Topological classification of binary trees using the Horton-Strahler index

Zoltán Toroczkai
Theoretical Division and Center for Nonlinear Studies, Los Alamos National Laboratory, Los Alamos, New Mexico 87545
(Received 27 August 2001; published 20 December 2001)

The Horton-Strahler (HS) index $r=\max(i,j)+\delta_{i,j}$ has been shown to be relevant to a number of physical (such as diffusion limited aggregation) geological (river networks), biological (pulmonary arteries, blood vessels, various species of trees), and computational (use of registers) applications. Here we revisit the enumeration problem of the HS index on the rooted, unlabeled, plane binary set of trees, and enumerate the same index on the ambilateral set of rooted, plane binary set of trees of n leaves. The ambilateral set is a set of trees whose elements cannot be obtained from each other via an arbitrary number of reflections with respect to vertical axes passing through any of the nodes on the tree. For the unlabeled set we give an alternate derivation to the existing exact solution. Extending this technique for the ambilateral set, which is described by an infinite series of nonlinear functional equations, we are able to give a double exponentially converging approximant to the generating functions in a neighborhood of their convergence circle, and derive an explicit asymptotic form for the number of such trees.

DOI: 10.1103/PhysRevE.65.016130 PACS number(s): 02.50.−r, 02.30.Lt, 05.40.−a, 89.20.−a

I. INTRODUCTION

Trees are ubiquitous structures which appear naturally in a large number of physical, chemical, biological, and social phenomena, such as river networks, diffusion limited aggregation, pulmonary arteries, blood vessels and tree species, social organizations, decision structures, etc. They also play an important role in computer science (use of registers and computer languages), in graph theory, and in various methods of statistical physics such as cluster expansions and renormalization group.

In spite of their apparent structural simplicity, and the large body of scientific work on trees (a sample of which is found in Refs. [1–23], and references therein), they still offer challenges even related to the quantitative description of their topological structure. At the dawn of the science of complex networks [24], it is therefore rather important to have a complete understanding of all the tree structures and their properties.

A tree is defined as a set of points (vertices, nodes) connected with line segments (branches, or edges) such that there are no cycles or loops (a connected graph without cycles). For the simplest (unlabeled) rooted plane binary tree, each vertex has exactly three connecting branches, except for one vertex which is distinguished from all the others by having only two connecting branches coined as the root (R) of the tree, and a certain number of vertices with a single connecting branch called the ''leaves.'' The height of the tree is defined by the maximum number of levels starting from the root (which has height 0), and it can be calculated as the maximum number of branches one has to pass to reach the root from its vertices (since the leaves have only one branch, it means that this longest excursion must start from one of the leaves). The paths from the leaves to the root define a natural direction on the tree (similarly to the river flow) which is always towards the levels of lower height. A tree of height k we call *complete*, if it has 2^k leaves each being a distance k from the root.

Let us now mention three applications of the mathematics of trees which are directly connected to the so-called Horton-Strahler index of the tree, which is the subject of interest of the present paper.

Originally, the Horton-Strahler index of a binary tree was introduced in the studies of natural river networks by Horton [8] and later refined by Strahler [9], as a way of indexing real river topologies, since river networks are topologically similar to binary trees. By definition, a leaf has a rank of 0 (some authors associate the value of 1), and a vertex has a rank of $r=r(i,j)$ where $r(i,j)$ is the index function with i and j being the ranks of the two connecting vertices from the level above. When

$$r(i,j)=\max(i,j)+\delta_{i,j}, \tag{1}$$

the index is called the Horton-Strahler (HS) index. The quantity of particular interest is the HS index of the root which thus categorizes the topological complexity of the whole tree. Several other quantities can be introduced in relation to the HS index. A *segment* of order k [10], or a *stream* of order k [11] is a *maximal* path of branches connecting vertices of HS index k, ending in a vertex with higher index. Let $\mathcal{S}_k(n,T)$ denote the number of segments of order k of a tree T with n leaves, and $\langle L_k(n,T)\rangle$ is the average physical length of a segment of order k (the average $\langle\cdot\rangle$ is taken on the tree T). The *bifurcation ratios* $\mathcal{B}_k(n,T)$ are defined as $\mathcal{B}_k(n,T)=\mathcal{S}_k(n,T)/\mathcal{S}_{k+1}(n,T)$, and the length ratios via $\mathcal{L}_k(n,T)=\langle L_{k+1}(n,T)\rangle/\langle L_k(n,T)\rangle$. Horton and Strahler have empirically observed that for river networks both the $\mathcal{S}_k(n)$ and $\langle L_k(n)\rangle$ tend to approximate a geometric series, $\mathcal{B}_k(n)\approx\mathcal{B}$ with $3\leqslant\mathcal{B}\leqslant5$ and $\mathcal{L}_k(n)\approx\mathcal{L}$ with $1.5\leqslant\mathcal{L}\leqslant3$. Such networks are called *topologically self-similar* [12]. The notion of HS index is further refined by introducing the *biorder* (i,j) of a vertex, representing the HS indices of its two children [13,10,12], and then studying the ramification matrix, with elements related to the number of vertices with a given biorder.

Another interesting application of the mathematics of binary trees and the HS index, is in the description of the

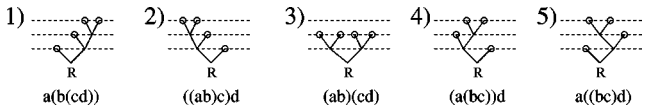

FIG. 1. The set of rooted, unlabeled binary plane trees corresponding to all the possible noncommutative, nonassociative bracketings of the four letter word $abcd$, $n=4$.

branched structure of diffusion-limited aggregates; see Ref. [12] and references therein. In this case the structures are grown on a substrate (which can be a point or a plane) by letting small particles diffuse towards the aggregate where they stick indefinitely at their point of first contact with the cluster. This creates complex and involved branched structures, whose topological complexity still remains a challenging problem to describe.

Finally, the last application we would like to mention is known as the *word bracketing problem* [4] which has obvious implications in computer science. Let us consider an alphabet of n letters, $\mathcal{A}=\{Y_1,Y_2,\ldots,Y_n\}$ and a word $S\equiv x_1x_2x_3\cdots x_{n-1}x_n$, $x_i\in\mathcal{A}$. A two-bracketing of the word S is a partition of its letters (by keeping their order) in groups of two units enclosed in brackets, where a unit can be a letter or a subpartition enclosed in brackets, such as $(x_1x_2)(x_3(x_4x_5))$, or $x_1(x_2(x_3(x_4x_5)))$, etc. The bracket (u_1u_2) between two units may be associated with a multiplicative composition law $(u_1u_2)=u_1\cdot u_2\in\mathcal{A}$. For example, let the alphabet $\mathcal{A}$ be all the positive integers, and the composition law be the regular multiplication of numbers. Then a bracketing of the multiple product S corresponds to one particular way of calculating S. A one-to-one correspondence can be made immediately to trees: let the letters $x_1,x_2,\ldots,x_n$ of the word S be associated with the leaves of a binary tree. To a particular bracketing of S it corresponds to a particular tree constructed by associating a lower level vertex to a bracketing (u_1u_2) (one may think of the brackets as representing the branches of the tree). The main question is how many ways are there to calculate such a product. If one assumes that the multiplication law is neither associative nor commutative, then the problem is referred to as the Catalan problem; see Ref. [4] for a number of solutions. The number of such bracketings is given by the Catalan numbers,

$$a_n=\frac{1}{m}\binom{2n-2}{n-1}.$$

The corresponding set of trees (see Fig. 1 for $n=4$) is in fact the set of rooted, unlabeled binary plane trees according to this bijection.

For later reference, we mention that the generating function $A(\xi)=\Sigma_{n=0}^{\infty}\xi^n a_n$ of the Catalan numbers obeys the equation $A^2-A+\xi=0$, with $A(0)=0$, so

$$A(\xi)=\frac{1}{2}(1-\sqrt{1-4\xi}). \tag{2}$$

The power series $A(\xi)$ converges within a disk of radius $a_c=1/4$.

The problem of enumerating trees becomes more difficult if the composition law is commutative, which was studied by Wedderburn and Etherington (WE) [5–7]. In the tree language, this means that two trees are considered identical if after a number of successive reflections with respect to the vertical axes passing through the vertices they can be transformed into each other and in this case they are said to be homeomorphic [3]. For the example shown in Fig. 1, there are only two such trees, since trees 1), 2), 4), and 5) can be transformed into each other. The trees that cannot be transformed into each other are called nonhomeomorphic. The set of nonhomeomorphic trees is called the set of *ambilateral* trees, [14,11]. Let the number of such trees with n leaves be denoted by w_n. The generating function (GF) defined as $W(\xi)=\Sigma_{n=0}^{\infty}\xi^n w_n$ obeys the nonlinear functional equation

$$W(\xi)=\xi+\frac{1}{2}W(\xi)^2+\frac{1}{2}W(\xi^2), \tag{3}$$

which has extensively been studied by Wedderburn [5]. Otter [3], studying a more general counting problem where the vertices can have at most m branches, comes to the conclusion that for the ambilateral trees, if n is large we have

$$w_n\sim cn^{-3/2}\gamma^n, \tag{4}$$

where $\gamma=2.483\,253\,5\cdots$. The method developed by Otter gives an iterative approach to c and γ. For example, γ is

$$\gamma=\lim_{n\to\infty}s_n^{2^{-n}}, \tag{5}$$

where $s_0=2$, $s_n=2+s_{n-1}^2$ so that for $n=4$ one already obtains an extremely close value of $\gamma\simeq(209\,091\,8)^{2^{-4}}$. Later, Bender developed a more general approach [15] deriving the same results. The coefficient c in Eq. (4) can also be computed: $c=0.318\,776\,62\cdots$.

The more practical application of the bracketing problem within computer science is the computation of arithmetic expressions by a computer. A general arithmetic expression involving only binary operators can simply be mapped onto a binary tree, called the syntax tree, which has as leaves the operands and the inner vertices the operators. A computer traverses this tree from the leaves towards the root and it uses registers to store the intermediate results. In general there are many ways of traversing such a tree, and the program that uses the *minimal* number of registers is the most efficient, or optimal, one. Ershov has shown [16] that the optimal code will use exactly as many registers to store the intermediate results as the HS index of the associated syntax tree.

In the present paper we investigate how the HS index is distributed on both the rooted, unlabeled, plane binary set of trees, and on the ambilateral set of binary trees. We first answer this question on the rooted, unlabeled, plane binary set, since it is simpler, but it will also provide us with a technique that can be extended to tackle the problem for the ambilateral set. For this set, the question was answered by Flajolet, Raoult, and Vuillemin [17], with a method somewhat similar to the one presented here. The enumeration

problem of the HS index on the ambilateral set is, however, inherently more difficult since it involves functional equations with nonlinear dependence in the argument similar to Eq. (3), and therefore an explicit solution in a closed form becomes impossible to attain. The derivation of an approximant formula for the number of ambilateral trees sharing the same HS index at the root is the main result of this paper.

The paper is organized as follows. First we present our derivation of the enumeration problem for the HS index on the unlabeled set in Sec. II, and then use this method of derivation from this case to develop a technique that can be used to attack the enumeration problem on the ambilateral set in the asymptotic limit, presented in Sec. III. Section IV is devoted to conclusions and outlook.

II. DISTRIBUTION OF THE HS INDEX ON THE UNLABELED SET

Let us observe that the root R of the tree has always two subtrees attached to it via the two branches, with k and $n-k$ leaves, respectively, $k=1,2,\ldots,n$. Let $N_n^{(r)}$ denote the number of unlabeled trees with n leaves that share the same HS index r at the root. A recursion is found for this number in the light of the observation above:

$$N_n^{(r)}=\sum_{k=1}^{n-1}\left\{N_k^{(r-1)}N_{n-k}^{(r-1)}+N_k^{(r)}\sum_{j=0}^{r-1}N_{n-k}^{(j)}+N_{n-k}^{(r)}\sum_{j=0}^{r-1}N_k^{(j)}\right\} \tag{6}$$

with the conventions $N_0^{(r)}\equiv 0$, $N_n^{(0)}\equiv\delta_{n,1}$, $N_1^{(r)}=\delta_{r,0}$. If the generating function for the variable n is defined as $D_r(\xi)=\Sigma_{n=0}^{\infty}\xi^n N_n^{(r)}$, then it obeys

$$D_r=D_{r-1}^2+2D_r\sum_{j=0}^{r-1}D_j,\quad r\geqslant 1,\quad D_0=\xi. \tag{7}$$

Next we give an exact solution to Eq. (7). Let us introduce the sum $B_r\equiv\Sigma_{j=0}^{r-1}D_j$, $r\geqslant 1$, $B_0=0$, $B_1=\xi$. Then $D_r=B_{r+1}-B_r$, and after rearranging the terms, Eq. (7) becomes $G_r=G_{r-1}$, where $G_r=B_r^2+B_{r+1}(1-2B_r)$. This means that $G_r=G_0=\xi$, i.e.,

$$B_r^2+B_{r+1}(1-2B_r)=\xi,\quad r\geqslant 0. \tag{8}$$

Note that the left hand side of Eq. (8) remains invariant to $B_r\mapsto 1-B_r$ which is another solution of Eq. (8). However, since in the case of the HS index $B_0=0$, this latter solution has to be dropped. If we make $2B_r\equiv 1-C_r$, Eq. (8) simplifies to $C_r^2-2C_rC_{r+1}=4\xi-1$, which after dividing both sides by $4\xi-1$, and introducing $Z_r\equiv C_r/\sqrt{4\xi-1}$, becomes

$$Z_{r+1}=\frac{Z_r^2-1}{2Z_r}. \tag{9}$$

Let us now write $Z_r=\cot(v_r)$, such that $v_0=\arctan(\sqrt{4\xi-1})$. Then Eq. (9) becomes $\cot(v_{r+1})=\cot(2v_r)$ which leads to $v_{r+1}=2v_r+\pi m$, $m\in\mathbb{Z}$, and which in turn is solved easily. Thus $Z_r=\cot(2^r v_0)$, so one finally obtains

$$D_r(\xi)=\frac{\sqrt{4\xi-1}}{2\sin(2^{r+1}\arctan\sqrt{4\xi-1})}. \tag{10}$$

Equation (10) is the exact solution to Eq. (7) in the complex ξ plane. On the real axis, within the radius of convergence a_c the above expression takes the form $D_r(\xi)=\sqrt{1-4\xi}/[2\sinh(2^{r+1}\arctan\sqrt{1-4\xi})]$, $\xi<a_c=1/4$. Since within the convergence disk one must have $\Sigma_{r=0}^{\infty}D_r(\xi)=A(\xi)$, we just obtained the identity [using Eq. (2)]

$$1+\sum_{r=1}^{\infty}\frac{1}{\sinh(2^r x)}=\coth(x),\quad x>0. \tag{11}$$

This identity can be checked to hold via more direct methods [25]. The singularities of $D_r(\xi)$ lie on the positive real axis at

$$\xi_k^{(r)}=\frac{1}{4\cos^2(k\pi/2^{r+1})},k=1,\ldots,2^r-1 \tag{12}$$

with an additional singularity at infinity (corresponding to $k=2^r$). We certainly have $\xi_k^{(r)}>a_c$. On the other hand, if one simply iterates Eq. (7) we obtain

$$D_r(\xi)=\frac{\xi^{2^r}}{2^r P_r(\xi)}, \tag{13}$$

where $P_r(\xi)$ is a polynomial in ξ of order 2^r-1: $P_1(\xi)=2^{-1}-\xi$, $P_2(\xi)=P_1(\xi)(2^{-1}-2\xi+\xi^2)$, $P_3(\xi)=P_2(\xi)(2^{-1}-4\xi+10\xi^2-8\xi^3+\xi^4)$, . . . , etc. One can find an explicit form for this polynomial from the general solution (10) if one invokes the identity [26] $\sin(nx)=n\sin x\cos x\Pi_{k=1}^{(n-2)/2}[1-\sin^2 x/\sin^2(k\pi/n)]$, so that Eq. (13) is recovered with: $P_r(\xi)=\Pi_{k=1}^{2^r-1}\cot^2(k\pi/2^{r+1})(\xi_k^{(r)}-\xi)$. It is easy to show, however, that $\Pi_{k=1}^{2^r-1}\cot^2(k\pi/2^{r+1})=1$, so the polynomial simplifies to

$$P_r(\xi)=\prod_{k=1}^{2^r-1}(\xi_k^{(r)}-\xi), \tag{14}$$

an expression valid on the whole complex ξ plane. Based on the explicit solution we obtained, one can give an exact form to the distribution of the HS index on the unlabeled set of trees, by inverting the generating function via

$$N_n^{(r)}=\frac{1}{2\pi i}\oint\frac{d\xi}{\xi^{n+1}}D_r(\xi)=\frac{1}{2\pi i}\oint\frac{d\xi}{\xi^{n+1}}\frac{\xi^{2^r}}{2^r P_r(\xi)}. \tag{15}$$

One can write

$$\frac{1}{2^r P_r(\xi)}=\sum_{j=1}^{2^r-1}\frac{A_j^{(r)}}{\xi_j^{(r)}-\xi}, \tag{16}$$

where $A_j^{(r)}=2^{-r}\Pi_{\substack{k=1\\k\neq j}}^{2^r-1}(\xi_k^{(r)}-\xi_j^{(r)})^{-1}$, $j=1,\ldots,2^r-1$. By Cauchy's theorem the integrals in Eq. (15) are readily performed, and one obtains

$$N_n^{(r)}=\begin{cases}\sum_{j=1}^{2^r-1} A_j^{(r)}[\xi_j^{(r)}]^{-(n-2^r+1)}, & n\geqslant 2^r\\ 0, & 0\leqslant n\leqslant 2^r-1.\end{cases} \tag{17}$$

From Eq. (16) it follows that $\Sigma_{j=1}^{2^r-1}A_j^{(r)}/\xi_j^{(r)}=2^{-r}/P_r(0)=1$. To obtain the last equality we used the form (14) and (12). Thus $N_{2^r}^{(r)}=1$, $r=1,2,\ldots$. The numbers $A_j^{(r)}$ can be calculated as follows. Observe that

$$A_j^{(r)}=\lim_{\xi\to\xi_j^{(r)}}\frac{\xi_j-\xi}{2^rP_r(\xi)}=-\frac{1}{2^rP_r'(\xi_j^{(r)})}, \tag{18}$$

where we used the L'Hôpital rule in the last equality. On the other hand, from Eqs. (13) and (10) it follows that $2^rP_r(\xi)=2\xi^{2^r}\sin(2^{r+1}\arctan\sqrt{4\xi-1})/\sqrt{4\xi-1}$. Taking the derivative of this equation at the point $\xi_j^{(r)}$, and inserting it in Eq. (18) it yields

$$A_j^{(r)}=(-1)^{j+1}\frac{4\xi_j^{(r)}-1}{2^{r+1}[\xi_j^{(r)}]^{2^r-1}}. \tag{19}$$

Thus we obtain from Eq. (17) the more explicit form,

$$N_n^{(r)}=\frac{1}{2^{r+1}}\sum_{j=1}^{2^r-1}(-1)^{j+1}\frac{4\xi_j^{(r)}-1}{[\xi_j^{(r)}]^n},\quad n\geqslant 2^r, \tag{20}$$

or using Eq. (12),

$$N_n^{(r)}=\frac{4^n}{2^{r+1}}\sum_{j=1}^{2^r-1}(-1)^{j+1}\sin^2\left(\frac{j\pi}{2^{r+1}}\right)\left[\cos\left(\frac{j\pi}{2^{r+1}}\right)\right]^{2n-2}, \tag{21}$$

an expression derived by Flajolet *et al.* [17]. Following this paper [17], our polynomials P_r can be simply connected to the Tchebycheff polynomial U [26], via the relation $P_r(\xi)=2^{-r}\xi^{2^r-1/2}U_{2^{r+1}-1}[1/(2\sqrt{\xi})]$.

If one employs the Poisson resummation formula for functions defined on a compact support (see Appendix B in Ref. [27]) on Eq. (21), an equivalent combinatorial expression can be derived in the form

$$N_{n+1}^{(r)}=\Sigma_{m=1}^{\infty}\nabla^2\left[\binom{2n}{n+k}\right](k)|_{k=1+(2m-1)2^r},$$

where $(\nabla f)(k)=f(k)-f(k-1)$ is the finite difference operator. For a different method, see Ref. [17].

Scaling limits. Next we briefly present the results of an asymptotic analysis on the $N_n^{(r)}$ numbers. Since $N_n^{(r)}$ is an enumeration result, it typically contains several scaling limits. In physical processes, during the growth of branched structures, usually only one of these limits is selected, and in frequent cases this limit has self-similar properties (such as for diffusion-limited aggregation (DLA), or for random generation of binary trees [18]). By definition, the family of trees that obey $\lim_{r\to\infty}(\ln[n(r)])/r=\text{const}\equiv\ln\mathcal{B}$ is called topologically self-similar [12], where $\mathcal{B}$ is the *bifurcation number*.

(i) $n\to\infty$ and r fixed. In this case the first term in Eq. (21) dominates the sum and the asymptotic behavior is given by $N_n^{(r)}\sim 2^{-r-1}\tan(\pi/2^{r+1})e^{n\ln[4\cos^2(\pi/2^{r+1})]}$. The rate of the exponential growth is a number between $\ln 2$ and $2\ln 2$.

(ii) $n\to\infty$, $r\to\infty$, $\sqrt{n}/2^r\to\infty$. Here the first term in Eq. (21) is still dominant (the rest being exponentially small corrections) and yields $N_n^{(r)}\sim\pi^2 2^{-3(r+1)}e^{n(2\ln 2-\pi^2/4^{r+1})}$. If $\sqrt{n}/2^r$ diverges with r slower than exponential, we have topological self-similarity with $\mathcal{B}=4$.

(iii) $n\to\infty$, $r\to\infty$, $\sqrt{n}/2^r\to d$, with some $0<d<\infty$. In this case the rest of the terms in Eq. (21) (after the first has been factored out) are of the type $j^2e^{-(j-1)\pi^2d^2}$ and the final expression is $N_n^{(r)}\sim A(d)4^n n^{-3/2}4^n$. The topological self-similarity is obvious with $\mathcal{B}=4$. The factor $A(d)$ is given by $A(d)=\pi^2e^{-\pi^2d^2}(1-e^{-\pi^2d^2})/(1+e^{-\pi^2d^2})^3$.

(iv) $n\to\infty$, $r\to\infty$, $\sqrt{n}/2^r\to 0$, and $n/2^r\to\infty$. In this case the analysis is performed easier from the combinatorial expression of $N_n^{(r)}$ and yields $N_n^{(r)}\sim\pi^{-1/2}n^{-5/2}e^{n2\ln 2-4^r/n}$.

III. DISTRIBUTION OF THE HS INDEX ON THE AMBILATERAL SET

Let us now analyze the same question on the set of ambilateral trees, and denote the number of ambilateral trees with n leaves and HS index r by $M_n^{(r)}$. We certainly must have the relation

$$\sum_{r=0}^{\infty}M_n^{(r)}=w_n. \tag{22}$$

Figure 2 gives the distribution of the HS index for n up to 32 and $r=2,3,4,5$. We can check easily that $M_n^{(0)}=\delta_{n,1}$, and $M_n^{(1)}=1-\delta_{n,1}$, so for simplicity these are not represented in the figure.

The numbers $M_n^{(r)}$ obey slightly more complicated recurrence relations since now the counting has to be done on a more restricted set. We must distinguish between odd and even n values. However, the two cases can be combined into one, if the convention $M_\nu^{(r)}=0$ for ν noninteger is adopted. The corresponding recurrence relation becomes

$$M_n^{(r)}=\sum_{0\leqslant k<j\leqslant n}\delta_{k+j,n}\left[M_k^{(r-1)}M_j^{(r-1)}+\sum_{s=0}^{r-1}(M_k^{(r)}M_j^{(s)}+M_k^{(s)}M_j^{(r)})\right]+M_{n/2}^{(r)}\sum_{s=0}^{r-1}M_{n/2}^{(s)}+\frac{1}{2}M_{n/2}^{(r-1)}(1+M_{n/2}^{(r-1)}). \tag{23}$$

n \ r	2	3	4	5	w_n
1					1
2					1
3					1
4	1				2
5	2				3
6	5				6
7	10				11
8	21	1			23
9	42	3			46
10	85	12			98
11	170	36			207
12	341	109			451
13	682	300			983
14	1365	813			2179
15	2730	2119			4850
16	5461	5442	1		10905
17	10922	13704	4		24631
18	21845	34142	23		56011
19	43690	84121	100		127912
20	87381	205741	424		293547
21	174762	499777	1617		676157
22	349525	1207954	5892		1563372
23	699050	2906779	20319		3626149
24	1398101	6970746	67531		8436379
25	2796202	16667409	216665		19680277
26	5592405	39757998	676214		46026618
27	11184810	94646973	2058825		107890609
28	22369621	224941421	6139668		253450711
29	44739242	533857098	17976046		596572387
30	89478485	1265538285	51801988		1406818759
31	178956970	2997087184	147192083		3323236238
32	357913941	7091960757	413083691	1	7862958391

FIG. 2. Particular values for the number of ambilateral trees with n leaves and HS index r. The shaded entries mean that there is no such tree.

The generating function $V_r(\xi)=\Sigma_{n=0}^{\infty}\xi^n M_n^{(r)}$ will thus obey

$$V_r(\xi)=\frac{1}{2}\,\frac{[V_{r-1}(\xi)]^2+V_{r-1}(\xi^2)}{1-\sum_{s=0}^{r-1}V_s(\xi)},\quad r\geqslant 1, \tag{24}$$

and $V_0(\xi)=\xi$. As a check for the correctness of Eq. (24), let us see if we recover the identity $\Sigma_{r=0}^{\infty}V_r(\xi)=W(\xi)$ [which follows from Eq. (22)]. Equation (24) is equivalent to $2V_r(\xi)-2\Sigma_{s=0}^{r-1}V_s(\xi)V_r(\xi)=[V_{r-1}(\xi)]^2+V_{r-1}(\xi^2)$. Introduce the temporary variable $G(\xi)=\Sigma_{r=0}^{\infty}V_r(\xi)$ and sum both sides of the equation over r, $r=1,2,\ldots,\infty$. One obtains $2(G(\xi)-\xi)-2\Sigma_{r=1}^{\infty}\Sigma_{s=0}^{r-1}V_s(\xi)V_r(\xi)=\Sigma_{r=0}^{\infty}[V_r(\xi)]^2+G(\xi^2)$. Using the identity $2\Sigma_{r=1}^{\infty}\Sigma_{s=0}^{r-1}V_sV_r=(\Sigma_{r=0}^{\infty}V_r)^2-\Sigma_{r=0}^{\infty}V_r^2$, one finds $G(\xi)=\xi+\frac{1}{2}[G(\xi)]^2+\frac{1}{2}G(\xi^2)$, which is precisely Eq. (3), showing that $G(\xi)=W(\xi)$, i.e., the relation $\Sigma_{r=0}^{\infty}V_r(\xi)=W(\xi)$ holds, indeed.

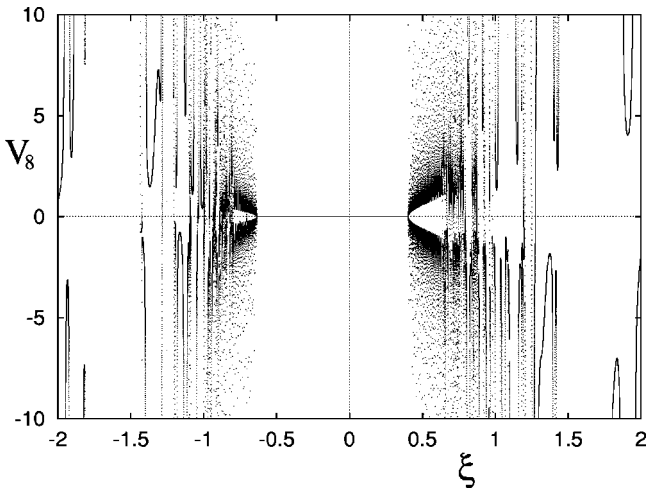

FIG. 3. The generating function $V_8(\xi)$ on the real axis. The function was evaluated in more than 1.3×10^5 points, and represented by dots.

In contrast to the previous case, the functional recurrence (24) cannot be treated in an exact analytical fashion due to the functional dependence on ξ^2. However, one can derive the asymptotic behavior and make statements that will lead to rather close approximations of the $M_n^{(r)}$ numbers. It is instructive to look at a few particular values, first:

$$V_1(\xi)=\frac{\xi^2}{1-\xi},$$

$$V_2(\xi)=\frac{\xi^4}{(1-2\xi)(1-\xi^2)}, \tag{25}$$

$$V_3(\xi)=\xi^8\,\frac{1-2\xi+\xi^2+2\xi^3-3\xi^4}{(1-2\xi)(1-2\xi^2)(1-\xi^4)(1-3\xi+4\xi^3-\xi^4)}.$$

Inverting $V_2(\xi)$, one obtains: $M_n^{(2)}=[2^{n-1}-3+(-1)^{n-4}]/6$, $n\geqslant 4$, which can be checked to hold; see Fig. 2. The result from the inversion of $V_3(\xi)$ is already so complicated that it is not worth presenting. As the index r increases, the polynomial expressions become more and more involved. Figure 3 shows the function $V_8(\xi)$ in the interval $[-2.0,2.0]$.

For every r, the power series for $V_r(\xi)$ has non-negative coefficients, $M_n^{(r)}\geqslant 0$. Based on a classic theorem of complex analysis, this means that on the circle of convergence, of radius $\alpha_r>0$, there will be a singularity of $V_r(\xi)$ at $\xi=\alpha_r$. Next we show that we have the ordering $0<\alpha_{r+1}<\alpha_r<1$ for $r\geqslant 2$, and the limit $\lim_{r\to\infty}\alpha_r$ exists and it is equal to $\alpha\equiv 1/\gamma=0.402\,697\,503\,6\cdots$. We shall use mathematical induction to prove the ordering. From the particular examples above it follows that $\alpha_2=0.5$, $\alpha_3=0.424\,507\cdots$.

Let us now assume that $\alpha_j<\alpha_{j-1}<1$ for all $j\leqslant r$, $j\geqslant 2$. We will show that $\alpha_{r+1}<\alpha_r$. Note that the radius of convergence for $V_j(\xi^2)$ is $\sqrt{\alpha_j}>\alpha_j$, if α_j is less than unity. By *reductio ad absurdum*, let us assume first, that $\alpha_{r+1}>\alpha_r$. This means, that $V_{r+1}(\xi)$ is analytic in α_r. From Eq. (24),

$$V_{r+1}(\xi)=\frac{V_r^2(\xi)+V_r(\xi^2)}{1-V_0(\xi)-\cdots-V_r(\xi)}. \tag{26}$$

By the argument above, $V_r(\xi^2)$ is analytic in α_r (its radius of convergence is $\sqrt{\alpha_r}>\alpha_r$, since by assumption $\alpha_r<\alpha_{r-1}<\cdots<\alpha_2=1/2<1$). In the denominator of Eq. (26), *all* functions V_j, $j=0,1,\ldots r-1$ are analytic in α_r, because by assumption they all have radii of convergence strictly larger than α_r. However, V_r is singular in α_r, and the singularities do not cancel in the numerator and denominator of Eq. (26), and thus V_{r+1} is singular in α_r, a contradiction. We are left to prove that $\alpha_{r+1}=\alpha_r$ cannot hold. Let us denote $B_r=\Sigma_{s=0}^{r}V_s$. Again, we assume that $\alpha_{r+1}=\alpha_r$ is true. It is easy to show that for any *finite* r, $|V_r(\alpha_r)|=\infty$. This means from the recurrence relation that

$$B_{r-1}(\alpha_r)=1 \tag{27}$$

[in the numerator of Eq. (24) we have only functions analytic at α_r]. Since $V_{r+1}(\xi)=[V_r^2(\xi)+V_r(\xi^2)]/[1-B_r(\xi)]$, from the assumption $\alpha_{r+1}=\alpha_r$ it would follow that the equation $B_r(x)=1$ *cannot* have any solutions (V_{r+1} is analytic within the circle of convergence) in the interval $0<x<\alpha_r$. [Note that in the interval $0<x<\alpha_r$, the numerator $V_r^2(\xi)+V_r(\xi^2)$ cannot be zero, since the power series V_r has only positive coefficients.] The equation $B_r(x)=1$ is equivalent to $B_{r-1}(x)+V_r(x)=1$. However, from Eq. (24) $1-B_{r-1}(x)=[V_r^2(x)+V_r(x^2)]/V_r(x)$. Thus the equation

$$V_r^2(x)=V_{r-1}^2(x)+V_{r-1}(x^2) \tag{28}$$

should have no solution in $0<x<\alpha_r$. If x is arbitrarily close to α_r, then $V_r^2(x)$ is arbitrarily large. However, since $\alpha_{r-1}>\alpha_r$, $V_{r-1}^2(x)$, and $V_{r-1}^2(x^2)$ are both bounded from above. Thus, for x sufficiently close to α_r, we have $V_r^2(x)>V_{r-1}^2(x)+V_{r-1}(x^2)$. On the other hand, the HS index of a tree T equals to the height of the largest, complete, balanced tree embedded in T. This means that $M_n^{(r)}=0$ for $n=0,1,2,\ldots,2^r-1$. Also, $M_{2^r}^{(r)}=1$. In other words, one must have $V_r(x)=x^{2^r}[1+\mathcal{O}(x)]$.

This means that $V_{r-1}^2(x)=x^{2^r}[1+\mathcal{O}(x)]$, $V_{r-1}(x^2)=x^{2^r}[1+\mathcal{O}(x^2)]$, and $V_r^2(x)=x^{2^r}x^{2^r}[1+\mathcal{O}(x)]$. Since $V_{r-1}^2(x)+V_{r-1}(x^2)=2x^{2^r}[1+\mathcal{O}(x)]$, there will always be an $x>0$, ($x<1$), sufficiently close to zero, such that $V_r^2(x)<V_{r-1}^2(x)+V_{r-1}(x^2)$. Therefore there must exist an $0<x<\alpha_r$, for which Eq. (28) holds, which is a contradiction. Thus we have proven that $0<\alpha_{r+1}<\alpha_r<1$, for all $r\geqslant 2$. As a matter of fact we have also shown that the radii of convergence satisfy

$$V_r^2(\alpha_{r+1})=V_{r-1}^2(\alpha_{r+1})+V_{r-1}(\alpha_{r+1}^2), \quad r\geqslant 1. \tag{29}$$

Since the series α_r is monotonically decreasing, and bounded from below, the limit $\alpha=\lim_{r\to\infty}\alpha_r$ exists.

We have shown that $\Sigma_{r=0}^{\infty}V_r(\xi)=W(\xi)$. Since the radius of convergence for the left hand side is the minimum of all the radii of the terms in the summation, i.e., α, it must equal to the radius of convergence for $W(\xi)$, which, as shown by Otter and Bender is $1/\gamma$, $\lim_{r\to\infty}\alpha_r=\alpha=1/\gamma=0.402\,697\,503\,6\cdots$. Taking the limit $r\to\infty$ in Eq. (27), we get

$$W(\alpha)=1 \tag{30}$$

[since by definition $B_r=\Sigma_{s=0}^{\infty}V_s$, so $\lim_{r\to\infty}B_r=W(\xi)$]. Or, using Eq. (3),

$$W(\alpha^2)=1-2\alpha, \tag{31}$$

an identity also shown by Bender. Equations (31) and (3) can simply be combined to give the iterative computation of α in the form already mentioned in the Introduction, as follows: If we make the temporary notation

$$U(\xi)\equiv[1-W(\xi)]/\sqrt{\xi}, \tag{32}$$

Eq. (3) takes the form

$$U(\xi^2)=2+U^2(\xi), \tag{33}$$

and Eq. (31) simplifies to

$$U(\alpha^2)=2. \tag{34}$$

Let $S(x)=2+x^2$. Then, from Eq. (33) $U(\xi^2)=S[U(\xi)]$, or $U(\xi)=S[U(\xi^{1/2})]=S\{S[U(\xi^{1/4})]\}=S\{\cdots S[U(\xi^{2^{-n}})]\cdots\}$, where there are a total of n compositions for the S function, n arbitrary. Let us now choose $\xi=\alpha^{2^{n+1}}$. This means $U(\alpha^{2^{n+1}})=S\{\cdots S[U(\alpha^2)]\cdots\}=S\{\cdots[S(2)]\cdots\}$, by virtue of Eq. (34). From Eq. (32), $U(\alpha^{2^{n+1}})=[1-W(\alpha^{2^{n+1}})]/\alpha^{2^n}$. We have shown previously that $\alpha<1$ (it is the limit of the monotonically decreasing series $\alpha_r<1$), therefore we have

$$\alpha=\lim_{n\to\infty}\left(\frac{1-W(\alpha^{2^{n+1}})}{s_n}\right)^{2^{-n}}=\lim_{n\to\infty}s_n^{-2^{-n}} \tag{35}$$

since $W(\alpha^{2^{n+1}})\to W(0)=0$, and where $s_n=S(s_{n-1})$, $s_0=2$, just as in the Introduction. The convergence is double exponential, very fast.

As in Sec. II, the asymptotic behavior of the $M_n^{(r)}$ numbers for relatively large n and r is governed by the innermost singularity of $V_r(\xi)$ on the real axis. The graph of V_8 shown in Fig. 3 suggests that the generating function is, in fact, well behaved in a certain interval to the right of the radius of convergence α_8; see also Fig. 4. The existence of this interval comes from the fact that the singularities of the term with nonlinear argument $V_{r-1}(\xi^2)$ in the numerator of Eq. (24) kick in only beyond the circle of convergence of $V_{r-1}(\xi^2)$, which is $\sqrt{\alpha_{r-1}}>\alpha_{r-1}$. Thus in the interval $\alpha_r<x<\sqrt{\alpha_{r-1}}$

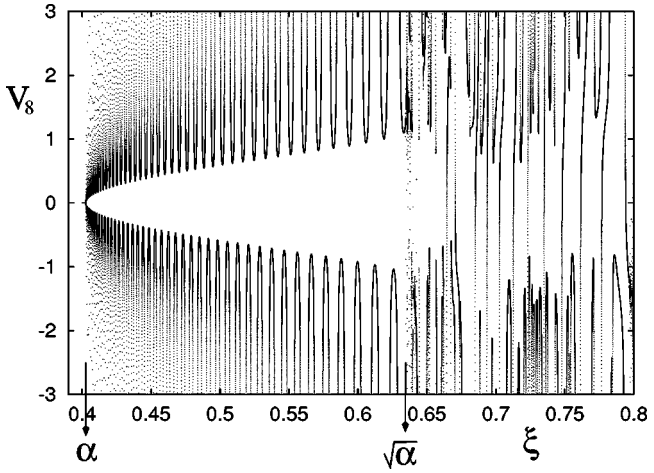

FIG. 4. A magnified region of Fig. 3. The arrows indicate the positions $\alpha=0.402\,69\cdots$ and $\sqrt{\alpha}=0.634\,58\cdots$ on the real axis.

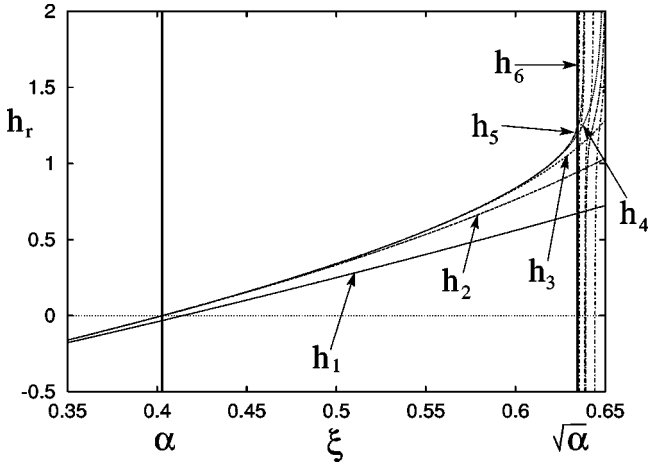

FIG. 5. The functions $h_r(\xi)$ are analytic on I. This figure shows $h_r(\xi)$ for $r=1,2,3,4,5,6$. The convergence on I to $h(\xi)$ is double exponentially fast. The thick vertical lines delimit the edges of the interval I. Close to α, the h_r functions cannot be distinguished on I for $r\geqslant 3$. To the right from I the h_r functions develop singularities. The point $\sqrt{\alpha}$ is a left accumulation point for the series of the left-most singularities of $h_r(\xi)$ as $r\to\infty$.

the term with the nonlinear argument is analytic, which ultimately is responsible for this nice behavior. Because $\alpha<\alpha_{r-1}$, for convenience we shall define the interval of this nice behavior to be $I=[\alpha,\sqrt{\alpha})$. In order to exploit this observation, we shall first rewrite the recurrence relation (24). Let us denote $H_r(\xi)=1-\Sigma_{s=0}^{r-1}V_s(\xi)$. With this notation, Eq. (24) takes the form $G_r(\xi)=G_{r-1}(\xi)$, $r\geqslant 1$, where $G_r(\xi)=H_r^2(\xi)-2H_r(\xi)H_{r+1}(\xi)+H_r(\xi^2)$. This leads to the new recurrence:

$$H_r^2(\xi)-2H_r(\xi)H_{r+1}(\xi)+H_r(\xi^2)=2\xi, \tag{36}$$

$H_0(\xi)=1$. This would be exactly solvable if it were not for the dependence on the nonlinear argument ξ^2. Note the resemblance to Eq. (8). Let $h_r(\xi)=2\xi-H_r(\xi^2)$, which is an analytic function in I. We also have $\Delta h_r(\xi)=h_r(\xi)-h_{r-1}(\xi)=V_r(\xi^2)=\xi^{2^r}[1+\mathcal{O}(\xi^2)]$, the latter equality being shown previously. This shows that in the interval I, the r dependence *weakens extremely fast*, double exponentially with increasing r. As a matter of fact, an upper estimate is

$$\Delta h_r(\xi)\leqslant\alpha^{2^{r-1}}. \tag{37}$$

In particular, $\Delta h_3(\xi)\leqslant 0.0263$, $\Delta h_4(\xi)\leqslant 0.000\,691\,6$, $\Delta h_5(\xi)\leqslant 4.79\times10^{-7}$, $\Delta h_6(\xi)\leqslant 2.28\times10^{-13}$, $\Delta h_7(\xi)\leqslant 5.22\times10^{-26}$, etc. Therefore, from the point of view of the asymptotic behavior, the h_r functions can be replaced by their asymptotic expression (as $r\to\infty$):

$$h(\xi)=W(\xi^2)+2\xi-1. \tag{38}$$

Figure 5 shows the functions h_r on the interval I for $r=1,2,3,4,5,6$.

Thus, instead of Eq. (36) we will consider

$$\overline{H}_r^2(\xi)-2\overline{H}_r(\xi)\overline{H}_{r+1}(\xi)=h(\xi). \tag{39}$$

The recurrence (39) in turn is easily solved in the way shown in Sec. I. The result is

$$\overline{H}_r(\xi)=\sqrt{h(\xi)}\cot\left[2^{r-r_0}\arctan\left(\frac{\sqrt{h(\xi)}}{\overline{H}_{r_0}(\xi)}\right)\right], \tag{40}$$

where r_0 for the moment is an arbitrary (positive integer) index. Recurrence (39) will become a good approximation to the recurrence (36) from an index r_0 on. The larger r_0 is the more accurate the approximation. Recurrence (39) is applied then with initial condition $\overline{H}_{r_0}(\xi)=H_{r_0}(\xi)$, which for modest r_0 values can be obtained by iterating Eq. (36) r_0 times.

What is the error we make when one replaces $h_{r_0}(\xi)$ with $h(\xi)$ on I? Summing the differences (37) from r_0+1 to infinity, one obtains the estimate $h(\xi)-h_{r_0}(\xi)\leqslant\alpha^{2^{r_0}}\Sigma_{m=0}^{\infty}\alpha^{2^{r_0}(2^m-1)}<\alpha^{2^{r_0}}/(1-\alpha^{2^{r_0}})$. Thus, for example, $h(\xi)-h_5(\xi)$ is smaller than 10^{-7}, $h(\xi)-h_6(\xi)$ is smaller than 10^{-13}, etc.

Therefore we can finally write on I:

$$V_r(\xi)\simeq\frac{\sqrt{W(\xi^2)+2\xi-1}}{\sin\{2^{r+1-r_0}\arctan[\sqrt{W(\xi^2)+2\xi-1}/H_{r_0}(\xi)]\}},$$

$$r>r_0,\quad \xi\in I. \tag{41}$$

In Fig. (6) we plot the right hand side (rhs) of Eq. (41) and the V_r function from iterating Eq. (24). Note that the approximation is very good, and it becomes virtually indistinguishable from the true function the closer ξ is to α. Larger r_0 values will also give better approximations, since the approximation is only applied from the r_0 index on. However, r_0 cannot be taken too high for approximation purposes, since it assumes that the exact expression of H_{r_0} (or V_{r_0}) is known. This makes only the modest r_0 values (less than 5) useful. On the other hand, expression (41) is very

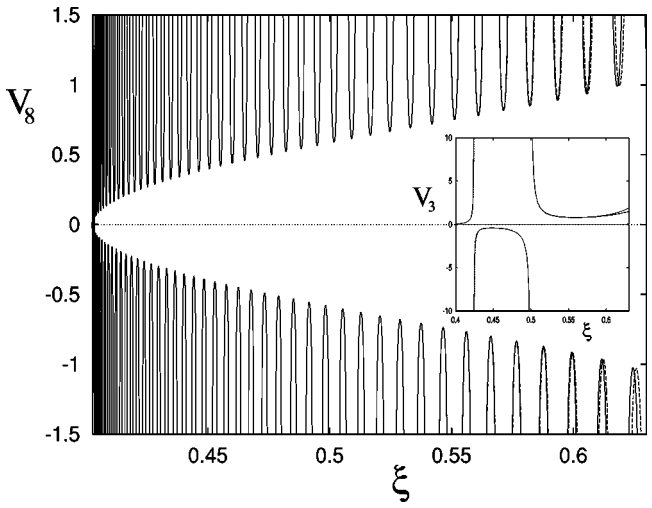

FIG. 6. The true $V_r(\xi)$ function (dashed line) from iterating Eq. (24), and the approximation in Eq. (41) (solid line) for $r=8$ with $r_0=3$, and $r=3$ with $r_0=2$ (the inset).

practical in analyzing the singularities of V and give rather close approximant expressions to these singularities. In particular, we see that within the interval I, Eq. (41) preserves the property that if $\alpha_{r'}$ is a singularity of $V_{r'}$ (or a zero of $H_{r'}$) then it is a singularity of V_r (or a zero of H_r), whenever $r>r'$. If one is interested in the asymptotic behavior, then a more tractable expression can be derived for the rhs of Eq. (41): the function $h(\xi)$ is analytic on the interval I, and since already for modest r values, the innermost singularity of V_r (denoted α_r) is extremely close to α, one can safely replace $h(\xi)$ in this neighborhood by $h(\xi)\simeq h'(\alpha)(\xi-\alpha)$.

This leads to the approximant

$$V_r(\xi)\simeq K_r(\xi)\equiv\frac{\mu\sqrt{\xi-\alpha}}{\sin[2^{r+1}\arctan(\theta\sqrt{\xi-\alpha})]},\quad \xi\in I \tag{42}$$

for sufficiently large r (here "large" means $r\geqslant 4$) where

$$\mu=\sqrt{h'(\alpha)},\quad \theta=\frac{\sqrt{h'(\alpha)}}{2^{r_0}H_{r_0}(\alpha)}. \tag{43}$$

Next, we compute $h'(\alpha)$. One can use a very similar method to the one employed to obtain Eq. (35), to give

$$h'(\alpha)=\lim_{n\to\infty}\frac{s_n}{s_0s_1\cdots s_{n-1}}=3.171\,055\,6\cdots, \tag{44}$$

so, $\mu=1.780\,745\,815\cdots$. If one computes θ for $r_0=3$, we have $H_3(\alpha)=(1-3\alpha+4\alpha^3-\alpha^4)/(1-2\alpha-\alpha^2+2\alpha^3)=0.164\,518\cdots$, and thus $\theta=1.353\,002\,2\cdots$. If we were to use $r_0=4$, then one would obtain $H_4(\alpha)=0.082\,262$, so $\theta=1.352\,952\,924\,5$ and slightly improve the approximation on θ. No significant improvement will be obtained with larger r_0 values. Figure 7 shows the agreement of the form given in Eq. (42). For clarity, we defined the function $f(z)$ given by

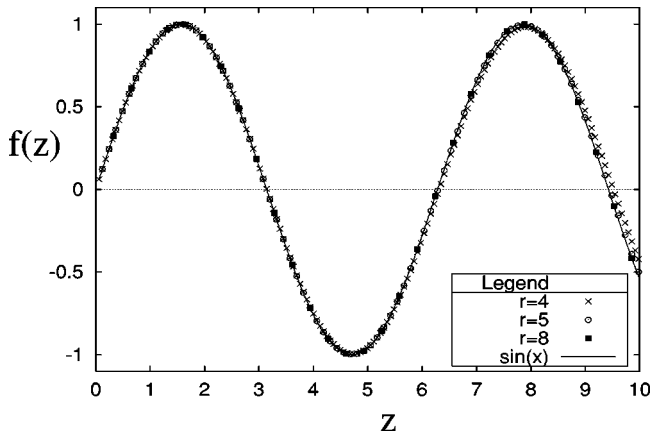

FIG. 7. The goodness of Eq. (42). For μ and θ we used the values derived in the text.

$$f(z)=\frac{\mu}{\theta}\frac{\tan\left(\dfrac{z}{2^{r+1}}\right)}{V_r\left[\alpha+\theta^{-2}\tan^2\left(\dfrac{z}{2^{r+1}}\right)\right]}. \tag{45}$$

Here we use the true V_r function using numerical iteration of Eq. (24), and evaluate it in the points $\xi=\alpha+\theta^{-2}\tan^2[z/(2^{r+1})]$. If the approximation (42) is good, then one should have $f(z)=\sin(z)$. As seen from Fig. 7, the approximation is already excellent for $r=4$ close to α (which corresponds to the $z=0$ point in this plot). The interval I in these transformed coordinates corresponds to $[0,2^{r+1}\arctan(\theta\sqrt{\sqrt{\alpha}-\alpha})]=(0,0.577\,435\,486\times 2^{r+1})$. There are no fitting parameters, we used for μ and θ the values derived above.

In order to obtain the approximation to the *number* $M_n^{(r)}$ of ambilateral trees with the same HS index at the root, we will have to invert Eq. (42). The singularities of the rhs of Eq. (42) are given by

$$\xi_k^{(r)}=\alpha+\theta^{-2}\tan^2\left(\frac{k\pi}{2^{r+1}}\right),k=1,2,3,\ldots,2^r-1 \tag{46}$$

[at the moment we do not care whether some of the singularities will fall outside the interval I, we just simply want to invert Eq. (42), and then at the end keep only those terms from the final expression that were generated by the singularities within I].

In a similar manner to the previous section, we first bring K_r to an inverted polynomial form:

$$K_r(\xi)=\frac{\mu[1+\theta^2(\xi-\alpha)]^{2^r}}{2^{r+1}\theta^{2^{r+1}-1}Q_r(\xi)}, \tag{47}$$

where Q_r is the polynomial: $Q_r(\xi)=\Pi_{k=1}^{2^r-1}(\xi_k^{(r)}-\xi)$. The case from the previous Sec. II corresponds to $\mu=1,\theta=2$, and $\alpha=1/4$. Thus if we denote by $\bar{M}_n^{(r)}$ the numbers coming from the inversion of $K_r(\xi)$, then

$$\bar{M}_n^{(r)}=\frac{\mu}{\theta^{2^{r+1}-1}}\frac{1}{2\pi i}\oint\frac{d\xi}{\xi^{n+1}}\frac{[1+\theta^2(\xi-\alpha)]2^r}{2^{r+1}Q_r(\xi)}. \quad (48)$$

We have

$$\frac{\mu}{2^{r+1}\theta^{2^{r+1}-1}Q_r(\xi)}=\sum_{j=1}^{2^r-1}\frac{A_j^{(r)}}{\xi_j^{(r)}-\xi}, \quad \text{with}$$

$$A_j^{(r)}=\frac{\mu}{2^{r+1}\theta^{2^{r+1}-1}}\prod_{\substack{k=1\\k\neq j}}^{2^r-1}\frac{1}{\xi_k^{(r)}-\xi_j^{(r)}}. \quad (49)$$

After performing the integrals, one obtains

$$\bar{M}_n^{(r)}=\sum_{j=1}^{2^r-1}A_j^{(r)}[\xi_j^{(r)}]^{-n-1}\sum_{m=0}^{\min\{n,2^r\}}\binom{2^r}{m}(1-\alpha\theta^2)^{2^r-m}[\theta^2\xi_j^{(r)}]^m. \quad (50)$$

This expression shows that the $\bar{M}_n^{(r)}$ may only *approximate* the $M_n^{(r)}$ numbers in a certain limit. This is seen from the fact that while one must have $M_n^{(r)}=0$ for $n<2^r$, and $M_{2^r}^{(r)}=1$, this is not respected by Eq. (50) (it would only be respected if $\alpha=\theta^{-2}$, however, this is not the case, and the reason behind this discrepancy is the neglected nonlinearity from the calculations). The limit, in which the approximation becomes good, is for r large (it means $r\geqslant 4$) and $n\gg 2^r$. In this case the sum over m can be performed, and one obtains

$$\bar{M}_n^{(r)}=\sum_{j=1}^{2^r-1}A_j^{(r)}[\xi_j^{(r)}]^{-n-1}[1+\theta^2(\xi_j^{(r)}-\alpha)]^{2^r}. \quad (51)$$

The $A_j^{(r)}$ numbers can be calculated in exactly the same way we did in the previous section. This leads to

$$A_j^{(r)}=(-1)^{j+1}\frac{\mu(\xi_j^{(r)}-\alpha)}{2^r\theta[1+\theta^2(\xi_j^{(r)}-\alpha)]^{2^r-1}}. \quad (52)$$

Inserting it into Eq. (51), it yields

$$\bar{M}_n^{(r)}=\frac{\mu}{2^r\theta}\sum_{j=1}^{2^r-1}(-1)^{j+1}\frac{[1+\theta^2(\xi_j^{(r)}-\alpha)](\xi_j^{(r)}-\alpha)}{[\xi_j^{(r)}]^{n+1}}. \quad (53)$$

As a check to the correctness of Eq. (53) we can take $\mu=1,\theta=2$, and $\alpha=1/4$ from the unlabeled case, to obtain Eq. (20). Equation (53) explicitly shows the contribution of each singularity. However, if we want to approximate the $M_n^{(r)}$ numbers, we should also account for the condition $\xi_j^{(r)}<\sqrt{\alpha}$. Using the expression (46), this leads to $j<J_r$, where

$$J_r=\frac{2^{r+1}}{\pi}\arctan(\theta\sqrt{\sqrt{\alpha}-\alpha})\simeq(0.183\,803\,525\,0\cdots)\times 2^{r+1}. \quad (54)$$

Thus, using again Eq. (46),

$$\bar{M}_n^{(r)}=\frac{\mu}{2^r\theta^3}\sum_{j=1}^{[J_r]}(-1)^{j+1}\frac{\tan^2\left(\frac{j\pi}{2^{r+1}}\right)\left[1+\tan^2\left(\frac{j\pi}{2^{r+1}}\right)\right]}{\left[\alpha+\theta^{-2}\tan^2\left(\frac{j\pi}{2^{r+1}}\right)\right]^{n+1}}. \quad (55)$$

When the asymptotic limit is generated by the innermost root $\alpha_r\simeq\xi_1^{(r)}$, i.e., by the first term in Eq. (55), one obtains for the topologically self-similar ambilateral trees, the scaling behavior

$$\bar{M}_n^{(r)}\sim\frac{2\mu\pi^2d^3}{\alpha\theta^3}e^{-\pi^2d^2/\alpha\theta^2}n^{-3/2}\gamma^n \quad (56)$$

and therefore $\mathcal{B}=\gamma=1/\alpha=2.483\,253\,5\cdots$.

Let us now see how well formula (55) approximates the $M_n^{(r)}$ numbers. To do this, we shall define the error $Q_n^{(r)}=[|\bar{M}_n^{(r)}-M_n^{(r)}|/M_n^{(r)}]\times 100\%$. For example, from the Table in Fig. 2, $M_{32}^4=413\,083\,691$. The formula above gives $\bar{M}_{32}^{(4)}=445\,781\,858$, and thus $Q_{32}^{(4)}=7.915\cdots\%$. Further error values: $Q_{100}^{(5)}=5.341\,32\ldots\%$, $Q_{800}^{(5)}=0.053\,91\cdots\%$, $Q_{800}^{(6)}=0.003\,551\cdots\%$.

IV. CONCLUSIONS AND OUTLOOK

Combinatorial enumeration of trees is typically difficult to solve when the set under enumeration obeys symmetry-exclusion principles, such as for the ambilateral case treated here. These symmetry-based constraints may arise in realistic situations and thus forces us to enumerate *classes* of subsets of trees. In the ambilateral case a class is defined as being formed by those binary trees that have the same number of leaves and HS index at the root and can be obtained one from another via successive reflections with respect to the nodes of the tree. Certainly, the symmetry operation defining the class must be an invariant transformation of the topological index (HS in our case). Another example of such symmetry-operation-generated class enumeration is the case of the ''leftist trees'' playing an important role in the representation of *priority queues*, shown by Crane [19], followed by Knuth [20], who gives their explicit definition. An elegant enumeration for the leftist trees, using generating function formalism, was only given very recently by Nogueira [21].

The existing solutions to such class enumerations on trees (such as ours and that of Flajolet *et al.* [17] and of Nogueira [21]) are obtained via methods tailored for the particularities

of the set and symmetry operation in question. It is desirable to have, however, at least on a formal level, a general encompassing theory of class enumerations of topological indices. In this direction, powerful methods such as that of the antilexicographic order method developed by Erdős and Székely [22] or the method of bijection to Schröder trees developed by Chen [23] may turn out to be effective after a suitable extension to include topological indices such as the Horthon-Strahler index. This, however, stands as an open problem.

ACKNOWLEDGMENTS

I am especially thankful to Eli Ben-Naim for introducing this problem to me, and for the many constructive suggestions while I was working on it. Useful discussions and comments from I. Benczik, T. Brown, W. Y. C. Chen, É. Czabarka, P. L. Erdős, M. Hastings, G. Istrate, and R. Mainieri are also gratefully acknowledged. This work was supported by the Department of Energy under Contract W-7405-ENG-36.

[1] A. Cayley, Collected Mathematical Papers, Cambridge, **3**, 242 (1889); **9**, 202 (1889); **11**, 365 (1889); **13**, 26 (1897).
[2] G. Pólya, Acta Math. **68**, 145 (1937).
[3] R. Otter, Ann. Math. **49**, 583 (1948).
[4] L. Comtet, *Advanced Combinatorics* (Reidel, Dordrecht, 1974).
[5] J.H.M. Wedderburn, Ann. Math. **24**, 121 (1922).
[6] I.H.M. Etherington, M. Gaz. **21**, 36 (1937).
[7] H. Harrary, G. Prins, Acta Math. **101**, 141 (1959).
[8] R.E. Horton, Bull. Geol. Soc. Am. **56**, 275 (1945).
[9] A.N. Strahler, Bull. Geol. Soc. Am. **63**, 1117 (1952); Trans., Am. Geophys. Union **38**, 913 (1957).
[10] J. Vannimenus and X.G. Viennot, J. Stat. Phys. **54**, 1529 (1989).
[11] J.W. Moon, Ann. Discr. Math. **8**, 117 (1980).
[12] T.C. Halsey, Europhys. Lett. **39**, 43 (1997); Phys. Today **53(11)**, 36 (2000).
[13] X.G. Viennot, in *Mélanges Offerts à M. P. Schutzenberger*, edited by D. Perrin and A. Lascoux (Hermes, Paris, 1989).
[14] J.S. Smart, Geol. Soc. Am. Bull. **80**, 1757 (1969).
[15] E.A. Bender, SIAM Rev. **16**, 485 (1974).
[16] A.P. Ershov, Commun. ACM **1**, 3 (1958).
[17] P. Flajolet, J.C. Raoult, and J. Vuillemin, Theor. Comput. Sci. **9**, 99 (1979).
[18] E. Ben-Naim, P.L. Krapivsky, and S.N. Majumdar, Phys. Rev. E **64**, 035101 (2001).
[19] C. A. Crane, Linear Lists and Priority Queues as Balanced Binary Trees, Tech. Rep. STAN-CS-72-259, Stanford University, 1972.
[20] D. Knuth, *The Art of Computer Programming* (Addison-Wesley, Reading, MA, 1973), Vol. 3.
[21] P. Nogueira, Discrete Appl. Math. **109**, 253 (2001).
[22] P.L. Erdős and L.A. Székely, Adv. Appl. Math. **10**, 488 (1989).
[23] W.Y.C. Chen, Proc. Natl. Acad. Sci. U.S.A. **87**, 9635 (1990).
[24] A.-L. Barabási, Phys. World **14(7)**, 33 (2001); S.H. Strogatz, Nature (London) **410**, 268 (2001).
[25] I. Benczik (private communication).
[26] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, edited by Alan Jeffrey (Academic Press, New York, 1994).
[27] Z. Toroczkai, G. Korniss, S. Das Sarma, and R.K.P. Zia, Phys. Rev. E **62**, 276 (2000).