Enumeration of Weighted Plane Trees

Alexander K. Zvonkin*

April 21, 2014

Abstract

In weighted trees, all edges are endowed with positive integral weight. We enumerate weighted bicolored plane trees according to their weight and number of edges.

1 Preliminaries

This note is not intended for a journal publication: it does not contain difficult results, and the proofs use only standard and well-known techniques. Furthermore, some of the results are already known. However, they have their place in the context of the study of weighted trees, see [5], [6].

Definition 1.1 (Weighted tree) A *weighted bicolored plane tree*, or a *weighted tree*, or just a *tree* for short, is a bicolored plane tree whose edges are endowed with positive integral *weights*. The sum of the weights of the edges of a tree is called the *total weight* of the tree.

The *degree* of a vertex is the sum of the weights of the edges incident to this vertex. Obviously, the sum of the degrees of black vertices, as well as the sum of the degrees of white vertices, is equal to the total weight n of the tree. Let the tree have p black vertices, of degrees $\alpha_1, \ldots, \alpha_p$, and q white vertices, of degrees $\beta_1, \ldots, \beta_q$, respectively. Then the pair of partitions (α, β), $\alpha, \beta \vdash n$, is called *passport* of the tree.

The *weight distribution* of a weighted tree is a partition $\mu \vdash n$, $\mu = (\mu_1, \mu_2, \ldots, \mu_m)$ where $m = p + q - 1$ is the number of edges, and μ_i, $i = 1, \ldots, m$ are the weights of the edges. Leaving aside the weights and considering only the underlying plane tree, we speak of a *topological tree*, which is a bicolored plane tree. Weighted trees whose weight distribution is $\mu = 1^n$ will be called *ordinary trees*: they coincide with the corresponding topological trees.

The adjective *plane* in the above definition means that our trees are considered not as mere graphs but as plane maps. More precisely, this means that the cyclic order of branches around each vertex of the tree is fixed, and changing this order will in general give a different tree. *All the trees considered in this paper will be endowed with the "plane" structure*; therefore, the adjective "plane" will often be omitted.

Example 1.2 (Weighted tree) Figure 1 shows an example of a weighted tree. The total weight of this tree is $n = 18$; its passport is $(\alpha, \beta) = (5^2 2^3 1^2, 7^1 6^1 4^1 1^1)$; the weight distribution is $\mu = 5^1 3^1 2^2 1^6$.

Definition 1.3 (Rooted tree) A tree with a distinguished edge is called a *rooted tree*, and the distinguished edge itself is called its *root*. We consider the root edge as being oriented from black to white.

The goal of this paper is the enumeration of rooted weighted (bicolored plane) trees.
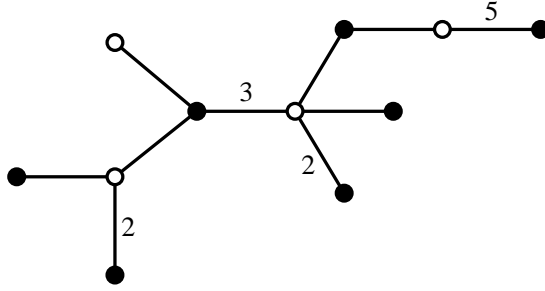
Figure 1: Weighted bicolored plane tree. The weights which are not indicated are equal to 1.

2 Statement of the main theorem

Theorem 2.1 (Enumeration of weighted trees) *Let a_n be the number of rooted weighted bicolored plane trees of weight n. Then the generating function $f(t) = \sum_{n \geq 0} a_n t^n$ is equal to*

$$
\begin{aligned}
f(t) &= \frac{1 - t - \sqrt{1 - 6\,t + 5\,t^2}}{2\,t} \\
&= 1 + t + 3\,t^2 + 10\,t^3 + 36\,t^4 + 137\,t^5 + 543\,t^6 + 2219\,t^7 + 9285\,t^8 + \ldots
\end{aligned} \tag{1}
$$

Numbers a_n satisfy the following recurrence relation:

$$
a_0 = 1, \quad a_1 = 1, \qquad a_{n+1} = a_n + \sum_{k=0}^{n} a_k a_{n-k} \quad for \quad n \geq 1. \tag{2}
$$

The asymptotic formula for the numbers a_n is

$$
a_n \sim \frac{1}{2}\sqrt{\frac{5}{\pi}} \cdot 5^n \, n^{-3/2}. \tag{3}
$$

Let $b_{m,n}$ be the number of rooted weighted bicolored plane trees of weight n with m edges. Then the generating function $h(s,t) = \sum_{m,n \geq 0} b_{m,n} s^m t^n$ is equal to

$$
\begin{aligned}
h(s,t) &= \frac{1 - t - \sqrt{1 - (2 + 4s)\,t + (1 + 4s)\,t^2}}{2st} \\
&= 1 + st + (s + 2s^2)\,t^2 + (s + 4s^2 + 5s^3)\,t^3 + (s + 6s^2 + 15s^3 + 14s^4)\,t^4 + \ldots
\end{aligned} \tag{4}
$$

The following is an explicit formula for the numbers $b_{m,n}$:

$$
b_{m,n} = \binom{n-1}{m-1} \cdot \mathrm{Cat}_m = \binom{n-1}{m-1} \cdot \frac{1}{m+1}\binom{2m}{m}, \tag{5}
$$

where Cat_m is the mth Catalan number.

Denote $|\mathrm{Aut}(T)|$ the order of the automorphism group of a tree T. Let c_n be the number of non-isomorphic non-rooted trees T of weight n, each counted with the factor $1/|\mathrm{Aut}(T)|$. Then

$$
c_n = \sum_T \frac{1}{|\mathrm{Aut}(T)|} = \sum_{m=1}^{n} \frac{b_{m,n}}{m}, \tag{6}
$$

where the first sum is taken over all the non-isomorphic non-rooted trees T of weight n.

The sequence a_n is listed in the On-Line Encyclopedia of Integer Sequences [4] as the entry A002212. It has many interpretations, some of them coming from chemistry. Among the various interpretations there are "multi-trees" (Roland Bacher, 2005) which correspond to our weighted trees. Almost all the above-stated formulas may also be found in [4].

Example 2.2 (Trees of weight 4) Figure 2 shows the trees of weight 4. There are ten trees in the picture, but in fact there are 16 non-isomorphic (non-rooted) trees of weight 4. Indeed, when we exchange black and white, four trees remain isomorphic to themselves while six others don't, so we must add to the set the six missing trees. Near each tree, the number of its possible rootings is indicated, with color exchange taken into account. We see that the total number of trees is 36, which is the coefficient a_4 in front of t^4 in $f(t)$, see (1). Among these 36 trees, there is one tree with one edge, six trees with two edges, 15 trees with three edges, and 14 trees with four edges. These are the coefficients of the polynomial $s + 6s^2 + 15s^3 + 14s^4$ which stands in front of t^4 in $h(s,t)$, see (4).

The number c_4, according to (6), is equal to

$$1 + \frac{6}{2} + \frac{15}{3} + \frac{14}{4} = 12\frac{1}{2}.$$

And, indeed, among the 16 non-isomorphic non-rooted trees there are ten asymmetric trees, four trees with the symmetry order 2, and two trees with the symmetry order 4, which gives

$$10 + 4 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} = 12\frac{1}{2}.$$
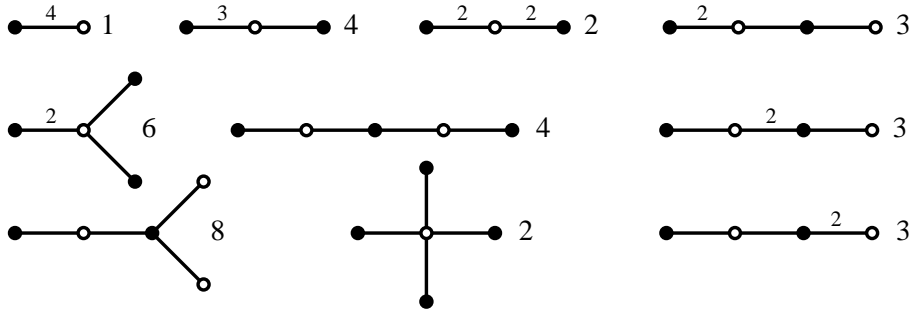
We leave the details to the reader.

Figure 2: Near each tree, the number of its possible rootings is indicated, with an eventual color exchange taken into account. The total number of rooted trees is 36.

3   Dyck words and weighted Dyck words

There is a standard way of encoding rooted topological (non-weighted) plane trees by *Dyck words* and *Dyck paths*. We start on the left bank of the root edge and go around the tree in the clockwise direction, writing the letter x when we follow an edge for the first time, and the letter y when we follow it the second time on its opposite side. A Dyck path corresponding to a Dyck word is a path on the plane which starts at the origin and takes a step $(1,1)$ for every letter x and a step $(1,-1)$ for every letter y. These objects may be easily characterized.

For a word w which is a concatenation of the three words, $w = u_1 u_2 u_3$ (either is allowed to be empty), we call u_1 a *prefix*, u_2 a *factor*, and u_3 a *suffix* of w.

Definition 3.1 (Dyck words and Dyck paths) A *Dyck word* is a word w in the alphabet $\{x,y\}$ such that $|w|_x = |w|_y$ (here $|w|_x$ and $|w|_y$ stand for the number of occurencies of x and y in w), while for any prefix u of w we have $|u|_x \geq |u|_y$. A *Dyck path* is a path on the plane which starts at the origin, takes steps $(1,1)$ and $(1,-1)$, and finishes on the horizontal axis, while always staying on the upper half-plane.

Figure 3 illustrates these notions. The root edge in the tree is shown by the thick line. Note that a Dyck word may be empty; then it corresponds to the tree consisting of a single vertex. Making an exception to the general rule, we do not color this vertex in black or white. Thus, there exists a single empty word, and a single tree without edges.

$$x\,x\,x\,x\,x\,y\,y\,y\,\mathbf{x}\,x\,y\,x\,y\,\mathbf{y}\,x\,y\,y\,y\,x\,x\,x\,y\,y\,y\,y\,x\,y$$
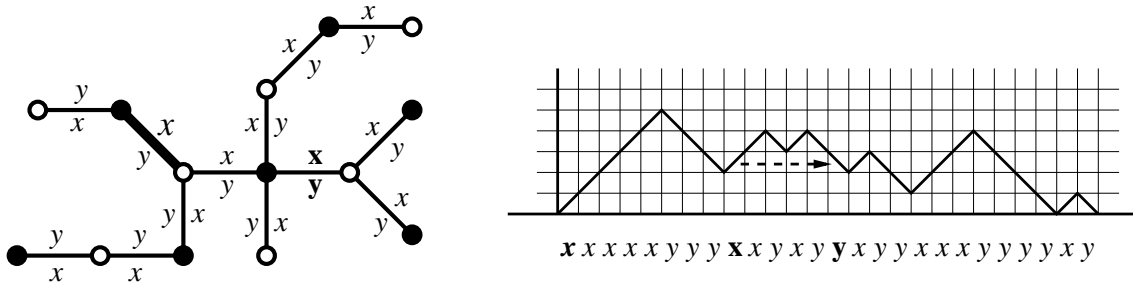
Figure 3: A rooted bicolored plane tree and its encoding by the corresponding Dyck word and Dyck path.

The following proposition is a trivial consequence of the construction.

Proposition 3.2 (Trees and Dyck words) *There is a bijection between rooted bicolored plane trees and Dyck words.*

There remains very little to do in order to describe weighted trees. There is a natural notion of coupling of the letters of a Dyck word: a *couple* is the pair of letters (x, y) standing on the opposite sides of the same edge. It is easy to recognize a couple in a word or in the path. Let x, y be a pair of letters in a Dyck word, where x comes before y in w. Consider the factor xuy of w which starts with x and terminates with y. Then the pair (x, y) forms a couple if and only if the factor u between x and y is a Dyck word. In a Dyck path, we take an ascending step corresponding to a letter x and go horizontally until we meet a descending step opposite to it: this step corresponds to the letter y which forms a couple with x. In Figure 3, an example of a couple, both on the tree and in the word, is indicated in a boldface font, and the dashed arrow shows how to find the "opposite" step. The couple (x, y) in which x is the very first letter of the Dyck word, corresponds to the root edge.

Now, returning to the weighted trees, we do the following: for every edge of the tree, we take the corresponding couple (x, y) and replace it with (x_i, y_i) where i is the weight of the edge.

Definition 3.3 (Weighted Dyck words) A *weighted Dyck word* is a word in the infinite alphabet $\{x_i, y_i\}_{i \geq 1}$ which is a Dyck word in which every couple of letters (x, y) is replaced by a certain couple (x_i, y_i). We say that a couple (x_i, y_i) has the weight i, and the weight of a word is the sum of the weights of all its couples.

Proposition 3.4 (Weighted trees and weighted Dyck words) *There is a bijection between rooted weighted bicolored plane trees and weighted Dyck words.*

4   Proof of the main theorem

Every non-empty Dyck word w has a unique decomposition of the form $w = xuyv$ where u and v are themselves Dyck words (maybe empty). Here, obviously, x is the first letter of w, and y is the letter coupled with it. The corresponding step in the Dyck path is the descending step of the first return of the path to the horizontal axis.

In the same way, every non-empty weighted Dyck word w has a unique decomposition of the form $x_i u y_i v$ for some $i \geq 1$, where u and v are weighted Dyck words.

Let $\mathcal{D}$ be the formal sum of all the weighted Dyck words, that is, the formal power series

$$\mathcal{D} = \varepsilon + x_1 y_1 + x_2 y_2 + x_1 y_1 x_1 y_1 + x_1 x_1 y_1 y_1 + x_3 y_3 + x_2 y_2 x_1 y_1 + x_2 x_1 y_1 y_2 + \ldots \qquad (7)$$

in non-commuting variables x_i, y_i, $i = 1, 2, \ldots$, where ε stands for the empty word. (In order to write down a series we must choose a total order on the words. A particular choice of the order is irrelevant. In (7), the words are ordered according to their weight.) Then, the above decomposition

4

of the words of $\mathcal{D}$ in the form $x_i u y_i v$ implies the following equation for $\mathcal{D}$:

$$\mathcal{D} = \varepsilon + x_1 \mathcal{D} y_1 \mathcal{D} + x_2 \mathcal{D} y_2 \mathcal{D} + \ldots = \varepsilon + \sum_{i=1}^{\infty} x_i \mathcal{D} y_i \mathcal{D}. \tag{8}$$

Now, do the following:

- replace each letter y_i in $\mathcal{D}$ by 1;

- replace each letter x_i in $\mathcal{D}$ by st^i;

- make the variables s and t commute.

Then, every word w in $\mathcal{D}$ is transformed into a word $s^m t^n$ where m is the number of occurencies of the letters x_i, $i \geq 1$, in w (or, equivalently, the number of edges of the weighted tree T_w corresponding to w), and n is the weight of w (or, equivalently, the total weight of T_w). Therefore, combining similar terms we get the generating function $h(s,t) = \sum_{m,n \geq 0} b_{m,n} s^m t^n$. At the same time, equation (8) is transformed into the following quadratic equation for $h(s,t)$:

$$h = 1 + s \left(\sum_{i=1}^{\infty} t^i \right) h^2 = 1 + \frac{st}{1-t} \cdot h^2. \tag{9}$$

Solving this equation, and choosing the sign in front of the square root in such a way as to avoid a singularity at zero, we obtain formula (4). Then, substituting $s = 1$ in (4) we get (1).

In order to obtain the asymptotic expression (3) for the numbers a_n it suffices to apply to $f(t)$ the ready-made formulas of asymptotic analysis of the coefficients of generating functions – see, for example, [1], Chapter VI. The only thing to note is that

$$1 - 6t + 5t^2 = (1 - t)(1 - 5t).$$

In order to prove (5), we proceed as follows. There are Cat_m topological rooted trees with m edges. Starting at the root edge, we go around a tree in the clockwise direction and attribute a non-zero weight to every newly encountered edge. There are $\binom{n-1}{m-1}$ ways to do that. Indeed, put n dots in a row, and distribute $m - 1$ separators among $n - 1$ places between the dots. This procedure splits the number n into m non-zero parts.

In order to prove the recurrence (2), consider separately the trees of weight $n + 1$ having the root edge of weight 1, and the trees of weight $n + 1$ having the root edge of weight $i \geq 2$. The weighted Dyck words corresponding to the trees of the first kind are of the form $x_1 u y_1 v$, where u and v are themselves weighted Dyck words. The sum of the weights of u and v is n; denoting the weight of u by k, so that the weight of v becomes $n - k$, and summing over the $k = 0, 1, \ldots, n$, we get the term $\sum_{k=0}^{n} a_k a_{n-k}$ of (2). Now, all the trees of weight $n+1$ having the root edge of weight $i \geq 2$ are obtained from the trees of weight n having the root edge of weight $i - 1$, by adding one unit to the weight of the root. This gives the term a_n in the right-hand part of (2).

Finally, (6) follows from the fact that there are m choices of a root edge in a tree T with m edges, but if this tree has non-trivial symmetries then some of these choices produce isomorphic rooted trees. The number of non-isomorphic rootings is $m/|\mathrm{Aut}(T)|$. Thus, dividing by m, we get the factor $1/|\mathrm{Aut}(T)|$.

Theorem 2.1 is proved. $\qquad\square$

5 Enumeration of ordinary trees according to their passport

Let $\lambda \vdash n$, $\lambda = (\lambda_1, \lambda_2, \ldots \lambda_k)$ be a partition of and integer n. Let us write λ in the *power notation*:

$$\lambda = 1^{d_1} 2^{d_2} \ldots n^{d_n}, \qquad \text{where} \qquad \sum_{i=1}^{n} d_i = k, \qquad \sum_{i=1}^{n} i \cdot d_i = n,$$

so that d_i is the number of the parts of λ equal to i. Denote

$$N(\lambda) \; = \; \frac{(k-1)!}{d_1! \, d_2! \, \ldots \, d_n!} \, . \tag{10}$$

The following theorem was proved in [7] (1964) and later generalized in [2] (1992):

Theorem 5.1 (Ordinary trees with a given passport) *The number of rooted ordinary bicolored plane trees with the passport (α, β) is equal to*

$$nN(\alpha)N(\beta). \tag{11}$$

Respectively, the number of the non-isomorphic ordinary bicolored plane trees with the passport (α, β), each one of them counted with the factor $1/|\mathrm{Aut}(T)|$, is

$$\sum_T \frac{1}{|\mathrm{Aut}\,(T)|} \; = \; N(\alpha)N(\beta) \tag{12}$$

where the sum is taken over the the non-isomorphic ordinary bicolored plane trees with the passport (α, β).

The above theorem is much more powerful than our Theorem 2.1. First of all, it gives an explicit formula; and, what is more important, it enumerates the trees not according to one or two parameters (as the weight and the number of edges in our case) but according to their passport. In the weighted case, a major difficulty in obtaining a similar formula stems from the fact that the same passport can be realized by a tree and by a forest, as one can see in a very simple example of Figure 4. Therefore, an inclusion-exclusion procedure might be unavoidable. This is indeed what takes place in [3].

Figure 4: The same passport $(5^1 3^1, 5^1 3^1)$ is realized by a forest and by a tree.

References

[1] **Flajolet P., Sedgewick R.** Analytic Combinatorics. – Cambridge Univ. Press, 2009, XIV+810 pp.

[2] **Goulden I. P., Jackson D. M.** The combinatorial relationship between trees, cacti and certain connection coefficients for the symmetric group. *Europ. J. Combinat.*, 1992, vol. **13**, 357–365.

[3] **Kochetkov Yu.** Enumeration of one class of plane weighted trees. – `arXiv:1310.6208v1`, 11 pp.

[4] **The On-Line Encyclopedia of Integer Sequences**, `http://oeis.org/`.

[5] **Pakovich F., Zvonkin A. K.** Minimum degree of the difference of two polynomials over $\mathbb{Q}$, and weighted plane trees. – `arXiv:1306.4141v1`. To appear in *Selecta Mathematica*, 2014.

[6] **Pakovich F., Zvonkin A. K.** Minimum degree of the difference of two polynomials over $\mathbb{Q}$. Part II: Davenport–Zannier triples. – In preparation (a preliminary version may be found at `http://www.labri.fr/perso/zvonkin/`).

[7] **Tutte W. T.** The number of planted plane trees with a given partition. – *Amer. Math. Monthly*, 1964, vol. **71**, no. 3, 272–277.