(t)$. On the other hand, we were unable to show uniform convergence in t or even almost sure continuity of the limit. If this could be done it would lead to a very elegant proof of the theorem of Section 6, based on the simple substitution of t by $tn^{-1/2}$ in (7.1).

3. This paper raises many questions: here we list some of them.

(a) Study of sums of independent random variables on partially ordered sets.

(b) Study of the speed of convergence of the $\Psi_n(t)$.

(c) Here we have studied a version of the central limit theorem. It would be interesting to study versions of the strong law of large numbers and of the law of the iterated logarithm.

(d) Renewal theory on a tree: one formulation could be the following: considering the $\{S_\tau\}$ as a real random function on the tree, the problem is to study the properties of the t-section of that function.

(e) The study in (d) could lead to a systematic treatment of the age-dependent Galton–Watson process.

(f) Generalization of Section 6 to the multitype Galton–Watson process.

We hope in the future to obtain some results leading toward the solution of these problems.

Added in proof. The conjectures mentioned in the introduction were proved in: A. J. Stam, On a conjecture by Harris. *Z. Wahrscheinlichkeitstheorie Ver. Geb.* **5** (1966), 202–206.

References

1. K. B. Athreya and P. Ney, "Branching Processes," Springer Verlag, New York, 1973, in press.
2. T. E. Harris, "The Theory of Branching Processes," Springer, Berlin/Göttingen/Heidelberg, 1963.
3. A. Joffe, Branching processes which perform a random walk, Symposium on applied stochastic processes, Rochester, New York, August 1971.
4. A. Joffe and A. R. Moncayo, On sums of independent random variables defined on a binary tree, *Bol. Soc. Mat. Mexicana* (1972), to appear.
5. A. Joffe and A. R. Moncayo, Asymptotic theorems for sums of independent random variables defined on a tree, *Bull. Amer. Math. Soc.*, to appear.

6. B. P. Kharlamov, The number of generations in a branching process with an arbitrary set of particle types, *Theor. Probability Appl.* **14** (1969), 432–449.
7. A. N. Kolmogorov, Über das logarithmisch normale Verteilungsgesetz der Dimensionen der Teilchen bei Zerstuckelung, *Dokl. Akad. Nauk SSSR* **31** (1941), 99–101.
8. P. E. Ney, The limit distribution of a binary cascade processes, *J. Math. Anal. Appl.* **10** (1965), 30–36.
9. P. E. Ney, The convergence of a random distribution function associated with branching processes, *J. Math. Anal. Appl.* **12** (1965), 316–327.