Journal of Statistical Physics, Vol. 125, No. 4, November 2006 (© 2006)
DOI: 10.1007/s10955-006-9187-7

Some New Results for McKean's Graphs with Applications to Kac's Equation

Ester Gabetta[1] and Eugenio Regazzini[1]

Received July 5, 2006; accepted July 18, 2006
Published Online: December 2, 2006

The main goal of the present paper is to sharpen some results about the error made when the Wild sums, used to represent the solution of the Kac analog of Boltzmann's equation, are truncated at the n-th stage. More precisely, in Carlen, Carvalho and Gabetta (*J. Funct. Anal.* **220**: 362–387 (2005)), one finds a bound for the above-mentioned error which depends on $(an^{\Lambda+\varepsilon})$. On the one hand, it is shown that Λ, the least negative eigenvalue of the linearized collision operator, is the best possible exponent. On the other hand, ε is an extra strictly positive number and a a positive coefficient which depends on ε too. Thus, it is interesting to check whether ε can be removed from the above bound. According to the aforesaid reference, this problem is studied here by means of the probability distribution of the depth of a leaf in a McKean random tree. In fact, an accurate study of the probability generating function of such a depth leads to conclude that the above bound can be replaced with $(a'n^{\Lambda})$.

KEY WORDS: depth of a leaf, depth of a tree, Kac's equation, McKean binary tree (or graph), rate of convergence of Wild sums, Stirling numbers (of the first kind), Wild convolution, Wild sum.

1. INTRODUCTION

Kac's equation describes the motion of a single molecule in a chaotic bath of like molecules moving on the line.(7,10,12,13) At time $t = 0$ the velocities of the molecules are considered as random quantities with a probability distribution satisfying a few specific conditions. In particular, they are assumed to be identically distributed with a common probability density function f_0 having finite mean energy. According to the Kac model, velocities turn out to be identically distributed

[1] Dipartimento di Matematica "F. Casorati", Università degli Studi di Pavia, Via Ferrata 1, I-27100 Pavia, Italy; e-mails: ester.gabetta@unipv.it, eugenio.regazzini@unipv.it.

at each time $t > 0$ and their common density $f(\cdot, t)$ satisfies the so-called *Kac's analog of Boltzmann's equation*

$$\begin{cases} \dfrac{\partial}{\partial t} f(v,t) = \dfrac{1}{2\pi} \displaystyle\int_{\mathbb{R}\times[0,2\pi)} \{ f(v\cos\theta - w\sin\theta, t) f(v\sin\theta + w\cos\theta, t) \\ \qquad\qquad - f(v,t) f(w,t)\}\, dw d\theta \\ f(v, 0^+) := f_0(v) \qquad\qquad\qquad\qquad (t > 0, v \in \mathbb{R}). \end{cases} \tag{1}$$

The right-hand side of the above equation is also well-known as collisional integral. In Ref. 17, McKean shows that, within the class of probability density functions on $\mathbb{R}$, (1) has a unique solution which can be expressed as *Wild's sum.*[22] After introducing the characteristic functions φ_0 and $\varphi(\cdot, t)$ of the initial velocity and of the velocity at time t of each particle, respectively, one can re-write (1) in the following terms

$$\begin{cases} \dfrac{\partial}{\partial t} \varphi(x,t) = \dfrac{1}{2\pi} \displaystyle\int_{[0,2\pi)} \varphi(x\sin\theta, t)\varphi(x\cos\theta, t) d\theta - \varphi(x,t) \\ \varphi(x, 0^+) := \varphi_0(x) \qquad\qquad\qquad\qquad (t > 0, x \in \mathbb{R}). \end{cases} \tag{2}$$

See Ref. 3. Throughout the present paper we frequently refer to (2) which is compatible with arbitrary initial probability measures. The case of initial data given by arbitrary probability distributions has been considered, for example, by Carlen and Lu.[6]

The Wild sum for the solution of (2) is defined by

$$\varphi(x,t) = \sum_{n\geq 1} e^{-t}(1-e^{-t})^{n-1} \hat{q}_n^+(x;\varphi_0) \quad (t \geq 0, x \in \mathbb{R}) \tag{3}$$

where functions $\hat{q}_n^+$s are found by recursion as

$$\hat{q}_n^+(x;\varphi_0) = \frac{1}{n-1} \sum_{j=1}^{n-1} \hat{q}_{n-j}^+(x;\varphi_0) \bullet \hat{q}_j^+(x;\varphi_0) \quad (n = 2, 3, \ldots) \tag{4}$$

with $\hat{q}_1^+ \equiv \varphi_0$ and the proviso that $\bullet$ stands for

$$\varphi_1 \bullet \varphi_2(x) := \frac{1}{2\pi} \int_0^{2\pi} \varphi_1(x\cos\theta)\varphi_2(x\sin\theta)\, d\theta \quad (x \in \mathbb{R}).$$

It is clear that if φ_1 and φ_2 are characteristic functions, then $\varphi_1 \bullet \varphi_2$ is a characteristic function too. The corresponding probability law is called *Wild convolution*. If ℓ_1 and ℓ_2 denote the probability laws associated with φ_1 and φ_2, respectively, then their Wild convolution will be denoted by $\ell_1 \circ \ell_2$.

There exist many studies about the convergence (as t diverges to infinity) of $f(\cdot, t)$ towards the normalized Maxwellian $M(v) = (2\pi)^{-1/2}\exp\{-v^2/2\}$,

$v \in \mathbb{R}$. In particular, in Ref. 5 one determines a bound for $\|q_n^+(f_0) - M\|_{L^1(\mathbb{R})}$ where q_n^+, i.e. any probability density corresponding to $\hat{q}_n^+$, is given by a suitable average over n-fold iterated Wild's convolutions. Moreover, one derives a new bound for $\|f(\cdot, t) - M\|_{L^1(\mathbb{R})}$. The former one is deduced from an analogous bound expressed in terms of a *weighted* χ*-metric* denoted by $|||\cdot|||$ whose definition is quoted in Sec. 6. See, for example, Sec. 14.2 in Rachev.[19] More precisely one has $|||q_n^+(f_0) - M||| \leq an^{\Lambda+\varepsilon}$ where ε is an arbitrary strictly positive number, a is a suitable constant (which depends on ε) and $\Lambda(= -1/4)$, the least negative eigenvalue for the linearized collision operator, turns out to be the best possible exponent. A key role in the study developed in Ref. 5 is played by a probabilistic construction of McKean[17,18] which leads to a very useful expression for q_n^+ through the introduction of certain tree graphs (*McKean graphs* or *McKean binary trees*). As far as some problems tackled in Ref. 5 are concerned, it should be stressed the importance of certain random quantities – i.e. functions defined on the set of all McKean trees – such as the depth of a leaf and the depth of a tree. Specifically, it can be shown that the statement of the above-quoted bounds depends on the structure of the probability distribution of the depth. The McKean representation expresses $\hat{q}_n^+$ as a weighted mean of n-fold "products" (in the sense of $\bullet$) of φ_0:

$$\hat{q}_n^+(x; \varphi_0) = \sum_{\gamma \in G(n)} p_n(\gamma) c_\gamma(x; \varphi_0) \quad (n = 1, 2, \ldots). \tag{5}$$

An explanation of (5) can start from $G(n)$ and γ. Here, $G(n)$ denotes the class of all McKean graphs with n leaves and γ stands for any of these graphs. They can be characterized by the fact that each node has either zero or two "children", a "left child" and a "right child". To illustrate the definition, a few elements of $G(8)$ are visualized in Fig. 1.

As far as the identification of c_γ is concerned, consider a sequence of (scattering) angles $(\theta_j)_{j\geq 1}$ in $[0, 2\pi)$ and, for any leaf l at level i of any γ in $G(n)$, look at the path which connects l and the "root" node , in ascending order. This path consists of i steps: the first one from l to its "parent" node, the second one from the "parent" to the "grandparent" of l, and so on. To such a path associate the product $\pi(l) = \alpha_1 \ldots \alpha_i$ where $\alpha_i = \alpha(\theta_i, \gamma)$ equals $\cos\theta_i$ if l is a "left child", or $\sin\theta_i$ if l is a "right child"; α_{i-1} equals $\cos\theta_{i-1}$ or $\sin\theta_{i-1}$ depending on the "parent" of l is, in turn, a "left child" or not, and so on. To illustrate this construction, consider leaf l_1 in γ_1 of Fig. 1. For such a leaf it turns out that $i = 3$, $\alpha_3 = \sin\theta_3$, $\alpha_2 = \cos\theta_2$, $\alpha_1 = \cos\theta_1$ and $\pi(l_1) = \sin\theta_3\cos\theta_2\cos\theta_1$. After setting $\pi(l) := 1$ if $n = 1$, it is worth recalling that equality

$$\sum_{l \in \gamma} \pi(l)^2 = 1 \tag{6}$$

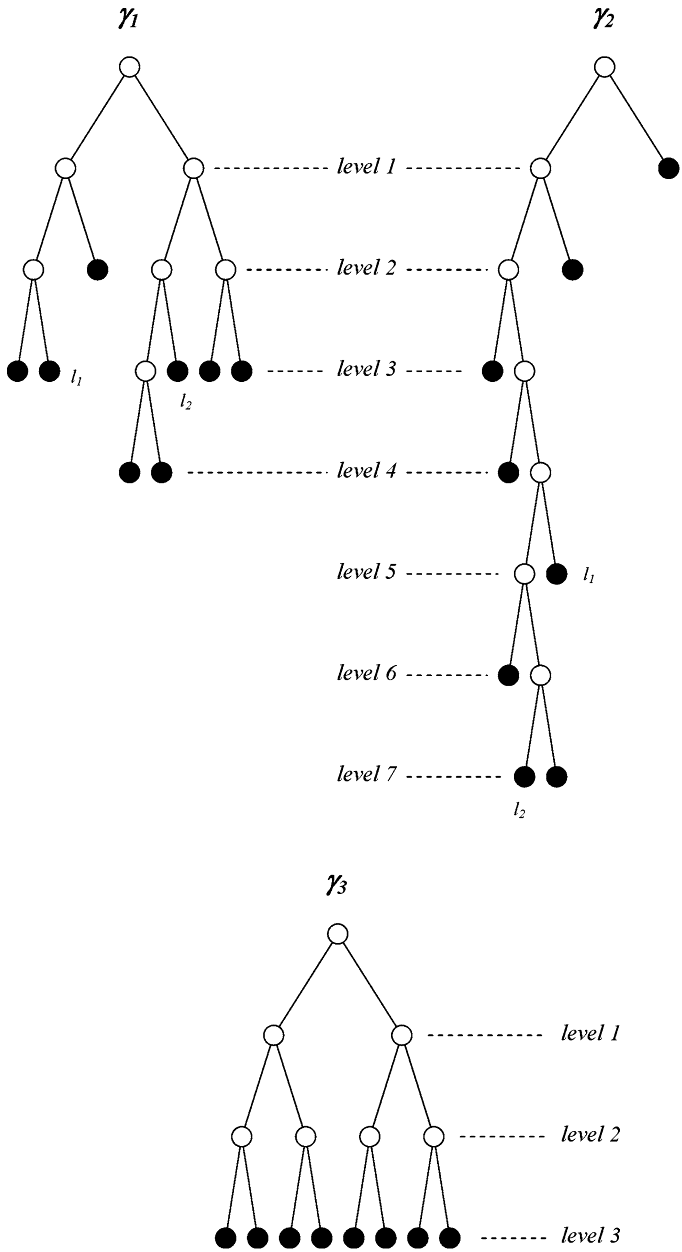

Fig. 1. Shaded (unshaded) circles stand for leaves (nodes).

holds for every γ in $G := \bigcup_{n\geq} G(n)$.[17] After stipulating the above conditions, it can be shown that c_γ can be written as

$$c_\gamma(x;\varphi_0) = \int_{[0,2\pi)^\infty} \left[\prod_{l\in\gamma} \varphi_0(x\pi(l))\right] u(d\theta)$$

for every γ in G, u being the product measure on the Borel σ-algebra $\mathcal{B}^\infty([0, 2\pi))$ of the product space $[0, 2\pi)^\infty$ which makes the coordinates stochastically independent and identically distributed according to the uniform probability on $[0, 2\pi)$. For this kind of construction, due to Kolmogorov, see, for example, Sec. 36 of Ref. 2.

Notice that γ can be split into a "left" subgraph γ_l and a "right" subgraph γ_r by removing the "root" node. Moreover, any element γ of $G(n)$ can be viewed as a tree with leaves labelled, from left to right, with the first n natural numbers. Specifically, the leaves are assumed to be ordered and labelled in the following way. The minimal leaf, designated by 1, is the one characterized by the circumstance that the path joining it to the root contains only "left children". Number 2 is attached to the minimal leaf of the right subgraph of the tree having "root" in the "parent" of leaf 1. The assignment process continues in this way until the completion of the ordering of the leaves of that very same subgraph. At this stage the process is extended in the same way to the subgraph of γ which has root in the "grandparent" of 1, and so on until the "root" of γ is reached. This last circumstance marks the beginning of the extension of the ordering to the right subgraph of γ, i.e. γ_r. In order to illustrate the process just described, note that rearrangement of the leaves of γ_1 in Fig. 1 according to the above ordering leads to assign number 2 to l_1 and number 6 to l_2. As far as γ_2 is concerned, leaves l_1 and l_2 must be labelled with 6 and 4, respectively. At this stage one is in a position to define the *depth of leaf* j in γ – $\delta_j(\gamma)$ in symbols – as the number of generations which separate j from the root of γ. In view of this definition, $\pi(j)$ can be written as $\alpha_1 \dots \alpha_{\delta_j}$. The quantity $\delta_{(1)}(\gamma) := \min\{\delta_1(\gamma), \dots, \delta_n(\gamma)\}$ corresponds to the above-mentioned concept of *depth of tree* γ.

The probability p_n on $G(n)$ induces probability distributions for δ_j and for $\delta_{(1)}$. The former one is determined in Proposition 2 of the present paper in the form

$$p_n\{\delta_j = d\} = \sum_k \frac{1}{(j-1)!(n-j)!} |s(j-1, d-k)| \cdot |s(n-j, k)| \qquad (7)$$

where $d = 1, \dots, n-1$ and $s(n, k)$ denotes a Stirling number of the first kind.[8,9]

The probability distribution of δ_j admits a couple of interesting interpretations. According to the former, the distribution of δ_j turns out to be connected with a well-known urn scheme explained, for istance, in Chapter 8 of Ref. 8.

Suppose that balls are successively drawn one after the other from an urn initially containing m white balls. After each trial the drawn ball is placed back in the urn along with s black balls. Then, if A_j denotes the event of drawning a white ball at the j-th trial, $j = 1, \dots, n$, setting $\theta = m/s$, one gets

$$\text{Prob}(A_j) = \frac{\theta}{\theta + j - 1} \quad (j = 1, \dots, n)$$

and one can prove that

$$\frac{|s(n,k)|\theta^k}{(\theta + n - 1)_n} \quad k = 0, 1, \dots, n \tag{8}$$

coincides with the probability of drawning k balls in n trials. This is the same as saying that the probability of the event $\{Y_1 + \cdots + Y_n = k\}$ is given by (8) when $Y_1, Y_2, \dots$ form a sequence of independent random variables such that

$$\text{Prob}\{Y_j = 1\} = \frac{\theta}{\theta + j - 1} = 1 - \text{Prob}\{Y_j = 0\} \quad (j = 1, 2, \dots).$$

Going back to (7), one notes that the probability distribution of δ_j is the convolution of two distributions of the same type as (8) with $\theta = 1$. This indicates that δ_j can be thought of as a sum of two independent random variables – say $\delta_{j,l}$ and $\delta_{j,r}$, respectively – $\delta_{j,l}$ having distribution (8) with $n = j - 1$ and $\delta_{j,r}$ the same distribution with $(n - j)$ in place of $n, \theta = 1$ in both cases. The latter interpretation starts from pointing out that $|s(n,k)|$ represents the number of permutations of n objects, with k orbits. See, for instance, Ref. 9. With reference to the physical model of interest, in the former interpretation, one can identify the collision with the appearance of a white ball. Then, $\delta_{j,l}$ ($\delta_{j,r}$, respectively) represents the random number of collisions of j with particles labelled with $1, \dots, j - 1$ (particles labelled with $j + 1, \dots, n$, respectively) under the assumption that the collision process is driven by the above-described urn scheme. In the latter interpretation one points out that the distribution of $\delta_{j,l}$ ($\delta_{j,r}$, respectively) is generated by the same mechanism as the one which regulates the number of orbits of the random permutations of $(j - 1)$ ($(n - j)$, respectively) objects when the same probability is attributed to all permutations at issue.

Expression (7) is derived through a new algorithm for evaluating $p_n(\gamma)$ explained in Sec. 2 and through the use of suitable probability generating functions such as

$$V_n(x, \xi) = \sum_{d=0}^{n-1} \xi^j x^d p_n\{\delta_j = d\}$$

which can be written as

$$V_n(x, \xi) = \sum_{j=1}^{n} \binom{x+n-j-1}{x-1} \binom{x+j-2}{x-1} \xi^j.$$

Now, for any c in (0, 1), $V_n(c/2, 1)$ plays a fundamental role in Ref. 5, where it is denoted by $T(n)$, to state the inequality

$$|||q_n^+(\cdot, f_0) - M||| \leq bT(n) \tag{9}$$

with $b = b(f_0)$. Moreover, the following bound for $T(n)$ is there obtained for any $\varepsilon > 0$,

$$T(n) \leq A(\varepsilon) \frac{n^\varepsilon}{n^{1-c}} \quad (c = 1 + \Lambda = 3/4). \tag{10}$$

Looking at the expression of $A(\varepsilon)$ provided in Ref. 4, p. 386, one notes that the above bound for $T(n)$ is not valid for $\varepsilon = 0$. In Proposition 8 of the present paper, by resorting to the above expression of $V_n(x, \xi)$ one gets an *exact* simple formula for $T(n)$, i.e.

$$T(n) = \frac{\Gamma(c+n-1)}{\Gamma(c)\Gamma(n)}. \tag{11}$$

where, by a well-known expansion of a ratio of gamma functions,

$$\frac{\Gamma(c+n-1)}{\Gamma(c)\Gamma(n)} = \frac{1}{\Gamma(c)} n^\Lambda \left\{1 + \frac{\Lambda(\Lambda-1)}{2n} + O\left(\frac{1}{n^2}\right)\right\}.$$

See also next Sec. 6 for more detail. Hence, combination of (9) with (11) gives

$$|||q_n^+(\cdot, f_0) - M||| \leq \frac{b}{\Gamma(c)} \frac{1}{n^{\frac{1}{4}}} \left|1 + \frac{5}{32n} + O\left(\frac{1}{n^2}\right)\right| \tag{12}$$

which, thanks to the elimination of ε, sharpens the basic inequality (2.16) in Ref. 5.

Bound (10) is also used, in Theorem 3.1 of Ref. 5, to estimate the coefficient $(1 - p_{n,k})$ of the "smooth" component in a decomposition of q_n^+,

$$p_{n,k} \leq \left(\frac{A}{(c/2)^{k-1}}\right) n^{-(1-2c)} \tag{13}$$

which holds true for every strictly positive c, and for some suitable constant A. Thanks to (11), (13) can be replaced with a new more precise inequality, i.e.

$$p_{n,k} \leq \frac{2^{k+\varepsilon}}{(k-1)!} \frac{\{C + \log(n-1)\}^{k-1}}{n-1} \tag{14}$$

where ε is any strictly positive number and C is the Euler–Mascheroni constant.

The paper is organised as follows. Section 2 explains the above-mentioned algorithm for the actual assessment of the probability distribution $p_n(\gamma)$, $\gamma \in G(n)$ and $n = 1, 2, \ldots$. Section 3 includes a description of a simple probabilistic framework which is useful for a precise formulation of the concept of depth. The exact form of the probability distribution (7) of the depth of a leaf is then deduced in Sec. 4. A few hints to the study of the probability distribution of the depth of a tree are given in Sec. 5. This distribution is involved in the aforesaid decomposition of q_n^+. The preliminary study of the law at issue yields the first term of a distinguished asymptotic expansion that we intend to develop and present in a separate paper, together with an analysis of the law of the height of a tree which is of interest, for example, in random search trees considered in computer science. Finally, Sec. 6 contains the proofs of (12) and (14).

2. EVALUATION OF $p_n(\gamma)$

The problem of determining $p_n(\gamma)$ in (5) can be solved by returning to (3)–(4) after putting

$$\mathfrak{I}(x; t, u) = \sum_{n\geq 0} \{u(1-e^{-t})\}^n \hat{q}_{n+1}^+(x; \varphi_0)$$

for every $t \geq 0$, x in $\mathbb{R}$ and u in [0, 1]. In fact, by (3)

$$\mathfrak{I}(x; t, u) = \varphi_0(x) + \sum_{n\geq 1} \{u(1-e^{-t})\}^n \frac{1}{n} \sum_{k=1}^{n} \frac{1}{2\pi} \cdot \int_0^{2\pi} \hat{q}_k^+(x\cos\theta; \varphi_0)\hat{q}_{n-k+1}^+(x\sin\theta; \varphi_0)\, d\theta$$

which, by the Cauchy notion of product of series and the identity $n^{-1} = \int_0^1 \sigma^{n-1} d\sigma$, becomes

$$\mathfrak{I}(x; t, u) = \varphi_0(x) + (1-e^{-t}) \int_0^u \left\{ \frac{1}{2\pi} \int_0^{2\pi} \mathfrak{I}(x\cos\theta; t, \sigma) \cdot \mathfrak{I}(x\sin\theta; t, \sigma) d\theta \right\} d\sigma \quad u \in [0, 1]. \tag{15}$$

Setting

$$\tilde{\varphi}(x; t, u) = \{1 - u(1-e^{-t})\}\mathfrak{I}(x; t, u)$$

and substituting in Eq. (15) gives

$$\tilde{\varphi}(x; t, u) = \{1 - u(1-e^{-t})\}\varphi_0(x) + u(1-e^{-t}) \cdot \int_0^u \left\{ \frac{1}{2\pi} \int_0^{2\pi} \tilde{\varphi}(x\cos\theta; t, \sigma)\tilde{\varphi}(x\sin\theta; t, \sigma) d\theta \right\} g(\sigma; u, t) d\sigma \tag{16}$$

where g stands for the probability density function

$$g(\sigma; u, t) = \frac{1 - u(1 - e^{-t})}{u\,\{1 - \sigma(1 - e^{-t})\}^2} \quad (0 < \sigma < u)$$

supported by $(0, u)$. It is clear that the solution of the primary problem (1) can be obtained from the solution of (16) by setting

$$\varphi(x, t) = \tilde{\varphi}(x; t, 1) \quad (x \in \mathbb{R}, t \geq 0).$$

Solving (16) instead of the primary problem has the advantage of bringing out a method for the determination of p_n which mimics the procedure to determine n-fold Wild convolutions. Specifically, think of (16) as fixed point problem and iterate to obtain

$$\tilde{\varphi}(x; t, u) = \sum_{n \geq 1} \sum_{\gamma \in G(n)} \tilde{p}_n(\gamma; t, u) c_\gamma(x; \varphi_0)$$

with

$$\tilde{p}_{|\gamma|}(\gamma; t, u) := \{1 - u(1 - e^{-t})\}\{u(1 - e^{-t})\}^{|\gamma|-1} p_{|\gamma|}(\gamma) \quad (\gamma \in G),$$

together with the fundamental relationship

$$\tilde{p}_{|\gamma|}(\gamma; t, u) = u(1 - e^{-t}) \int_0^u \tilde{p}_{|\gamma_l|}(\gamma_l; t, \sigma) \tilde{p}_{|\gamma_r|}(\gamma_r; t, \sigma) g(\sigma; u, t) d\sigma \qquad (17)$$

where $|\gamma|$ denotes the number of leaves of γ. After splitting both γ_l and γ_r in the same way, the process can be continued till one reaches leaves in the deepest level. Notice that, since $p_{|\gamma|}(\gamma) = 1$ if $\gamma \in G(1)$, one has $\tilde{p}_1(\gamma; t, u) = 1 - u(1 - e^{-t})$. These facts suggest the following procedure for the calculation of the p_n in a tree γ with n leaves and deepest level equal to i. Introduce the operation

$$f_1 * f_2(\sigma) := \sigma(1 - e^{-t}) \int_0^\sigma g(x; \sigma, t) f_1(x, t) f_2(x, t) dx$$

for any pair of function f_1 and f_2 defined on $(0, 1) \times (0, +\infty)$ in such a way that the integral is finite. Next, set $\lambda_1(u, t) := \{1 - u(1 - e^{-t})\}$ in each of the leaves of γ, find the leftmost pair of leaves at level i, erase this pair of leaves which makes the former "parent" node a leaf, and write down $\lambda_2(u, t) := \lambda_1(\cdot, t) * \lambda_1(\cdot, t)(u)$ in the new leaf. After erasing all the leaves at level i, in this way, proceed to erase pairs of leaves at level $(i - 1)$ and, for any pair, write down

$$\lambda_3(u, t) := \lambda_l(\cdot, t) * \lambda_r(\cdot, t)(u) \qquad (18)$$

in the leaf which replaces the corresponding "parent" node, where l (r, respectively) can be 1 or 2 according to whether the left (right, respectively) leaf of the pair was a leaf of γ or the "parent" node of a pair erased in the previous step. Then, after erasing all the leaves at level $(i-1)$, proceed to erase pairs of leaves at level $(i-2)$ in the same way and, for any pair, write down (18) in the leaf which replaces its "parent" node, keeping in mind that both l and r can belong to $\{1, 2, 3\}$. Once this has been done until only the "root" is left, one has written $\tilde{p}_{|\gamma|}(\gamma; t, u)$ in the "root".

Example 1. Consider tree γ_1 in Fig. 1, where $n = 8$ and $i = 4$. One wants to evaluate $p_8(\gamma_1)$. According to the above procedure, one starts with the number to be written in the parent node appearing at level 3:

$$\sigma_3(1-e^{-t}) \int_0^{\sigma_3} \{1-\sigma_4(1-e^{-t})\}^2 \frac{1-\sigma_3(1-e^{-t})}{\sigma_3\{1-\sigma_4(1-e^{-t})\}^2} d\sigma_3$$
$$= \sigma_3(1-e^{-t})\{1-\sigma_3(1-e^{-t})\}.$$

Now there are three subgraphs with leaves at level 3 (see Fig. 2). After erasing their leaves, one has the following values in the leaves named 1, 2, 3, 4:

$$\sigma_2(1-e^{-t})\{(1-e^{-t})\},$$
$$1-\sigma_2(1-e^{-t}),$$
$$\sigma_2(1-e^{-t}) \int_0^{\sigma_2} \{1-\sigma_3(1-e^{-t})\}^2 \sigma_3(1-e^{-t}) \frac{1-\sigma_2(1-e^{-t})}{\sigma_2\{1-\sigma_3(1-e^{-t})\}^2} d\sigma_3$$
$$= \frac{1}{2}\{\sigma_2(1-e^{-t})\}^2\{1-\sigma_2(1-e^{-t})\},$$

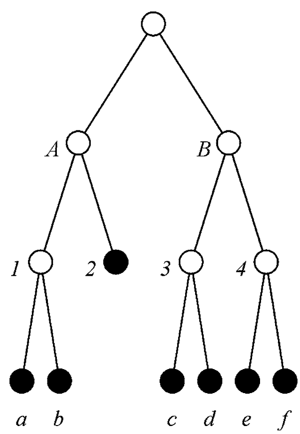

Fig. 2. Scheme for the evaluation of $P_8(\gamma_1)$ in Example 1.

and

$$\sigma_2(1-e^{-t})(1-\sigma_2)\{(1-e^{-t})\}.$$

respectively. So, only two graphs remain. They must be erased and their values are written in A and B, respectively:

$$\frac{1}{2}\{\sigma_1(1-e^{-t})\}^2\{1-\sigma_1(1-e^{-t})\} \qquad (\text{value of } \gamma_{1l})$$

$$\frac{1}{2\cdot 4}\{\sigma_1(1-e^{-t})\}^4\{1-\sigma_1(1-e^{-t})\} \qquad (\text{value of } \gamma_{1r}).$$

Hence,

$$\begin{aligned}\tilde{p}_8(\gamma_1;t,u) &= u(1-e^{-t})\frac{1}{2\cdot 2\cdot 4}\int_0^u \{1-\sigma_1(1-e^{-t})\}^2 \\ &\quad \{\sigma_1(1-e^{-t})\}^6\frac{1-u(1-e^{-t})}{u\{1-\sigma_1(1-e^{-t})\}^2}d\sigma_1 \\ &= \frac{1}{2\cdot 2\cdot 4\cdot 7}\{u(1-e^{-t})\}^7\cdot\{1-u(1-e^{-t})\}\end{aligned}$$

and

$$p_8(\gamma_1) = \frac{1}{2^4\cdot 7}.$$

It should be observed that

$$\sum_{\gamma\in G(n)} p_n(\gamma) = 1 \tag{19}$$

holds for every $n = 1, 2, \ldots$. In fact, combination of (3) and (5), recalling that φ is a characteristic function, gives

$$1 = \varphi(0,t) = \sum_{n\geq 1} e^{-t}(1-e^{-t})^{n-1}\sum_{\gamma\in G(n)} p_n(\gamma)$$

i.e.

$$\frac{1}{1-\xi} = \sum_{n\geq 1}\xi^{n-1}\sum_{\gamma\in G(n)} p_n(\gamma) \quad (\xi = 1-e^{-t}). \tag{20}$$

Now write $(1-\xi)^{-1} = \sum_{n\geq 1}\xi^{n-1}$ for every ξ in $(0, 1)$ to conclude that the coefficient of ξ^{n-1} in the right-hand side of (20) equals one for every $n = 1, 2, \ldots$.

Further relevant properties of the p_n will be discovered and studied in the next sections.

3. PROBABILISTIC FRAMEWORK

Before proceeding to the proof of the main results, it is worthwhile trying to introduce a suitable probabilistic setting. As a matter of fact, positivity of the $\tilde{p}_{|\gamma|}$ and (19) lead to think of $\tilde{p}_{|\gamma|}(\gamma; t, 1)$ as the probability of running into tree γ, i.e. into a specific physical situation. In view of this, look at G as a space of *outcomes* and, since G is countable, consider the class $\mathcal{G}$ of all subsets of G as the family of all *events* pertaining to an ideal experiment with outcomes in G. Examples of significant events will be described in connection with a few relevant instances of random elements, i.e. functions defined on G. First consider the random number ν defined to be the function from G to $\mathbb{N} = \{1, 2, \ldots\}$ which to each γ in G assigns the number $|\gamma|$ of leaves. As far as the depth $\delta_j = \delta_j(\gamma)$ is concerned, stipulate that $\delta_j(\gamma) = 0$ both if $j > |\gamma|$ and if $|\gamma| = 1 = j$. It should be observed that the random vector $(\nu, \delta_1, \ldots, \delta_{|\gamma|})$ is a one-to-one mapping of G into the range. Thus, each element of G can be characterized through a specific determination of such a random vector. In the previous notation, the *depht of a tree*, or *minimal depth*, is defined to be the random number

$$\gamma \mapsto \delta_{(1)}(\gamma) := \min\{\delta_1(\gamma), \ldots, \delta_{|\gamma|}(\gamma)\} \quad (\gamma \in G)$$

and the *height of a tree* by

$$\gamma \mapsto \delta_{(|\gamma|)}(\gamma) := \max\{\delta_1(\gamma), \ldots, \delta_{|\gamma|}(\gamma)\} \quad (\gamma \in G).$$

Once the measurable space $(G, \mathcal{G})$ has been specified, in order to define on it a probability measure $P^{(t)}$ which is consistent with the coefficients $\tilde{p}_n(\gamma; t, 1)$ appearing in the expression of $\tilde{\varphi}(x; t, 1)$, it is enough to put

$$P^{(t)}(A) := \sum_{\gamma \in A} e^{-t}(1 - e^{-t})^{|\gamma|-1} p_{|\gamma|}(\gamma) \quad (A \subset G).$$

This, in turn, yields

$$\begin{aligned}
P^{(t)}(\{\gamma\}) &= e^{-t}(1 - e^{-t})^{|\gamma|-1} p_{|\gamma|}(\gamma) \quad (\gamma \in G),\\
P^{(t)}(G(n)) &= \sum_{\gamma \in G(n)} P^{(t)}(\{\gamma\})\\
&= \sum_{\gamma \in G(n)} e^{-t}(1 - e^{-t})^{n-1} p_n(\gamma)\\
&= e^{-t}(1 - e^{-t})^{n-1} \quad \text{(in view of (19))}.
\end{aligned}$$

Since $G(n) = \{\nu = n\}$, it turns out that the probability distribution $P_\nu^{(t)}$ of ν is characterized by

$$P_\nu^{(t)}(\{n\}) = e^{-t}(1 - e^{-t})^{n-1} \quad (n = 1, 2, \ldots).$$

The computation of the probability distributions of depths is more involved and will be considered in next Secs. 4 and 5. It is based on specific difference equations concerned with conditional probabilities. In order to pave the way for their understanding, note that, under $P^{(t)}$, $p_n(\gamma)$ represents the conditional probability of $\{\gamma\}$ given that γ is assumed to belong to $G(n)$, that is

$$P^{(t)}(\{\gamma\} \mid G(n)) = \frac{P^{(t)}(\{\gamma\} \cap G(n))}{P^{(t)}(G(n))} = \frac{e^{-t}(1-e^{-t})^{n-1} p_n(\gamma)}{e^{-t}(1-e^{-t})^{n-1}} \quad (\gamma \in G)$$

with $p_n(\gamma) = 0$ if $\gamma \notin G(n)$. Hence, for any $A \subset G$,

$$P^{(t)}(A \mid G(n)) = \sum_{\gamma \in A \cap G(n)} p_n(\gamma).$$

In general, the evaluation of probability $P^{(t)}$ is made sensibly easy by the use of (17). As an example, one provides a new brief proof of an important result stated in Ref. 4 as Lemma 2.1. It says that, *for every n, the probability distribution of the number of leaves in the left subgraph γ_l of a graph γ with n leaves is uniform.* After defining the random number $\nu^{(l)}$ by $\nu^{(l)}(\gamma) = |\gamma_l|$ for every γ in G, one can restate the previous result in the following form:

$$P^{(t)}(\nu^{(l)} = j \mid G(n)) = \frac{1}{n-1} \quad (j = 1, \ldots, n-1; n = 2, 3, \ldots). \tag{21}$$

In fact, $P^{(t)}\{\nu^{(l)} = j, \nu = n\} = P^{(t)}\{\nu^{(l)} = j, \nu^{(r)} = n-j\}$ holds true for every j in $\{1, \ldots, n-1\}$, provided that $\{r_1, r_2\}$ is understood as a simplified form for $r_1 \cap r_2$. Now, by (17) with $u = 1$, and the ensuing discussion,

$$\begin{aligned} P^{(t)}\{\nu^{(l)} = j, \nu^{(r)} = n-j\} &= (1-e^{-t}) \int_0^1 \{1 - \sigma(1-e^{-t})\} \\ &\quad \cdot \{\sigma(1-e^{-t})\}^{j-1} \cdot \{1-\sigma(1-e^{-t})\} \\ &\quad \cdot \{\sigma(1-e^{-t})\}^{n-j-1} \cdot g(\sigma; 1, t) d\sigma \\ &= \frac{1}{n-1} e^{-t}(1-e^{-t})^{n-1} = \frac{1}{n-1} P^{(t)}\{\nu = n\} \end{aligned} \tag{22}$$

which proves (21).

In view of the obvious equality

$$P^{(t)}(\nu^{(l)} = 0 \mid G(1)) = 1$$

and from (22) one gets

$$P^{(t)}\{\nu^{(l)} = j\} = \sum_{n\geq 1} P^{(t)}(\nu^{(l)} = j \mid G(n)) e^{-t}(1-e^{-t})^{n-1}.$$

Thus,

$$P^{(t)}\{\nu^{(l)} = 0\} = e^{-t}$$

whilst, for any j in $\mathbb{N}$,

$$P^{(t)}\{\nu^{(l)} = j\} = e^{-t}(1-e^{-t})^j \sum_{n\geq 0} \frac{(1-e^{-t})^n}{n+j}. \tag{23}$$

The series in the right-hand side of (23), as a function of $(1-e^{-t})$, is named *Lerch trascendent*, denoted by $\phi(1-e^{-t}, 1, j)$. It admits an integral representation leading to

$$\begin{aligned} P^{(t)}\{\nu^{(l)} = j\} &= e^{-t}(1-e^{-t})^j \phi(1-e^{-t}, 1, j) \\ &= e^{-t}(1-e^{-t})^j \int_0^{+\infty} \frac{e^{-(j-1)x}}{e^x - 1 + e^{-t}} dx \quad (j = 1, 2, \ldots). \end{aligned} \tag{24}$$

Moreover,

$$\lim_{t\to+\infty} \frac{\phi(1-e^{-t}, 1, j)}{t} = 1$$

holds for every j in $\mathbb{N}$. See, for example, Ref. 16.

4. DISTRIBUTION OF THE DEPTH OF A LEAF

The present section, as well as the next one, is concerned with some results on the depth of a leaf and of a tree, respectively, to be used in the study of the convergence of Wild's sums. On account of this, the first part deals with the meaning of depth of a leaf in connection with the notion of Wild's n-fold product. For the sake of explanatory clearness, refer to the particular graph γ_1 of Fig. 1 which has linked to it the n-fold "product"

$$(((\varphi_0 \bullet \varphi_0) \bullet \varphi_0) \bullet (((\varphi_0 \bullet \varphi_0) \bullet \varphi_0) \bullet (\varphi_0 \bullet \varphi_0))).$$

This representation shows that the one-to-one correspondence between $G(n)$ and the set of all n-fold "products" regards, in point of fact, the correspondence between $G(n)$ and "bracketings", the arrangement of brackets being crucial because of the nonassociativity of the Wild convolution. Notice that the number of enclosures coincides with the number of nodes of the corresponding graph. Assigning to each φ_0, appearing in a specific "product", the number which marks the corresponding leaf in the tree linked to such a product, makes clear that the *depth of the leaf at* *issue equals the number of the pairs of brackets which enclose the φ_0 labelled in the aforesaid way*. Enclosures, in each graph, as explained in Sec. 13 of Ref. 18, are, in turn, associated with collisions in pairs of particles according to the ramification of each graph. Thus the depth of leaf j can be thought of as the number of the collisions of the corresponding molecule with the remaining $(n-1)$ particles in the same bath. Firstly, one considers the problem of determining

$$P^{(t)}(\delta_j = d \mid G(n)) = \sum_{\gamma \in \{\delta_j = d\} \cap G(n)} p_n(\gamma)$$

for any integer d in $\mathbb{N}_0$. After extending p_n, by additivity, to any subset of G, one has

$$\begin{aligned} P^{(t)}(\delta_j = d \mid G(n)) &=: p_n\{\delta_j = d\} \\ &= \mathbb{1}_{\{0\}}(d) \quad \text{if } n = 1 \end{aligned}$$

while, for any $n \geq 2$,

$$\begin{aligned} p_n\{\delta_j = d\} &= \sum_k p_n(\{\delta_j = d\} \cap \{\gamma \in G(n) : |\gamma_l| = k\}) \qquad (25) \\ &= \sum_k p_n(\delta_j = d \mid \gamma \in G(n), |\gamma_l| = k) p_n\{\gamma \in G(n), |\gamma_l| = k\} \\ &= \sum_{k=1}^{n-1} p_n(\delta_j = d \mid \gamma \in G(n), |\gamma_l| = k) \frac{1}{n-1} \quad \text{(from (21)).} \end{aligned}$$

Moreover,

$$p_n(\delta_j = d \mid \gamma \in G(n), |\gamma_l| = k) = \begin{cases} p_n(\delta_j(\gamma_l) = d - 1 \mid \gamma \in G(n), |\gamma_l| = k) \\ \qquad \text{if } k \geq j \\ p_n(\delta_{j-k}(\gamma_r) = d - 1 \mid \gamma \in G(n), |\gamma_l| = k) \\ \qquad \text{if } k < j \end{cases}$$

where, from (17) and its direct consequences explained in Sec. 2,

$$p_n(\delta_j(\gamma_l) = d - 1 \mid \gamma \in G(n), |\gamma_l| = k) = p_k\{\delta_j = d - 1\} \qquad (26)$$

and

$$p_n(\delta_{j-k}(\gamma_r) = d - 1 \mid \gamma \in G(n), |\gamma_l| = k) = p_{n-k}\{\delta_{j-k} = d - 1\}. \qquad (27)$$

These elementary facts can be used , together with the method of generating functions, to prove

Proposition 2. *For any n in $\mathbb{N}$, let j and d be elements of $\{1, \dots, n\}$ and $\{0, \dots, n-1\}$, respectively. Then,*

$$P^{(t)}(\delta_j = d \mid G(1)) = \mathbb{1}_{\{0\}}(d)$$

and, for any $n = 2, 3, \dots$,

$$P^{(t)}(\delta_j = d \mid G(n)) = \sum_k \frac{1}{(j-1)!(n-j)!} |s(j-1, d-k)| \cdot |s(n-j, k)|$$

where $s(n, k)$ denotes a Stirling number of the first kind (hence, $s(0, 0) = 1$, $s(n, 0) = 0$ if $n > 0$ and $s(n, k) = 0$ if $k > n$).

The Stirling numbers of the first and second kind are the coefficients of the expansions of the factorials into powers and of the powers into factorials, respectively. In particular,

$$\begin{aligned}(t)_n : &= t(t-1)\dots(t-n+1) = \sum_{k=0}^{n} s(n,k)t^k \quad n = 1, 2, \dots \\ &= 1 \quad \text{if } n = 0.\end{aligned}$$

See Chapter 8 of Ref. 8 for a comprehensive treatment of Stirling numbers.

Proof of Proposition 2: It is enough to deal with the case of $n \geq 2$. Combine (25) with (26)–(27) to obtain

$$(n-1)p_n\{\delta_j = d\} = \sum_{k=1}^{j-1} p_{n-k}\{\delta_{j-k} = d-1\} + \sum_{k=j}^{n-1} p_k\{\delta_j = d-1\} \qquad (28)$$

for every $j = 1, \dots, n$, with $\sum_{k=1}^{0} = \sum_{k=n}^{n-1} := 0$. Multiply (28) by x^d and sum over $d = 0, \dots, n-1$ to obtain

$$(n-1)g_{j,n}(x) = x\left\{\sum_{k=1}^{j-1} g_{j-k,n-k}(x) + \sum_{k=j}^{n-1} g_{j,k}(x)\right\}$$

with

$$\begin{aligned}g_{j,n}(x) &:= \sum_{d=0}^{n-1} p_n\{\delta_j = d\}x^d \quad j = 1, \dots, n;\ n = 2, 3, \dots \\ g_{1,1}(x) &= 1.\end{aligned} \qquad (29)$$

Now multiply both sides of the last difference equation by ξ^j and sum over $j = 1, \ldots, n$. After setting

$$V_n(x, \xi) := \sum_{j=1}^{n} \xi^j g_{j,n}(x) \quad n = 2, 3, \ldots$$
$$V_1(x, \xi) := \xi \tag{30}$$

one gets

$$(n-1)V_n(x, \xi) = x \sum_{j=1}^{n-1} (1 + \xi^{n-j}) V_j(x, \xi).$$

Finally, take any x and ξ in the unit disk, multiply the above equation by z^n with $|z| < 1$, sum over $n = 2, 3, \ldots$ and put

$$V(z; x, \xi) := \sum_{n \geq 2} z^n V_n(x, \xi). \tag{31}$$

This gives

$$\sum_{n\geq 2} (n-1) V_n(x, \xi) z^n = \frac{1 + \xi - 2\xi z}{(1-z)(1-z\xi)} xz V(z; x, \xi) + \frac{1 + \xi - 2\xi z}{(1-z)(1-z\xi)} \xi x z^2$$

i.e.

$$z \frac{\partial}{\partial z} V(z; x, \xi) = \left(1 + \frac{1 + \xi - 2\xi z}{(1-z)(1-z\xi)} xz\right) V(z; x, \xi) + \frac{1 + \xi - 2\xi z}{(1-z)(1-z\xi)} \xi x z^2.$$

For any z in $(0, 1)$, the general solution of this equation can be written as

$$V(z; x, \xi) = \frac{z}{((1-z)(1-z\xi))^x} \{c - \xi\, ((1-z)(1-z\xi))^x\}$$

and, in view of well-known expressions for the Stirling numbers (see, for example, Ref. 9), one obtains

$$V(z; x, \xi) = cz \sum_{n\geq 0} z^n \sum_{k=0}^{n} \xi^k \frac{1}{(n-k)!k!} \sum_{j=0}^{n-k} |s(n-k, j)| x^j \sum_{r=0}^{k} |s(k, r)| x^r - z\xi$$

where, in view of (31), the coefficient of z must be zero, which is tantamount to saying that c must equal ξ. Hence,

$$\begin{aligned} V(z; x, \xi) &= \sum_{n\geq 2} z^n \sum_{j=1}^{n} \xi^j \frac{1}{(n-j)!(j-1)!} \sum_{k=0}^{n-j} |s(n-j, k)| x^k \sum_{r=0}^{j-1} |s(j-1, r)| x^r \\ &= \sum_{n\geq 2} z^n \sum_{j=1}^{n} \xi^j \frac{1}{(n-j)!(j-1)!} \sum_{d\geq 0} x^d \sum_{k+r=d} |s(n-j, k)||s(j-1, r)| \end{aligned}$$

which, via (31), entails

$$V_n(x,\xi) = \sum_{j=1}^{n} \xi^j \frac{1}{(n-j)!(j-1)!} \sum_{d\geq 0} x^d \sum_k |s(n-j,k)| \cdot |s(j-1,d-k)|$$

and, via (30),

$$g_{j,n}(x) = \sum_{d\geq 0} x^d \frac{1}{(n-j)!(j-1)!} \sum_k |s(n-j,k)| \cdot |s(j-1,d-k)|.$$

At this stage, compare this expression with (29) to complete the proof. □

By resorting to the former of the interpretations of (7) given in Sec. 1, it is immediate to prove

Proposition 3. *For any n in $\mathbb{N}\setminus\{1\}$:*

$$E_n(\delta_1) = \frac{1}{n-1} + \frac{1}{n-2} + \cdots + \frac{1}{2} + 1 = E_n(\delta_n)$$

and

$$E_n(\delta_j) = 1 + \frac{1}{2} + \cdots + \frac{1}{(j-1)} + \left\{1 + \cdots + \frac{1}{(n-j)}\right\}$$

if $1 < j < n$. Analogously,

$$Var_n(\delta_j) = \frac{1}{2}\cdot\frac{1}{2} + \cdots + \frac{1}{n-2}\cdot\frac{n-3}{n-2}\cdot\frac{1}{n-1}\cdot\frac{n-2}{n-1} = Var_n(\delta_n)$$

and

$$Var_n(\delta_j) = \sum_{k=1}^{j-1} \frac{1}{k}\left(1-\frac{1}{k}\right) + \sum_{k=1}^{n-j} \frac{1}{k}\left(1-\frac{1}{k}\right).$$

This section continues with some results concerning the (unconditional) distribution of δ_j. In other words, one wants to determine $P^{(t)}\{\delta_j = d\}$ for any j in $\mathbb{N}$ and for every d in $\mathbb{N}_0 := \mathbb{N}\cup\{0\}$. Clearly,

$$\begin{aligned}
P^{(t)}\{\delta_j = d\} &= \sum_{n\geq 1} e^{-t}(1-e^{-t})^{n-1} \cdot p_n\{\delta_j = d\} \\
&= \mathbb{1}_{\{0\}}(d) \sum_{n<j} e^{-t}(1-e^{-t})^{n-1} + \mathbb{1}_{\mathbb{N}}(d) \sum_{n\geq j} e^{-t}(1-e^{-t})^{n-1} \\
&\quad \cdot \frac{1}{(j-1)!(n-j)!} \sum_k |s(j-1,d-k)| \cdot |s(n-j,k)|
\end{aligned}$$

$$= 1\!\!1_{\{0\}}(d)\left\{1-(1-e^{-t})^{j-1}\right\} + 1\!\!1_{\mathbb{N}}(d)\frac{e^{-t}}{(j-1!)}(1-e^{-t})^{j-1}$$

$$\cdot \sum_{k} |s(j-1,d-k)| \cdot \sum_{\sigma} \frac{|s(\sigma,k)|}{\sigma!}(1-e^{-t})^{\sigma}$$

$$= 1\!\!1_{\{0\}}(d)\left\{1-(1-e^{-t})^{j-1}\right\} + 1\!\!1_{\mathbb{N}}(d)(1-e^{-t})^{j-1}\frac{e^{-t}}{(j-1!)}$$

$$\cdot \sum_{k} |s(j-1,d-k)| \frac{t^k}{k!}$$

(see Corollary 8.1 in Ref. 8).

Hence, one is in a position to state

Proposition 4. *For any j in $\mathbb{N}$ and d in $\mathbb{N}_0$ one has*

$$P^{(t)}\{\delta_j = d\} = \begin{cases} 1-(1-e^{-t})^{j-1} & \textit{if } d=0 \\ (1-e^{-t})^{j-1} \displaystyle\sum_{m=0}^{d\wedge(j-1)} \frac{|s(j-1,m)|}{(j-1)!} \cdot \frac{e^{-t}t^{d-m}}{(d-m)!} & \textit{if } d \in \mathbb{N} \end{cases}$$

which entails

$$P^{(t)}(\delta_j = d \mid \cup_{n\geq j} G(n)) = \sum_{m=0}^{d\wedge(j-1)} \frac{|s(j-1,m)|}{(j-1)!} \cdot \frac{e^{-t}t^{d-m}}{(d-m)!} \quad (d \in \mathbb{N}).$$

The last expression says that the conditional probability distribution of δ_j, under the hypothesis that a tree has at least j leaves, is the same as the distribution of the sum of two independent random variables, the former being distributed according to (8) with $\theta = 1$, $n = j-1$, the latter having the Poisson distribution with parameter t. Thus,

$$E^{(t)}(\delta_j \mid \cup_{n\geq j} G(n)) = 1 + \cdots + \frac{1}{j-1} + t$$

$$\mathrm{Var}^{(t)}(\delta_j \mid \cup_{n\geq j} G(n)) = \frac{1}{2}\cdot\frac{1}{2} + \cdots + \frac{1}{j-1}\cdot\left(1-\frac{1}{j-1}\right) + t \quad \text{if } j \geq 2$$

and

$$E^{(t)}(\delta_1) = E^{(t)}(\delta_1 \mid \cup_{n\geq 1} G(n)) = t$$

$$\mathrm{Var}^{(t)}(\delta_1) = \mathrm{Var}^{(t)}(\delta_1 \mid \cup_{n\geq 1} G(n)) = t$$

which entails

Proposition 5. *For any j in $\mathbb{N}\setminus\{1\}$ and d in $\mathbb{N}_0$, one has*

$$E^{(t)}(\delta_j) = \left\{1 + \cdots + \frac{1}{j-1} + t\right\}(1-e^{-t})^{j-1}$$

and

$$\mathrm{Var}^{(t)}(\delta_j) = (1-e^{-t})^{j-1}\left[\frac{1}{2}\cdot\frac{1}{2} + \cdots + \frac{1}{j-1}\cdot\left(1 - \frac{1}{j-1}\right) + t \right.$$
$$\left. + (1-(1-e^{-t})^{j-1})\cdot\left(1 + \cdots + \frac{1}{j-1} + t\right)^2\right].$$

5. REMARKS ABOUT THE DISTRIBUTION OF THE DEPTH OF A TREE

Recall that the term *depth of a tree* designates the random variable $\delta_{(1)}$ as defined in Secs. 1 and 3. In some branches of science an important role is played by the concept of *height of a tree* which corresponds to the random variable

$$\gamma \mapsto \delta_{(|\gamma|)}(\gamma) := \max\{\delta_1(\gamma), \ldots, \delta_{|\gamma|}(\gamma)\} \quad (\gamma \in G).$$

See, for example, Refs. 14, 15, and 20, and the references quoted therein. In computer science research, the knowledge of the probability distribution of $\delta_{(|\gamma|)}$ is required in connection with data compression schemes. Here, one hints to the computation of the probability distribution of $\delta_{(1)}$ in order to get a slight improvement in decomposition (3.1) in Ref. 5. For further motivation, see the final part of the next section. The above-quoted arguments can be worked out to tackle the problem of approximating the distribution of $\delta_{(1)}$. Thus, the resulting approach will be developed in a forthcoming paper. A starting point for determining exact forms of the distribution of $\delta_{(1)}$ is contained in the following

Proposition 6. *Set*

$$q_{n,k} := P^{(t)}(\delta_{(1)} \geq k \mid G(n)) \tag{32}$$

for any $k = 0, 1, \ldots$ and $n = 1, 2, \ldots$. Then, the recursion relation

$$q_{n,k} = \frac{1}{n-1}\sum_{\nu=1}^{n-1} q_{\nu,k-1}q_{n-\nu,k-1}$$

holds for every $k = 1, 2, \dots$ *and* $n = 1, \dots,$ *with the initial condition* $q_{n,0} = 1$ $(n = 1, 2, \dots)$*. Moreover,* $q_{n,k} = 0$ *if* $n < 2^k$*, and*

$$\psi_k(z) = \frac{1}{z} \sum_{n \geq 2^k} z^n q_{n,k} \quad (|z| < 1)$$

satisfies

$$\psi_k'(z) = \psi_{k-1}^2(z) \quad (|z| < 1, k = 1, 2, \dots)$$

with the initial condition $\psi_0(z) = (1 - z)^{-1}$ *for* $|z| < 1$.

Proof: The nonlinear recursion relation is a direct consequence of the definition of $\delta_{(1)}$ combined with the usual computational method for conditional probabilities described in Secs. 2–3. To get the above differential equation, multiply $(n - 1)q_{n,k} = \sum_{\nu=1}^{n-1} q_{\nu,k-1}q_{n-\nu,k-1}$ by z^n and sum over $n = 2, 3, \dots$ to obtain

$$\begin{aligned}\sum_{n\geq 2} nz^n q_{n,k} - \sum_{n \geq 2} z^n q_{n,k} &= \sum_{n\geq 2} z^n \sum_{\nu=1}^{n-1} q_{\nu,k-1}q_{n-\nu,k-1}\\ &= \sum_{\nu \geq 1} q_{\nu,k-1} z^\nu \sum_{n \geq 1+\nu} z^{n-\nu} q_{n-\nu,k-1}\\ &= z^2 \psi_{k-1}(z) \quad (k = 1, 2, \dots)\end{aligned}$$

i.e.

$$z\frac{d}{dz}[z\psi_k(z)] - z\psi_k(z) = z^2\psi_{k-1}(z).$$

$\square$

A recursion relation of the same type as (32) can be found in Ref. 15 in connection with the probability law of $\delta_{(|\gamma|)}$. In Ref. 1, Proposition 6 is utilized to determine an exact expression for the probability distribution of $\delta_{(1)}$. Here, one gives an upper bound for $p_n\{\delta_{(1)} \leq d\}$.

For the sake of notational simplicity, set

$$A_j(d) := \{\delta_j \leq d\} \text{ and } A_{i_1,\dots,i_m}(d_1, \dots d_m) := A_{i_1}(d_1) \cap \dots \cap A_{i_m}(d_m).$$

Then, one can write

$$p_n\{\delta_{(1)} \leq d\} = p_n\left(\bigcup_{j=1}^{n} A_j(d)\right)$$

and, from the *inclusion and exclusion principle* (see, e.g., Sec. 4.5 in Ref.11),

$$p_n\left(\bigcup_{j=1}^{n} A_j(d)\right) = \sum_{j=1}^{n} p_n(A_j(d)) - \sum_{1\le j_1<j_2\le n} p_n(A_{j_1,j_2}(d,d) + \cdots$$
$$+(-1)^{n-1} p_n(A_{1,\ldots,n}(d,\ldots,d)).$$

At this stage, the *Bonferroni inequalities* (see Ref. 11 once again)

$$\begin{aligned}
p_n\{\delta_{(1)}\le d\} &\le \sum_{j=1}^{n} p_n(A_j(d))\\
p_n\{\delta_{(1)}\le d\} &\ge \sum_{j=1}^{n} p_n(A_j(d)) - \sum_{1\le j_1<j_2\le n} p_n(A_{j_1,j_2}(d,d))\\
p_n\{\delta_{(1)}\le d\} &\le \sum_{j=1}^{n} p_n(A_j(d)) - \sum_{1\le j_1<j_2\le n} p_n(A_{j_1,j_2}(d,d))\\
&\quad - \sum_{1\le j_1<j_2<j_3\le n} p_n(A_{j_1,j_2,j_3}(d,d,d))\\
\ldots \quad &\ldots \quad \ldots
\end{aligned}$$

say that if one stops the inclusion and exclusion formula after an even (odd, respectively) number of sums one gets a lower (upper, respectively) bound. Here, one confines oneself to exhibiting an exact form for the first sum, derived from Proposition 2 via well-known properties of Stirling's numbers.

Proposition 7. *The following equality is valid for every d in $\mathbb{N}$:*

$$\sum_{j=1}^{n} p_n\{\delta_j\le d\} = \frac{1}{(n-1)!}\sum_{\nu=0}^{d} 2^{\nu}\,|s(n-1,\nu)|\,.$$

Moreover, setting $B_n(d) := \{C+\log(n-1)\}^{d-1}/(n-1)$, where C stands for the Euler–Mascheroni constant, given any $\varepsilon>0$ there is $\bar{n}$ in $\mathbb{N}$ such that

$$p_n\{\delta_{(1)}\le d\} \le \frac{2^{d+\varepsilon}}{(d-1)!}B_n(d)$$

holds for all $n\ge\bar{n}$.

Proof: By elementary arguments,

$$\sum_{j=1}^{n} p_n\{\delta_j \le d\}$$

$$= \sum_{j=1}^{n}\sum_{\nu=0}^{d} p_n\left\{\delta_j = \nu\right\}$$

$$= \sum_{j=1}^{n}\sum_{\nu=0}^{d}\sum_{k} \frac{1}{(j-1)!(n-j)!} \cdot |s(j-1,\nu-k)| \cdot |s(n-j,k)|$$

(from Proposition 4)

$$= \sum_{\nu=0}^{d}\sum_{k} \frac{1}{(n-1)!}\sum_{j}\binom{n-1}{j-1} \cdot |s(j-1,\nu-k)| \cdot |s(n-j,k)|$$

$$= \sum_{\nu=0}^{d}\sum_{k} \frac{1}{(n-1)!}\sum_{j}\binom{n-1}{j} \cdot |s(j,\nu-k)| \cdot |s(n-j-1,k)|$$

$$= \frac{1}{(n-1)!}\sum_{\nu=0}^{d}\sum_{k}\binom{\nu}{k} \cdot |s(n-1,\nu)| \quad \text{(see, e.g., Sec. 8.7 of Ref. 8)}$$

$$= \frac{1}{(n-1)!}\sum_{\nu=0}^{d} 2^{\nu}\,|s(n-1,\nu)|\,.$$

Passing to the second part, recall that the Stirling numbers of the first kind admit asymptotic expressions such as

$$|s(n+1,k+1)| \sim n!\frac{[\log(n+1)+C]^k}{k!}.$$

See Ref. 8. Then,

$$\frac{1}{(n-1)!}\sum_{\nu=0}^{d} 2^{\nu}|s(n-1,\nu)| \sim \frac{2^d}{(d-1)!}\frac{1}{n-1}\{C+\log(n-1)\}^{d-1} \quad (n \to +\infty)$$

which, combined with Bonferroni's inequalities, completes the proof. □

For an alternative upper bound for $p_n\{\delta_{(1)} \le d\}$, see Ref. 1. In any case, in order to get sharper evaluations of $q_{n,k}$, one ought to investigate the order of smallness of sums

$$\sum_{1\le i_1<\cdots<i_m\le n} p_n(A_{i_1,\ldots,i_m}(d_1,\ldots,d_m))$$

whose exact expression could be deduced from multivariate generating functions. This subject will be developed in the paper at the present time being prepared, previously mentioned.

6. APPLICATION TO RATES OF CONVERGENCE OF WILD SUMS

In seeking bounds on the error made when the Wild summation for solutions of the Boltzmann equation for a gas of Maxwellian molecules is truncated at the nth stage, in Refs. 4 and 5 one puts special emphasis on the expectation $T(n)$, with respect to p_n, of the random variable $W(\gamma) := \sum_{j=1}^{n}(c/2)^{\delta_j(\gamma)}$ when c is any element of interval (0, 1),

$$T(n) := E_n(W) = \sum_{j=1}^{n} E_n\left(\left(\frac{c}{2}\right)^{\delta_j}\right)$$

where E_n represents expectation with respect to p_n. From the previous analysis based on generating functions, it easily follows that $T(n)$ is nothing else than the value at $(x,\xi) = (c/2, 1)$ of $V_n(x,\xi)$, i.e.

$$T(n) = V_n(c/2, 1) \quad (n \in \mathbb{N}).$$

Lemma 1.4 in Ref. 4 states the upper bound (10) for T_n. As far as the computation of $A(\varepsilon)$ is concerned, in the proof of this lemma one defines

$$\begin{aligned} N_0 &= \min\left\{n \in \mathbb{N} : \frac{c}{c+\varepsilon}\left(\frac{n}{n-1}\right)^{1-c-\varepsilon} \le 1\right\} \\ &= \left\langle 1 + \frac{1}{\left(\frac{c+\varepsilon}{c}\right)^{1-c-\varepsilon} - 1}\right\rangle \end{aligned}$$

with $\langle x\rangle$ standing for the smallest integer $\ge x$, and one obtains

$$A(\varepsilon) = \max\{k^{1-c+\varepsilon}T(k) : k = 1,\ldots,N_0\}. \tag{33}$$

Notice that, in view of the definition of N_0, (10) would be devoid of sense for $\varepsilon = 0$. In any case, the importance of (10) is that it leads to state Theorem 1.9 in Ref. 4 in the following form: *If ϕ is any convex functional on probability density functions such that*

$$\phi(f \circ g) \le \frac{c}{2}(\phi(f) + \phi(g)) \tag{34}$$

is in force for all pairs (f, g) *of densities with zero mean and unit variance and for some* c *in* $(0, 1)$*, then for any* $\varepsilon > 0$ *there is a constant* b *such that*

$$\phi(q_n^+(\cdot; f_0)) \leq bA(\varepsilon)\frac{n^\varepsilon}{n^{1-c}}. \tag{35}$$

Moreover, in Ref. 5, which deals with the Kac model, one considers all probability densities with finite forth moment, and one determines a distinguished convex functional ϕ^* satisfying (34) with $c = \Lambda + 1 + \eta$, Λ $(= -1/4)$ being the least negative eigenvalue of the operator $\mathcal{L}$ defined in Sec. 1 of Ref. 5, and η being any positive number satisfying $0 < \Lambda + 1 + \eta < 1$. It is important to recall that ϕ^* satisfies

$$\phi^*(q_n^+(\cdot, f_0)) = |||q_n^+(\cdot, f_0) - M|||$$

with $M = M(x) = (2\pi)^{-1/2}\exp\{-x^2/2\}$, $x \in \mathbb{R}$, and

$$|||g||| = \sup_{\xi \neq 0} \frac{\left|\int_{\mathbb{R}} e^{i\xi x} g(x)dx\right|}{\xi^4}$$

whenever $\int_{\mathbb{R}} x^k g(x)dx = 0$ if $k = 0, 1, 2, 3$ and $\int_{\mathbb{R}} x^4 |g(x)|\, dx < +\infty$. Hence, from (35) with $c = \Lambda + 1 + \eta$, and in view of the arbitrariness of η, one can write

$$|||q_n^+(\cdot, f_0) - M||| \leq bA(\varepsilon)n^{\Lambda+\varepsilon} \tag{36}$$

which expresses the decay of the error in norm $||| \cdot |||$. See Theorem 2.2 in Ref. 5.

The above description is enough to point out that (36) could be improved if it were possible to replace (10) with a more accurate estimate of $T(n)$. As a matter of fact, this can be done by virtue of the arguments developed in Sec. 4.

Proposition 8. *For every* x *and* c *in* $(0, 1)$ *one has*

$$V_n(x, \xi) = \sum_{j=1}^{n} \xi^j \sum_k \frac{|s(n-j, k)|}{(n-j)!} x^k \sum_m \frac{|s(j-1, m)|}{(j-1)!} x^m$$

$$= \sum_{j=1}^{n} \binom{x + n - j - 1}{x - 1} \binom{x + j - 2}{x - 1} \xi^j$$

and, therefore,

$$T(n) := V_n\left(\frac{c}{2}, 1\right) = \binom{c - 2 + n}{c - 1} = \frac{\Gamma(c + n - 1)}{\Gamma(c)\Gamma(n)}$$

which admits the asymptotic expansion

$$T(n) \sim \frac{1}{n^{1-c}} \frac{1}{\Gamma(c)} \sum_{k \geq 0} A_k(1 - c)n^{-k} \quad (n \to \infty)$$

with

$$A_0(1-c) \equiv 1, \quad A_k(1-c) = \frac{1}{k}\sum_{m=0}^{k-1}\binom{1-c-m}{k-m+1}A_m(1-c) \quad (k = 1, 2, \ldots).$$

Recall that the above asymptotic expansion must be understood in the sense that

$$T(n) = \frac{1}{n^{1-c}}\frac{1}{\Gamma(c)}\left\{1 + A_1(1-c)\frac{1}{n} + \cdots + A_N(1-c)\frac{1}{n^N}\right\}$$
$$+ o\left(\frac{A_N(1-c)}{\Gamma(c)}\frac{1}{n^{N+1-c}}\right) \quad (n \to \infty).$$

Proof of Proposition 8: From Sec. 4,

$$V_n(x,\xi) = \sum_{j=1}^{n}\xi^j M_j, \quad M_j := \sum_{d\geq 0}x^d\sum_k \frac{|s(n-j,k)|}{(n-j)!}\frac{|s(j-1,d-k)|}{(j-1)!}.$$

Therefore, M_j is the probability generating function of the convolution of probabilities $|s(n-j,k)|/(n-j)!$ with probabilities $|s(j-1,m)|/(j-1)!$ and, hence, it coincides with the product of the generating functions of these sequences, i.e.:

$$M_j = \frac{(n-j-1+x)_{n-j}}{(n-j)!}\frac{(j-2+x)_{j-1}}{(j-1)!} = \binom{x+n-j-1}{x-1}\binom{x+j-2}{x-1}.$$

In particular,

$$V_n(x,1) = \sum_{j=1}^{n}\binom{x+n-j-1}{x-1}\binom{x+j-2}{x-1} = \binom{2x+n-2}{2x-1}$$

and

$$T(n) = V_n\left(\frac{c}{2},1\right) = \binom{c+n-2}{c-1} = \frac{\Gamma(c+n-1)}{\Gamma(c)\Gamma(n)}.$$

Finally, the second part of Proposition 8 follows from the asymptotic expansion of a ratio of gamma functions proved in Ref. 21. □

Now, thanks to Proposition 8, with $1-c = 1/4$, and the brief discussion before it, Theorem 2.2 in Ref. 5 can be restated in a sharper form as it follows:

For any probability density f_0 satisfying $\int_{\mathbb{R}} x^4 f_0(x)\,dx < +\infty$ there is a finite constant a' so that for all n,

$$|||q_n^+(\cdot, f_0) - M||| \leq a' n^{\Lambda}.$$

For the definition of a', see (12) and Proposition 8.

The final remark is about an application of Proposition 7. As mentioned in Sec. 1, such an application is related to a decomposition of q_n^+ proved in Ref. 5. Indeed, it turns out that q_n^+ can be written as a mixture (i.e., a convex combination) of probability density functions $B_{n,k}$ and $U_{n,k}$ with weights

$$q_{n,k+1} = P^{(t)}(\delta_{(1)} \geq k+1 \mid G(n))$$

and

$$p_{n,k} := 1 - q_{n,k+1}$$

respectively. Now, since $B_{n,k}$ is "smooth", $q_{n,k+1}$ gives a quantitative measure of the "portion of smoothess" in q_n^+ and, therefore, it ought to be high and increase rather fastly when n diverges. In point of fact, Theorem 3.1 in Ref. 5 states that, under suitable regularity conditions for the initial datum in (1), given any c in $(0, 1/2)$, one can determine some constant γ depending only on k and on the Linnik functional (sometimes identified with the Fisher information) of the initial datum, such that

$$||B_{n,k}||_{H^{k/2}(\mathbb{R})} \leq \gamma. \tag{37}$$

Moreover, $U_{n,k}$ satisfies

$$|||U_{n,k} - M||| \leq \phi \tag{38}$$

and there is a positive number A depending only on c so that (13) holds true.

Now, in view of Proposition 7, Theorem 3.1 in Ref. 5 can be simplified and improved in the following terms:

Let the initial datum in (1) be a probability density function with finite Linnik functional. Then (37)–(38) hold true and, for any $\varepsilon > 0$, there is $\bar{n} = \bar{n}(\varepsilon)$ such that

$$p_{n,k} \leq \frac{2^{k+\varepsilon}}{(k-1)!} \frac{\{C + \log(n-1)\}^{k-1}}{n-1}$$

is valid for every $n > \bar{n}$, C being the Euler–Mascheroni constant.

ACKNOWLEDGMENTS

The authors would like to thank Eric A. Carlen and Maria C. Carvalho for stimulating discussions and helpful suggestions, and two anonymous referees. This work was partially supported by the *Ministero dell'Istruzione, dell'Università e della Ricerca* (MIUR National Project "Problemi matematici delle teorie cinetiche" COFIN 2004), the European Community Human Potential Program under Contract HPRN-CT-2002-00282 (HYKE) and the IMATI (CNR - Pavia, Italy).

REFERENCES

1. F. Bassetti, E. Gabetta and E. Regazzini, On the depth of the trees in the McKean representation of Wild's sums. To appear in *Transport Theor. Stat. Phys.* (2006).
2. P. Billingsley, *Probability and Measure, Third edition* (J. Wiley & S., New York, 1995).
3. A. V. Bobylev, Exact solutions of the nonlinear Boltzmann equation and the theory of relaxation of a Maxwellian gas. *Teoret. Mat. Fiz.* **60**: 280–310 (1984).
4. E. A. Carlen, M. C. Carvalho and E. Gabetta, Central limit theorem for Maxwellian molecules and truncation of the Wild expansion. *Comm. Pure Appl. Math.* **53**: 370–397 (2000).
5. E. A. Carlen, M. C. Carvalho and E. Gabetta, On the relation between rates of relaxation and convergence of Wild sums for solutions of the Kac equation. *J. Funct. Anal.* **220**: 362–387 (2005).
6. E. A. Carlen and X. Lu, Fast and slow convergence to equilibrium for Maxwellian molecules via Wild sums. *J. Statist. Phys.* **112**: 59–134 (2003).
7. C. Cercignani, *Theory and Application of the Boltzmann Equation* (Elsevier, New York, 1975).
8. C. A. Charalambides, *Enumerative Combinatorics* (Chapman & Hall/CRC, Boca Raton, 2002).
9. L. Comtet, *Analyse Combinatoire*, Tome second (Presses universitaires de France, 1970).
10. P. Diaconis and L. Saloff-Coste, Bounds for Kac's Master Equation. *Comm. Math. Phys.* **209**: 729–755 (2000).
11. W. Feller, *An Introduction to Probability Theory and its Applications*, Vol I, 3rd edition (Wiley, New York, 1968).
12. M. Kac, Foundations of Kinetic Theory. *Proc. 3rd Berkeley Sympos.*, J. Neyman ed., **3**: 171–197 (1956).
13. M. Kac, *Probability and Related Topics in Physical Science* (Wiley Interscience, New York, 1959).
14. C. Knessl and W. Szpankowski, Asymptotic behavior of the height in a digital search tree and the longest phrase of the Lempel-Ziv scheme. *SIAM J. Comput.* **30**: 923–964 (2000).
15. C. Knessl and W. Szpankowski, The height of a binary search tree: The limiting distribution perspective. *Theoret. Comput. Sci.* **289**: 649–703 (2002).
16. W. Magnus, F. Oberhettinger and R. P. Soni, *Formulas and Theorems for the Special Functions of Mathematical Physics* (Springer, Berlin, 1966).
17. H. P. McKean Jr., Speed of approach to equilibrium for Kac's caricature of a Maxwellian gas. *Arch. Rational Mech. Anal.* **21**: 343–367 (1966).
18. H. P. McKean Jr., An exponential formula for solving Boltzmann's equation for a Maxwellian gas. *J. Combinatorial Theory* **2**: 358–382 (1967).
19. S. T. Rachev, *Probability Metrics and the Stability of Stochastic Models* (J. Wiley & S., Chichester, 1991).
20. B. Reed, The height of a random binary search tree. *J. ACM* **50**: 306–332 (2003).
21. F. G. Tricomi and A. Erdélyi, The asymptotic expansion of a ratio of Gamma functions. *Pacific J. Math.* **1**: 133–142 (1951).
22. E. Wild, On Boltzmann's equation in the kinetic theory of gases. *Proc. Cambridge Philos. Soc.* **47**: 602–609 (1951).