On Han's Hook Length Formulas for Trees

William Y.C. Chen[1] Oliver X.Q. Gao[2] Peter L. Guo[3]

Center for Combinatorics, LPMC-TJKLC
Nankai University, Tianjin 300071, P.R. China

[1]chen@nankai.edu.cn, [2]oliver@cfc.nankai.edu.cn, [3]lguo@cfc.nankai.edu.cn

Submitted: Mar 22, 2011; Accepted: Jun 22, 2011; Published: Jul 29, 2011
Mathematics Subject Classifications: 05A15, 05A19

Abstract

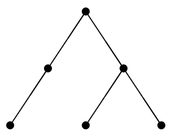
Recently, Han obtained two hook length formulas for binary trees and asked for combinatorial proofs. One of Han's formulas has been generalized to k-ary trees by Yang. Sagan has found a probabilistic proof of Yang's extension. We give combinatorial proofs of Yang's formula for k-ary trees and the other formula of Han for binary trees. Our bijections are based on the structure of k-ary trees associated with staircase labelings.

Keywords: hook length formula, k-ary tree, bijection, staircase labeling.

1 Introduction

Motivated by the hook length formula of Postnikov [6], Han [4] discovered two hook length formulas for binary trees. Han's proofs are based on recurrence relations. He raised the question of finding combinatorial proofs of these two formulas [3, 4]. Yang [9] generalized one of Han's formulas to k-ary trees by using generating functions. A probabilistic proof of Yang's formula has been found by Sagan [7]. By extending Han's expansion technique to k-ary trees, Chen, Gao and Guo [1] gave another proof for Yang's formula. The objective of this paper is to give combinatorial proofs of Yang's formula for k-ary trees and the other formula of Han for binary trees.

Recall that a k-ary tree is a rooted unlabeled tree where each vertex has exactly k subtrees in linear order, where we allow a subtree to be empty. When $k = 2$ (resp., $k = 3$), a k-ary tree is called a binary (resp., ternary) tree. A complete k-ary tree is a k-ary tree for which each internal vertex has exactly k nonempty subtrees. The hook length of a vertex u in a k-ary tree T, denoted by h_u, is the number of vertices of the subtree rooted at u. The hook length multi-set $\mathcal{H}(T)$ of T is defined to be the multi-set of hook lengths of all vertices of T. For example, Figure 1 gives an illustration of the hook length multi-set of a binary tree.



T $\quad \mathcal{H}(T) = \{1, 1, 1, 2, 3, 6\}$

Figure 1: The multi-set of hook lengths of a binary tree.

Let B_n be the set of all binary trees with n vertices. Han [4] discovered the following formulas. He also gave derivations of these formulas in [3] by using the expansion technique.

Theorem 1.1 (Han [4]) *For each positive integer n, we have*

$$\sum_{T \in B_n} \frac{1}{\prod_{h \in \mathcal{H}(T)} h 2^{h-1}} = \frac{1}{n!} \tag{1.1}$$

and

$$\sum_{T \in B_n} \frac{1}{\prod_{h \in \mathcal{H}(T)} (2h+1) 2^{2h-1}} = \frac{1}{(2n+1)!}. \tag{1.2}$$

As pointed out by Han [4], the above two formulas have a special feature that the hook lengths appear as exponents. Yang [9] extended the above formula (1.1) to k-ary trees.

Theorem 1.2 (Yang [9]) *For any positive integers n and k, we have*

$$\sum_{T} \prod_{h \in \mathcal{H}(T)} \frac{1}{h k^{h-1}} = \frac{1}{n!}, \tag{1.3}$$

where the sum ranges over k-ary trees with n vertices.

To give a combinatorial proof of (1.3), we shall define a set $S(n, k)$ of *staircase arrays* on $[k] = \{1, 2, \ldots, k\}$. More precisely, we shall represent an array in $S(n, k)$ in the form $(C_0, C_1, \ldots, C_{n-1})$, where $C_0 = \emptyset$ and for $1 \leqslant i \leqslant n-1$, C_i is a vector of length i with each entry in $[k]$.

We introduce the notion of staircase labelings of a k-ary tree, and we show that the sequences in $S(n, k)$ are in one-to-one correspondence with k-ary trees with n vertices associated with staircase labelings. This leads to a bijective proof of formula (1.3). Based on this bijection, we also obtain a combinatorial interpretation of formula (1.2).

2 A combinatorial proof of (1.3)

Our combinatorial proof of Yang's formula (1.3) is based on the following reformulation

$$\sum_T \frac{n!k^{1+2+\cdots+n}}{\prod_{h\in\mathcal{H}(T)} hk^h} = k^{1+2+\cdots+(n-1)}. \tag{2.1}$$

It is clear that the right-hand side of (2.1) equals the number of sequences in $S(n,k)$. As will be seen, the left hand-side of (2.1) equals the number of k-ary trees with n vertices associated with staircase labelings. We shall give a bijection between $S(n,k)$ and the set of k-ary trees with n vertices associated with staircase labelings.

More precisely, a staircase labeling of a k-ary tree is defined as follows. For a k-ary tree T with n vertices, we use a set $\{C_0, C_1, \ldots, C_{n-1}\}$ of vectors on $[k]$ to label the vertices of T, where C_i contains i elements in $[k]$. Moreover, we impose the following restrictions: for any vertex u with label C_i and a descent (not necessarily a child) v with label C_j, we have $i < j$, that is, the labels on any path from the root to a leaf have increasing lengths; and the $(i+1)$-st entry of C_j is determined by the relative position of the child of u on the path from u to v among its siblings. To be more specific, if the r-th child of u is on the path from u to v, then the $(i+1)$-st entry of C_j is set to be r.

For example, Figure 2 gives a staircase labeling of a ternary tree. For the label of any vertex, the entries that are determined by the labels of its ancestors are written in boldface.

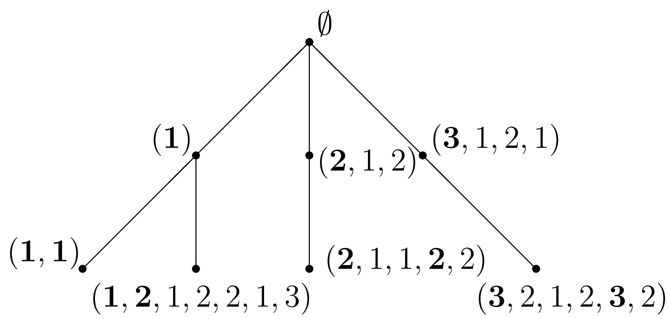


Figure 2: A staircase labeling of a ternary tree

Let $I(n,k)$ denote the set of k-ary trees with n vertices associated with staircase labelings. The following lemma shows that $|I(n,k)|$ is equal to the left-hand side of (2.1).

Lemma 2.1 *For* $n \geqslant 1$,

$$|I(n,k)| = \sum_T \frac{n!k^{1+2+\cdots+n}}{\prod_{h\in\mathcal{H}(T)} hk^h}, \tag{2.2}$$

where the sum ranges over k*-ary trees with* n *vertices.*

Proof. Let $P \in I(n,k)$ be a k-ary tree with a staircase labeling. Suppose that the labels of P are $C_0, C_1, \ldots, C_{n-1}$, where C_i is a vector of length i. Define Q to be the k-ary tree obtained from P by replacing a label C_i with i. Clearly, Q is an increasing k-ary tree in the sense that the label of any internal vertex is smaller than the labels of its children.

We shall consider the question of determining the number of k-ary trees P with n vertices associated with staircase labelings that correspond to a given increasing k-ary tree Q. Clearly, P and Q have the same underlying k-ary tree, denoted by T. In other words, we shall compute the number of staircase labelings of a k-ary tree T with given label length for each vertex. For any vertex u of T, let f_u denote the number of vertices on the path from the root to u. We claim that there are

$$k^{1+\cdots+n-\sum_{u\in T} f_u} \tag{2.3}$$

staircase labelings of T such that a vertex with label i in Q is associated with a vector C_i of length i. To prove (2.3), let u_i be the vertex of Q with label i. Recalling the definition of a staircase labeling, we need to determine how many entries in C_i that are determined by the ancestors of u_i. It can be seen that there are $f_{u_i} - 1$ entries of C_i that are determined by the ancestors of u_i. The other entries can be any element in $[k]$. Hence there are $k^{i+1-f_{u_i}}$ choices for C_i. This implies (2.3).

Note that the number in (2.3) does not depend on the specific increasing labeling of the k-ary tree T. To compute the number of staircase labelings of a k-ary tree T, it suffices to determine the number of increasing labelings of T. It is known that the number of increasing labelings of T equals

$$\frac{n!}{\prod_{h\in\mathcal{H}(T)} h},$$

see Knuth [5] or Gessel and Seo [2]. So we deduce that

$$|I(n,k)| = \sum_T \frac{n!}{\prod_{h\in\mathcal{H}(T)} h} k^{1+\cdots+n-\sum_{u\in T} f_u}, \tag{2.4}$$

where T ranges over k-ary trees with n vertices.

To obtain formula (2.2), we need to establish the following relation

$$\sum_{u\in T} h_u = \sum_{u\in T} f_u. \tag{2.5}$$

This can be justified by observing that both sides of (2.5) count the number of ordered pairs (u,v), where v is a descendant of u in T under the assumption that u is a descendant of itself. Substituting (2.5) into (2.4), we arrive at (2.2). This completes the proof. ∎

We have the following correspondence.

Theorem 2.1 *There is a bijection between $S(n,k)$ and $I(n,k)$.*

Proof. The map φ from $I(n,k)$ to $S(n,k)$ is straightforward, that is, for $P \in I(n,k)$ with a labeling set $\{C_0, C_1, \ldots, C_{n-1}\}$, define

$$\varphi(P) = (C_0, C_1, \ldots, C_{n-1}).$$

We proceed to give the inverse map ϕ from $S(n,k)$ to $I(n,k)$. Given a sequence $(C_0, C_1, \ldots, C_{n-1})$ in $S(n,k)$, we aim to construct a k-ary tree with n vertices associated with a staircase labeling by using the labels $C_0, C_1, \ldots, C_{n-1}$.

The map ϕ can be described as a recursive procedure. Let v_0 be a vertex with label $C_0 = \emptyset$. Clearly, v_0 and its label C_0 form a k-ary tree with a staircase labeling. Let $C_1 = (c_1)$. Adding a vertex v_1 as the c_1-th child of v_0 and assigning the label C_1 to v_1, we get a k-ary tree labeled by C_0 and C_1, denoted by P_1. It can be easily checked that P_1 is a k-ary tree with a staircase labeling. Assume that P_{m-1} ($m \geqslant 2$) is a k-ary tree with a staircase labeling with vertices $v_0, v_1, \ldots, v_{m-1}$ such that for $0 \leqslant i \leqslant m-1$, the vertex v_i has label C_i. Now we construct a k-ary tree with a staircase labeling, denoted by P_m, by adding the vertex v_m to P_{m-1} and assigning the label C_m to v_m.

To determine the position of v_m, we start with the root v_0. Let $C_m = (c_1, c_2, \ldots, c_m)$. If the c_1-th subtree of v_0 is empty, then we add the vertex v_m to P_{m-1} as the c_1-th child of v_0. Otherwise, we arrive at the c_1-th child of v_0, denoted by v_j. Note that the label of v_j is C_j. If the c_{j+1}-th subtree of v_j is empty, then we add the vertex v_m to P_{m-1} as the c_{j+1}-th child of v_j. If the c_{j+1}-th subtree of v_j is not empty, then we arrive at the c_{j+1}-th child of v_j. Repeating this process, we get a k-ary tree P_m labeled by $C_0, C_1, \ldots, C_m$. It is clear that P_m is a k-ary tree with a staircase labeling.

Thus, we obtain a k-ary tree $\phi(C_0, C_1, \ldots, C_{n-1}) = P_{n-1}$, labeled by $C_0, C_1, \ldots, C_{n-1}$. It can be checked that the maps φ and ϕ are inverses of each other. This completes the proof. ∎

In particular, for $k = 2$, the proof of Theorem 2.1 reduces to a combinatorial proof of Han's formula (1.1) for binary trees. Figure 3 gives an illustration of the bijection ϕ for $n = 6$, $k = 2$ and

$$(C_0, C_1, \ldots, C_5) = (\emptyset, (2), (2,1), (1,2,2), (1,2,2,1), (2,2,1,1,2)) \in S(6,2).$$

3 A combinatorial interpretation of (1.2)

In this section, we apply the bijection ϕ constructed in the previous section to give a combinatorial interpretation of formula (1.2). To this end, we reformulate (1.2) in terms of complete binary trees.

Clearly, one can add $n+1$ leaves to a binary tree with n vertices to form a complete binary tree with $2n+1$ vertices. Moreover, a vertex u with hook length h_u in a binary tree becomes an internal vertex with hook length $2h_u + 1$ in the corresponding complete binary tree. Denote by B^c_{2n+1} the set of complete binary trees with $2n+1$ vertices. Then

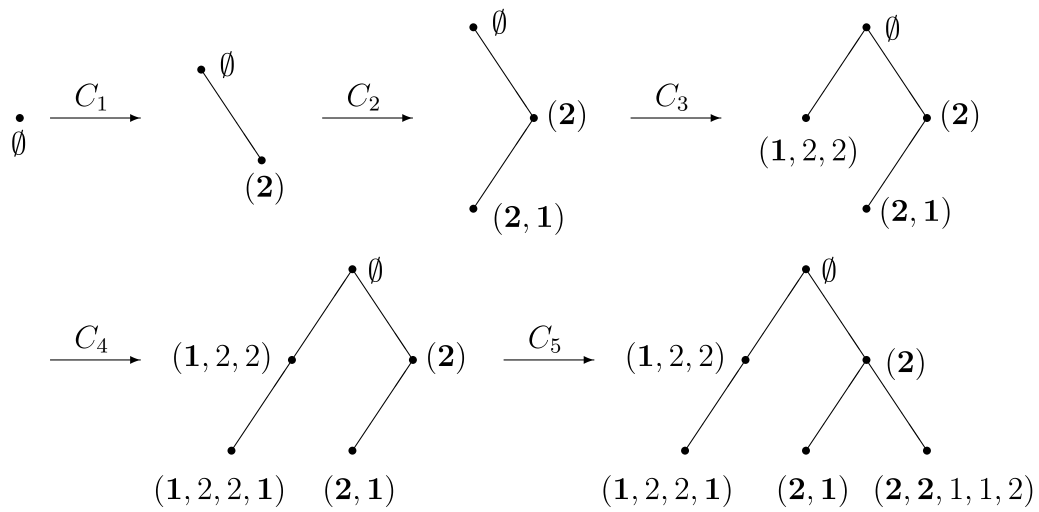


Figure 3: An illustration of the bijection ϕ.

(1.2) is equivalent to the following formula

$$\sum_{T\in B_{2n+1}^c} \frac{1}{\prod_{h\in\mathcal{H}(T)} h2^{h-1}} = \frac{1}{2^n(2n+1)!}. \tag{3.1}$$

In fact, our combinatorial interpretation of (3.1) is based on the following form

$$\sum_{T\in B_{2n+1}^c} \frac{(2n+1)!2^{1+2+\cdots+(2n+1)}}{\prod_{u\in\mathcal{H}(T)} h2^h} = \frac{2^{1+2+\cdots+2n}}{2^n}. \tag{3.2}$$

Combinatorial proof of (3.2). By the argument in the proof of Lemma 2.1, we see that the left-hand side of (3.2) is equal to the number of complete binary trees with $2n+1$ vertices associated with staircase labelings. Let $S'(2n+1,2)$ be the set of sequences in $S(2n+1,2)$ corresponding to complete binary trees with staircase labelings under the bijection ϕ. By the construction of ϕ, we shall give an explanation of the following relation

$$|S'(2n+1,2)| = \frac{1}{2^n}|S(2n+1,2)|. \tag{3.3}$$

Since $|S(2n+1,2)| = 2^{1+2+\cdots+2n}$, we are led to a combinatorial proof of (3.2).

It remains to prove (3.3). To this end, we shall construct a sequence of subsets $M_0, M_1, \ldots, M_n$ of $S(2n+1,2)$ such that

$$S(2n+1,2) = M_0 \supset M_1 \supset \cdots \supset M_n = S'(2n+1,2),$$

and for $1 \leqslant i \leqslant n$,

$$|M_i| = \frac{1}{2}|M_{i-1}|.$$

Let us begin with the definition of the subset M_1 of M_0. Let $(C_0, C_1, \ldots, C_{2n})$ be a sequence in M_0, and let T be the corresponding binary tree with a staircase labeling under the bijection ϕ. If both subtrees of the root of T have an odd number of vertices, then we choose this sequence $(C_0, C_1, \ldots, C_{2n})$ to be in M_1.

We proceed to prove the following relation

$$|M_1| = \frac{1}{2}|M_0|. \tag{3.4}$$

Let $(C_0, C_1, \ldots, C_{2n})$ be a sequence in M_1. Denote by T the corresponding binary tree with a staircase labeling under the bijection ϕ. Assume that for $1 \leqslant i \leqslant 2n$, s_i is the first entry of the vector C_i. By the construction of ϕ, if $s_i = 1$ (resp., $s_i = 2$), then there is a vertex with label C_i in the left (resp., right) subtree of the root of T. Since both subtrees of the root of T have an odd number of vertices, there is an odd number of 1's (or, equivalently, 2's) among $s_1, s_2, \ldots, s_{2n}$. Consider the set $\{1,2\}^{2n}$ of vectors of length $2n$ with entries in $\{1,2\}$. It is clear that there are as many vectors in $\{1,2\}^{2n}$ with an odd number of 1's as vectors in $\{1,2\}^{2n}$ with an even number of 1's. This implies that $|M_1| = |M_0 \backslash M_1|$, and hence we obtain (3.4).

In general, we can define the subset M_{j+1} of M_j for $j \geqslant 1$. Let $(C_0, C_1, \ldots, C_{2n})$ be a sequence in M_j, and let T be the corresponding binary tree with a staircase labeling under the bijection ϕ. Suppose that the vertices of T are $v_0, v_1, \ldots, v_{2n}$ such that the vertex v_i is labeled by C_i. Let $v_{t_0}, v_{t_1}, v_{t_2}, \ldots$ be the internal vertices of T such that the indices are arranged in increasing order, that is, $t_0 < t_1 < t_2 < \cdots$. If both subtrees of v_{t_j} have an odd number of vertices, then this sequence $(C_0, C_1, \ldots, C_{2n})$ is put in M_{j+1}. Using the same argument as that for (3.4), we deduce that

$$|M_{j+1}| = \frac{1}{2}|M_j|.$$

Let $(C_0, C_1, \ldots, C_{2n})$ be a sequence in M_n, and let T be the corresponding binary tree associated with a staircase labeling under the bijection ϕ. It can be seen that T is a binary tree with a staircase labeling such that both subtrees of any internal vertex have an odd number of vertices. It follows that T is a complete binary tree with a staircase labeling, which implies that $M_n = S'(2n+1, 2)$. This completes the proof. ∎

Acknowledgments. We wish to thank the referee for helpful suggestions. This work was supported by the 973 Project, the PCSIRT Project of the Ministry of Education, and the National Science Foundation of China.

References

[1] W.Y.C. Chen, O.X.Q. Gao and P.L. Guo, Hook length formulas for trees by Han's expansion, Electron. J. Combin. 16 (1) (2009), R62.

[2] I.M. Gessel and S. Seo, A Refinement of Cayley's formula for trees, Electron. J. Combin. 11 (2) (2006), R27.

[3] G.-N. Han, Discovering new hook length formulas by expansion technique, Electron. J. Combin. 15 (1) (2008), R133.

[4] G.-N Han, New hook length formulas for binary trees, Combinatorica 30 (2) (2010), 253-256.

[5] D.E. Knuth, The Art of Computer Programming, Vol. 3, Sorting and Searching, 2nd ed., Addison Wesley Longman, 1998.

[6] A. Postnikov, Permutohedra, associahedra, and beyond, Int. Math. Res. Notices (2009), 1026-1106.

[7] B.E. Sagan, Probabilistic proofs of hook length formulas involving trees, Sém. Lothar. Combin., Special issue dedicated to the memory of Pierre Leroux, 61A (2009), Art. B61Aa.

[8] S. Seo, A combinatorial proof of Postnikov's identity and a generalized enumeration of labeled trees, Electron. J. Combin. 11 (2) (2006), N3.

[9] L.L.M. Yang, Generalizations of Han's hook length identities, arXiv:math.CO/0805.0109.