Introduction to Hook Length Formula

Mehtaab Sawhney*

July 22, 2016

Abstract

USAMO 2016 Problem 2 provided controversy on the nature of Hook Length Formula (HLF) and whether it was "elementary". The attached note gives a variant of a proof by Bandlow [1] regarding HLF and has been streamlined so that it requires a mere two limits and shows a possible route of how a contestant could have proved the formula even on contest day. (Although this would arguably have been more remarkable than solving all problems on the test itself.)

1 Introduction

For the sake of brevity we will only briefly define the terms necessary for the proof. A Young Tableaux is a grid of left-justified squares such that the rows have a weakly decreasing number of squares as one goes down. Suppose there are n squares in the Young Tableau, then the Hook Length Formula counts the number of permutations of 1 through n in the grid so that the numbers increase as one goes from left to right on the rows and up to down on the columns. (Also known as Standard Young Tableaux.)

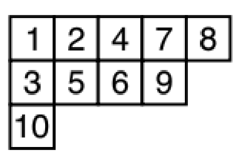

Figure 1: Young Tableaux (Taken from https://www.wikipedia.org/)

Finally the type of a Young Tableaux is a list of the number of squares in each row. This above diagram has type $(5, 4, 1)$. Note that the type corresponding directly to a partition of n, and we denote the shape as λ. The key thing to note is that in any Standard Young Tableaux, n is at the end of one of the rows such

*Student at the University of Pennsylvania

that the row below is **strictly** shorter than the row above. Finally the hook length in "Hook Length Formula" is the sum of the number of squares strictly below or to right of a given square plus 1. In particular the box filled with a 2 in the above figure has hook length 5 while the square filled in with 6 has hook length 2.

2 Elementary Proof of Hook Length's Formula

Bandlow [1] gives an elementary proof of Hook Length Formula, and this note mimics that proof with a less technical presentation.

Theorem 1 *Let the number of Standard Young Tableaux with n squares and shape λ, be Q_λ. Then*

$$Q_\lambda = \frac{n!}{\prod_{t\in\lambda} L(t)}$$

with $L(t)$ denoting the hook length of squares t in the λ.

Proof: The proof is by induction on n. For $n = 1$ the statement is trivial. Suppose that the theorem holds for all $n < k$, and now we prove it for k. Suppose λ has the type $(b_1 + \ldots + b_\ell, \ldots, b_1 + \ldots + b_\ell, b_1 + \ldots + b_{\ell-1}, \ldots, b_1 + \ldots + b_{\ell-1}, \ldots, b_1, \ldots, b_1)$ where there are a_i rows with length $b_1 + \ldots + b_i$. Let the s^{th} block be the a_s rows of length $b_1 + \ldots + b_s$. The only possible position for k is the bottom-right corner of one of these blocks. Note that after removing one of these corners then the remainder of the squares also form a Standard Young Tableaux. Finally let λ_s be λ without the bottom-right corner of the s^{th} block. Note that the hook length of only the squares in the same row or column as the removed square are affected. Then it follows that

$$R_s = \frac{\prod_{t\in\lambda} L(t)}{\prod_{t\in\lambda_s} L(t)}$$

$$= \left(\frac{a_1 + \ldots + a_{s-1} + b_1 + \ldots + b_s}{a_1 + \ldots + a_{s-1} + b_2 + \ldots + b_s}\right)\left(\frac{a_2 + \ldots + a_s + b_2 + \ldots + b_s}{a_2 + \ldots + a_{s-1} + b_3 + \ldots + b_s}\right)\cdots\left(\frac{a_s + b_{s-1} + b_s}{a_{s-1} + b_s}\right)a_s b_s$$

$$\times\left(\frac{a_s + a_{s+1} + b_{s+1}}{a_s + b_{s+1}}\right)\left(\frac{a_s + a_{s+1} + a_{s+2} + b_{s+1} + b_{s+2}}{a_s + a_{s+1} + b_{s+1} + b_{s+2}}\right)\left(\frac{a_s + \cdots + a_\ell + b_{s+1} + \ldots + b_s}{a_s + \ldots + a_{\ell-1} + b_{s+1} + \ell + b_\ell}\right)$$

by straightforward computation. The inductive step is equivalent to

$$X_{(\{a_i,b_i\},\ell)} = \sum_{s=1}^{\ell} R_s = k = \sum_{i=1}^{\ell} a_i(b_1 + \ldots + b_i).$$

This identity will now be proved by induction on ℓ. Note that $\ell = 1$ is trivial and then the key is to use

$$\left(\frac{a_1 + \ldots + a_s + b_1 + \ldots + b_s}{a_1 + \ldots + a_{s-1} + b_2 + \ldots + b_s}\right) = 1 + \frac{b_1}{a_1 + \ldots + a_{s-1} + b_2 + \ldots + b_s}.$$

for $s \geq 2$. Now it follows that

$$X_{(\{a_i,b_i\},\ell)} - X_{(\{a_i,b_i\}/\{a_1,b_1\},\ell-1)}$$
$$= b_1 a_1 \left(\frac{a_1+a_2+b_2}{b_2+a_1}\right) \ldots \left(\frac{b_2+\ldots+b_\ell+a_1+\ldots+a_\ell}{b_2+\ldots+b_\ell+a_1+\ldots+a_{\ell-1}}\right)$$
$$+ b_1 a_2 \left(\frac{b_2}{a_1+b_2}\right)\left(\frac{a_2+a_3+b_3}{a_2+b_3}\right) \ldots \left(\frac{a_2+\ldots+a_\ell+b_3+\ldots+b_\ell}{a_2+\ldots+a_{\ell-1}+b_3+\ldots+b_\ell}\right)$$
$$\vdots$$
$$+ b_1 a_\ell \left(\frac{a_2+\ldots+a_{\ell-1}+b_2+\ldots+b_\ell}{a_1+\ldots+a_{\ell-1}+b_2+\ldots+b_\ell}\right) \ldots \left(\frac{b_\ell}{a_{\ell-1}+b_\ell}\right).$$

Thus to prove the identity by induction it suffices to show that

$$X_{(\{a_i,b_i\},\ell)} - X_{(\{a_i,b_i\}/\{a_1,b_1\},\ell-1)} = b_1(a_1+\ldots+a_\ell).$$

Now denote $s_1 = 0$ and $s_t = a_1 + \ldots + a_{t-1} + b_2 + \ldots + b_t$ for $t \geq 2$. The inductive step is equivalent to

$$\sum_{i=1}^{\ell} a_i = \sum_{i=1}^{\ell} a_i \prod_{1 \leq j \neq k \leq n}^{\ell} 1 + \frac{a_i}{s_j - s_i}.$$

This is a well-known identity from [2] and the following proof is almost identical to the one given in the paper.

Lemma 1 *For $c \notin \{b_1 \ldots, b_n\}$ then*

$$\prod_{k=1}^{n} \frac{x+a_k-b_k}{x-b_k} = \prod_{k=1}^{n} \frac{c+a_k-b_k}{c-b_k} + \sum_{k=1}^{n} \frac{a_k(x-c)}{(b_k-c)(x-b_k)} \prod_{1 \leq j \neq k \leq n} \frac{b_k+a_j-b_j}{b_k-b_j}.$$

PROOF: Let

$$f(x) = \prod_{k=1}^{n} (x+a_k-b_k)$$

and

$$w(x) = (x-c) \prod_{k=1}^{n} (x-b_k).$$

Using Lagrange Interpolation of $f(x)$ at $c, b_1, \ldots, b_k$ then

$$f(x) = f(c) \prod_{k=1}^{n} \frac{x-b_k}{c-b_k} + \sum_{k=1}^{n} \frac{f(b_k)w(x)}{(x-b_k)w'(b_k)}$$

and using

$$f(b_k) = a_k \prod_{1 \leq j \neq k \leq n} (b_k+a_j-b_j)$$

it follows that

$$\prod_{k=1}^{n} \frac{x+a_k-b_k}{x-b_k} = \prod_{k=1}^{n} \frac{c+a_k-b_k}{c-b_k} + \sum_{k=1}^{n} \frac{a_k(x-c)}{(b_k-c)(x-b_k)} \prod_{1\leq j\neq k\leq n} \frac{b_k+a_j-b_j}{b_k-b_j}.$$

Therefore the lemma follows. The final trick is to take a pair of limits. Taking the limit as $c \to \infty$ it follows that

$$\prod_{k=1}^{n} \left(1+\frac{a_k}{x-b_k}\right) = 1 + \sum_{k=1}^{n} \frac{a_k}{x-b_k} \prod_{1\leq j\neq k\leq n} \left(1+\frac{a_j}{b_k-b_j}\right).$$

Multiplying by x it follows that

$$x\left(\prod_{k=1}^{n} \left(1+\frac{a_k}{x-b_k}\right) - 1\right) = \sum_{k=1}^{n} \frac{a_k x}{x-b_k} \prod_{1\leq j\neq k\leq n} 1+\frac{a_j}{b_k-b_j}.$$

Finally taking the limit as $x \to \infty$ then

$$\sum_{i=1}^{n} a_k = \sum_{i=1}^{n} a_k \prod_{1\leq j\neq k\leq n} 1+\frac{a_j}{b_k-b_j}.$$

Shifting the index n to ℓ and taking the sequencs $\{a_i\}$ and $\{s_i\}$ then

$$\sum_{i=1}^{\ell} a_i = \sum_{i=1}^{\ell} a_i \prod_{1\leq j\neq k\leq n}^{\ell} 1+\frac{a_i}{s_j-s_i}$$

as desired. Therefore the inductive step is complete and the result follows by induction. ■

References

[1] Jason Bandlow, *An elementary proof of the hook formula*, Electron. J. Combin **15** (2008), no. 1.

[2] R William Gosper, Mourad EH Ismail, Ruiming Zhang, et al., *On some strange summation formulas*, Illinois Journal of Mathematics **37** (1993), no. 2, 240–277.