A CAYLEY-TYPE IDENTITY FOR TREES

RAN J. TESSLER

Abstract. We prove a weighted generalization of the formula for the number of plane vertex-labeled trees.

1. Introduction

It is well known that the number of vertex-labeled trees on \( n \) vertices is \( n^{n-2} \). The formula was discovered by Carl Wilhelm Borchardt in 1860 [Borchardt(1860)] and was extended by Cayley in [Cayley(1889)]. Since then many proofs of this formula were given, and many extensions were found. A beautiful well-known extension is the following weighted Cayley formula.

Theorem 1.1. Let \( T_n \) be the set of vertex-labeled trees with \( n \) vertices labeled by \([n] = \{1, \ldots, n\}\). Associate a variable \( x_i \) to every \( i \in [n] \), and associate the monomial \( \prod_{i \in [n]} x_i^{d_T(i)} \) to \( T \in T_n \), where \( d_T(i) \) is the degree of \( i \) in \( T \). Then

\[
\sum_{T \in T_n} \prod_{i \in [n]} x_i^{d_T(i)} = \prod_{i=1}^{n} x_i \left( \sum_{i=1}^{n} x_i \right)^{n-2}.
\]

An identity closely related to Cayley’s formula is the formula, due to Leroux and Miloudi, [Leroux and Miloudi(1992)] (see also [Callan(2014)] for a short proof) which says that for \( n \geq 2 \) there are \( \binom{2n-3}{n-1} \) vertex-labeled plane trees on \( n \) vertices. By a plane tree we mean an abstract tree enriched with cyclic orders for the edges which emanate from each vertex.

In this note we prove a ”weighted version” for this formula, namely

Theorem 1.2. Associate a variable \( x_i \) to every \( i \in [n] \), and for integers \( m \geq 1 \) and \( a \) denote by \( \binom{x+m+a}{x+a} \) the polynomial

\[
\prod_{i=a+1}^{a+m} (x+i),
\]

we extend the definition to \( m = 0 \) by writing \( \binom{x+a}{x+a} = 1 \). For \( n \geq 2 \) it holds that

\[
\sum_{T \in T_n} \prod_{i \in [n]} \left( \frac{x_i + d_T(i) - 1}{x_i - 1} \right) = \prod_{i=1}^{n} x_i \left( \sum_{i \in [n]} x_i + 2n - 3 \right) \left( \sum_{i \in [n]} x_i + n - 1 \right).
\]

For example, for \( n = 2 \) the left and right hand sides of the formula give \( x_1x_2 \). For \( n = 3 \) the formula gives

\[
x_1x_2x_3(x_1 + x_2 + x_3 + 3).
\]

Dividing both sides by \( \prod x_i \) and substituting \( x_1 = \ldots = x_n = 0 \) gives precisely the Leroux-Miloudi formula. The weighted Leroux-Miloudi formula was used in [Luria and Tessler(2016)] to calculate and prove the threshold for the appearance of spanning 2–spheres in the Linial-Meshulam model for random 2–complexes.

1.1. Acknowledgements. R.T. is supported by Dr. Max Rössler, the Walter Haefner Foundation and the ETH Zürich Foundation.
2. Proof of the formula

When \( n = 2 \) the formula trivially holds (in fact, correctly interpreted, the formula extends to \( n = 1 \)). Our proof will be inductive. By dividing both sides by \( \prod x_i \) the theorem is seen to be equivalent to proving, for \( n \geq 2 \),

\[
\sum_{T \in T_n} \prod_{i \in [n]} \left( x_i + d_T(i) - 1 \right) = \left( \sum_{i \in [n]} x_i + 2n - 3 \right) \cdot \left( \sum_{i \in [n]} x_i + n - 1 \right).
\]

Substitute \( y_i = x_i + 1 \). Then translates to

\[
L_n(y_1, \ldots, y_n) := \sum_{T \in T_n} \prod_{i \in [n]} \left( y_i + d_T(i) - 2 \right) = \left( \sum_{i \in [n]} y_i + n - 3 \right) \cdot \left( \sum_{i \in [n]} y_i - 1 \right) =: R_n(y_1, \ldots, y_n).
\]

Both the left hand and right hand side are polynomials of degree \( n - 2 \) in \( n \) variables. Thus, any monomial does not contain at least one of the variables. Hence, (2) will follow from proving that for each \( i = 1, \ldots, n \)

\[
L_n(y_1, \ldots, y_n)|_{y_i=0} = R_n(y_1, \ldots, y_n)|_{y_i=0}.
\]

Since \( L_n, R_n \) are in addition symmetric, it is enough to prove (3) for \( i = n \). As

\[
\left( \sum_{i=1}^n y_i \right)|_{y_n=0} = \sum_{i=1}^{n-1} y_i
\]

we have

\[
R_n|_{y_n=0} = (n - 3 + \sum_{i=1}^{n-1} y_i)R_{n-1}.
\]

The induction will therefore follow if we could show that

\[
L_n|_{y_n=0} = (n - 3 + \sum_{i=1}^{n-1} y_i)L_{n-1}.
\]

Denote by \( w(T) \) the summand in (3) which corresponds to the tree \( T \),

\[
w(T) = \prod_{i \in [n]} \left( y_i + d_T(i) - 2 \right).
\]

Observe that if \( y_n = 0 \) then \( w(T) = 0 \) whenever \( d_T(n) > 1 \). Thus,

\[
L_n(y_1, \ldots, y_n)|_{y_n=0} = \sum_{T \in T_n'} \prod_{i \in [n]} \left( y_i + d_T(i) - 2 \right),
\]

where \( T_n' \subseteq T_n \) is the collection of trees in which \( n \) is a leaf. For \( t \in T_n' \) let \( a(t) \in [n-1] \) be the single neighbour of \( n \) and let \( t(T) \in T_{n-1} \) be the tree obtained from erasing the vertex \( n \). It hold that

\[
w(T) = (y_{a(t)} + d_T(a(t)) - 2)w(t(T)) = (y_{a(t)} + d_{t(T)}(a(t)) - 1)w(t(T)).
\]

Since the sum of degrees of vertices in a graph is twice the number of edges, and the number of edges in a tree on \( m \) vertices is \( m-1 \) (3) yields, for any \( T \in T_{n-1} \),

\[
\sum_{T' \in t^{-1}(T)} w(T') = \sum_{a \in [n-1]} (y_a + d_T(a) - 1)w(T) = \left( \sum_{a=1}^{n-1} y_a + n - 3 \right)w(T) = L_{n-1}
\]

Putting (3) and the definition of \( L_n \) together

\[
L_n|_{y_n=0} = \sum_{T \in T_n'} \prod_{i \in [n]} \left( y_i + d_T(i) - 2 \right) = \left( \sum_{a=1}^{n-1} y_a + n - 3 \right) \sum_{T \in T_{n-1}} w(T) = L_{n-1}
\]

which is precisely (3).
References

[Borchardt(1860)] C. B. Borchardt. ber eine Interpolationsformel fr eine Art Symmetrischer Functionen und ber Deren Anwendung. Math. Abh. der Akademie der Wissenschaften zu Berlin, pages 1–20, 1860.

[Callan(2014)] D. Callan. A quick count of plane (or planar embedded) labeled trees. 2014. URL https://oeis.org/A006963/a006963_1.pdf

[Cayley(1889)] A. Cayley. A theorem on trees. Quart. J. Pure Appl. Math., 23:376–378, 1889.

[Leroux and Miloudi(1992)] P. Leroux and B. Miloudi. Généralisations de la formule d’Otter. Ann. Sci. Math. Québec, 16 (1):53–80, 1992.

[Luria and Tessler(2016)] Z. Luria and R. J. Tessler. A Sharp Threshold for Spanning 2-Spheres in Random 2-Complexes. 2016.

Institute for Theoretical Studies, ETH Zürich