On Han’s Hook Length Formulas for Trees

William Y.C. Chen\(^1\), Oliver X.Q. Gao\(^2\) and 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

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 with staircase labelings.

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

AMS Classification: 05A15, 05A19

1 Introduction

Motivated by the hook length formula of Postnikov \(^5\), 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 \(^8\) 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 \(^6\). 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 \(^\square\) gives an illustration of the hook length multi-set 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.
Figure 1: The multi-set of hook lengths of a binary tree.

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

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

and

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

As pointed out by Han [4], the above two formulas have a special feature that the hook lengths appear as exponents. Yang [8] extended the above formula (1.1) to $k$-ary trees.

Theorem 1.2 (Yang [8]) For any positive integers $n$ and $k$, we have

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

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 $1 \leq i \leq n - 1$, $C_i$ is a vector of length $i$ with each entry in $[k]$.

We shall show that the sequences in $S(n, k)$ are in one-to-one correspondence to $k$-ary trees with $n$ vertices whose labels form a staircase sequence in $S(n, k)$. Such $k$-ary trees are called $k$-ary trees with staircase labelings. This leads to a bijective proof of formula (1.3). Based on this bijection, we further 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

$$\sum_{T} \frac{n! k^{1+2+\ldots+n}}{\prod_{h \in \mathcal{H}(T)} h k^h} = k^{1+2+\ldots+(n-1)}.$$
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 whose labels form a staircase array. Such a $k$-ary tree is called a $k$-ary tree with a staircase labeling. Furthermore, we shall give a one-to-one correspondence 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: The labels are vectors on $[k]$ with distinct lengths. In other words, the labels can be written as $C_0, C_1, \ldots, C_{n-1}$, where $C_i$ is a vector on $[k]$ of length $i$. 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 is set to be $r$.

For example, Figure 2 gives a staircase labeling of a ternary tree, where 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](image)

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 $n \geq 1$,

$$|I(n, k)| = \sum_{T} \frac{n! k^{1+2+\ldots+n}}{\prod_{h \in H(T)} h^k},$$

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 \( P \) in \( I(n, k) \) 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

\[
k^{1 + \cdots + n - \sum_{u \in T} f_u}
\]

(2.3)

staircase labelings of \( T \) such that the label corresponding to a vertex with label \( i \) in \( Q \) has length \( i \). To prove (2.3), let \( u_i \) be the vertex of \( Q \) with label \( i \). Assume that \( P \) is a \( k \)-ary tree with a staircase labeling for which the vertex \( u_i \) has label \( C_i \). Notice that \( C_i \) is of length \( i \). 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 \( f_u - 1 \) entries of \( C_i \) that are determined by the ancestors of \( u_i \). The other entries can be any choice of the elements in \( [k] \). Consequently, there are \( k^{i+1-f_u} \) 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 \( k \)-ary trees with staircase labelings, it suffices to determine the number of increasing labelings of \( T \). It is known that the number of increasing labelings of \( T \) equals

\[
\frac{n!}{\prod_{h \in H(T)} h}
\]

see, for example, Gessel and Seo [2]. So we deduce that

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

(2.4)

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

We now need to establish the following relation

\[
\sum_{u \in T} h_u = \sum_{u \in T} f_u.
\]

(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 \( T \) and we adopt 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 2.2** There is a bijection between \( S(n, k) \) and \( I(n, k) \).
Proof. The map $\varphi$ 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 now proceed to give the reverse map $\phi$ 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 $\{C_0, C_1, \ldots, C_{n-1}\}$.

The map $\phi$ 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 $P_{m-1}$ ($m \geq 2$) is a $k$-ary tree with a staircase labeling with vertices $v_0, v_1, \ldots, v_{m-1}$ such that for $0 \leq i \leq m-1$ 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 at the root $v_0$. Let $C_m = (c_1, c_2, \ldots, c_m)$. If the $c_1$-th child 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 $v_{i_0}$. Consider the label $C_{i_0}$ of $v_{i_0}$. If the $c_{i_0+1}$-th child of $v_{i_0}$ is empty, then we add the vertex $v_m$ to $P_{m-1}$ as the $c_{i_0+1}$-th child of $v_{i_0}$. Otherwise, we arrive at the $c_{i_0+1}$-th child of $v_{i_0}$. Repeating this process, we finally get a $k$-ary tree $P_m$ with labeled by $C_0, C_1, \ldots, C_m$. It is clear that $P_m$ is a $k$-ary tree with a staircase labeling.

By the above procedure, 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 $\varphi$ and $\phi$ are inverses of each other. This completes the proof. \[ \blacksquare \]

In particular, for $k = 2$, the proof of Theorem 2.2 reduces to a combinatorial proof of Han’s formula (1.1) for binary trees. Figure 3 gives an illustration of the bijection $\phi$ 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 shall apply the bijection $\phi$ constructed in the previous section to give a combinatorial interpretation of formula (1.2). To this end, we need to 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
Figure 3: An illustration of the bijection \( \phi \).

The tree becomes an internal vertex with hook length \( 2h_u + 1 \) in the corresponding complete binary tree. Denote by \( B_{2n+1} \) the set of complete binary trees with \( 2n+1 \) vertices. Then (1.2) is equivalent to the relation

\[
\sum_{T \in B_{2n+1}} \frac{1}{\prod_{h \in H(T)} h2^{h-1}} = \frac{1}{2^n(2n+1)!}.
\]  

(3.1)

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

\[
\sum_{T \in B_{2n+1}} \frac{(2n+1)!2^{1+2+\cdots+(2n+1)}}{\prod_{u \in H(T)} h2^h} = \frac{2^{1+2+\cdots+2n}}{2^n}.
\]  

(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 staircase labelings of complete binary trees with \( 2n+1 \) vertices. Let \( S'(2n+1, 2) \) the set of sequences in \( S(2n+1, 2) \) corresponding to complete binary trees with staircase labelings under the bijection \( \phi \). By the construction of \( \phi \), we shall give an explanation of the fact that

\[
|S'(2n+1, 2)| = \frac{1}{2^n} |S(2n+1, 2)|.
\]  

(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 \( M_0, M_1, \ldots, M_n \) such that

\[ S(2n+1, 2) = M_0 \supset M_1 \supset \cdots \supset M_n = S'(2n+1, 2), \]
and for $1 \leq i \leq n$,

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

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 under the bijection $\phi$. If both subtrees of the root of $T$ (labeled with $C_0$) have an odd number of vertices, then we choose this sequence $(C_0, C_1, \ldots, C_{2n})$ to be in $M_1$. By the construction of $\phi$, it can be easily seen that $(C_0, C_1, \ldots, C_{2n})$ belongs to $M_1$ if and only if there is an odd number of 1’s among $s_1, s_2, \ldots, s_{2n}$, where $s_i$ is the first entry of $C_i$. Since for any set of an even number of elements, there are as many subsets with an odd number of elements as subsets with an even number of elements, we deduce that

$$|M_1| = \frac{1}{2} |M_0|. \quad (3.4)$$

Similarly, we can define the subset $M_2$ of $M_1$. Let $(C_0, C_1, \ldots, C_{2n})$ be a sequence in $M_1$, and let $T$ be the corresponding binary tree under the bijection $\phi$. Suppose that the vertices of $T$ are $v_0, v_1, \ldots, v_{2n}$ and a vertex $v_i$ is labeled by $C_i$. Clearly, $v_0$ is the root of $T$ labeled by $C_0$. Suppose that $v_j$ is an non-root internal vertex with $j$ being minimum. It is clear that $v_j$ is a child of $v_0$. If both subtrees of $v_j$ have an odd number of vertices, then we choose this sequence $(C_0, C_1, \ldots, C_{2n})$ to be in $M_2$. Using the above argument for (3.4), we obtain that

$$|M_2| = \frac{1}{2} |M_1|.$$  

In general, we can define the subset $M_{j+1}$ of $M_j$ for $j \geq 1$. Let $(C_0, C_1, \ldots, C_{2n})$ be a sequence in $M_j$, and let $T$ be the corresponding binary tree under the bijection $\phi$. Suppose that the vertices of $T$ are $v_{t_0}, v_{t_1}, \ldots, v_{t_{2n}}$ such that a 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 defined to be in $M_{j+1}$. By the above reasoning, we see that

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

Once the subset $M_n$ has been determined, we get a binary tree with a staircase labeling such that both subtrees of any internal vertex have an odd number of vertices. In other words, we obtain a complete binary tree with a staircase labeling. It follows that $M_n = S'(2n + 1, 2)$. This completes the proof. 

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

[6] 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.

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

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