Hopf algebras and topological recursion

João N Esteves

CAMGSD, Departamento de Matemática, Instituto Superior Técnico, Av. Rovisco Pais 1, 1049-001 Lisboa, Portugal

E-mail: joao.n.esteves@tecnico.ulisboa.pt

Received 7 July 2015, revised 25 August 2015
Accepted for publication 14 September 2015
Published 12 October 2015

Abstract
We consider a model for topological recursion based on the Hopf algebra of planar binary trees defined by Loday and Ronco (1998 Adv. Math. 139 293–309). We show that extending this Hopf algebra by identifying pairs of nearest neighbor leaves, and thus producing graphs with loops, we obtain the full recursion formula discovered by Eynard and Orantin (2007 Commun. Number Theory Phys. 1 347–452).

Keywords: Hopf algebras, topological recursion, trees, graphs

1. Introduction

The use of graphs, in particular of trees, binary trees and planar binary trees, in mathematical physics has a long tradition. The canonical examples are perhaps Feynman diagrams but the connection with Hopf algebras of trees started with the works of Connes and Kreimer [6–8] that describe the combinatorics of the procedure of extracting subdivergences in quantum field theory, known as the BPHZ renormalization procedure [5]. Another approach to using graphs in QFT, and in particular in QED, considering binary trees, planar or not, was followed by Brouder and Frabetti [3, 4, 14]. Later it was understood that these two approaches are very similar, and in some cases equivalent, and are related to quasisymmetric and non-commutative quasisymmetric functions, see for instance [1, 12, 13, 15].

In this paper we show how the Hopf algebra of planar binary trees defined by Loday and Ronco [17] can be seen as a representation of the vector space generated by correlation functions that obey the Eynard–Orantin recursion formula. These correlation functions are graded by the Euler characteristic and we can consider for each degree the vector space over \( \mathbb{Q} \) generated by them and then take the direct sum of these vector spaces for all degrees. First we consider planar binary trees of order \( n \), that is with \( n \) vertices and \( n + 1 \) leaves, as a representation of the genus \( g = 0 \) correlation function \( W_g(p, p_1, \ldots, p_{n-1}) \) of Euler

---

1 The author was supported by Fundação para a Ciência e a Tecnologia through the grant SFRH/BPD/77123/2011. He also wishes to thank José Mourão and Nicolas Orantin for fruitful discussions.
characteristic $\chi = 2 - 2g - k$ equal to $-n$. Here the Euler characteristic is one of the Riemann or topological surfaces of genus $g$ and $k$ punctures or borders to which the correlation functions $W_k(p, p_1, \ldots, p_k)$ are usually related in some concrete problems. For $g = 0$ we label the root with $p$ and the $n + 1$ leaves with the $p_1, \ldots, p_{n+1}$ variables. Then, by connecting the nearest neighbors’ leaves with a single edge and reducing the number of pairs of labels in the same way as increasing the genus, we obtain graphs with loops that we see as a representation of higher genus correlation functions with the same Euler characteristic.

As an example, take $W_3^2(p)$ which has $\chi = -3$. Its underlying generating trees are planar binary trees of order three, which are also models for $W_3^2(p, p_1, p_2, p_3, p_4)$: identifying pairs of nearest neighbor leaves in the left and right branches independently, we get the second term of topological recursion. Identifying pairs of leaves each taken from the left and the right branches gives the first term. Note that in this case not every planar binary tree of order three gives $W_3^2$. In fact, the first tree of figure 1 does not give a genus 2 correlation function by identifying the nearest neighbor leaves in opposite branches.

2. The topological recursion of Eynard and Orantin

The topological recursion formula discovered by Eynard and Orantin has its origin in matrix models, for general reviews see for instance [9, 11]. In the Hermitian 1-matrix form of the theory the purpose is to compute connected correlation functions $W_{k+1}$ depending on a set of variables $p, p_1, \ldots, p_k$.
\[ W_{k+1}(p, p_1, \ldots, p_k) = \left\{ \text{Tr} \frac{1}{p - M} \text{Tr} \frac{1}{p_1 - M} \ldots \text{Tr} \frac{1}{p_k - M} \right\}_{c} \]  

starting with \( W_1(p) \) and \( W_2(p, p_1) \). These functions, which are solutions of the so-called loop equations, are only well defined over Riemann surfaces because in \( \mathbb{C} \) they are multivalued. They admit an expansion in the order \( N \) of the random matrix \( M \), with components \( W_{k+1}(p, p_1, \ldots, p_k) \) related to a definite genus. We will not be concerned here with the actual computation of correlation functions in specific models.

Let \( K = (p_1, \ldots, p_k) \) be a vector of variables. For instance, in concrete cases these can be coordinates of punctures on Riemann surfaces, labels of borders on topological surfaces or variables in matrix models, but we just leave them as labels of leaves of planar binary trees or of graphs obtained from planar binary trees. We assign the label \( p \) to the root of a tree or of a graph with loops obtained from a tree. The topological recursion formula is

\[ W_{k+1}(p, K) = \sum_{\text{branchpoints}} \text{Res}_{p \to a} K_{p}(q, \bar{q}) \left( W_{k+2}^{-1}(q, \bar{q}, K) + \sum_{L \cup M = K, h = 0}^{g} W_{[L]+1}^{h}(q, L) W_{[M]+1}^{r-h}(\bar{q}, M) \right) \]

where the sum is restricted to terms with an Euler characteristic equal to or smaller than 0. For instance, if \( h = 0 \) then \( |L| \geq 1 \). For a very clear exposition about this setup from the point of view of algebraic geometry, see for instance [10], but some comments are in order. The branch points are the ones from a meromorphic function \( x \) defined on a so-called spectral curve \( \mathcal{E}(x, y) = 0 \). The recursion kernel \( K_{p}(q, \bar{q}) \) is, roughly speaking, a meromorphic \((1,1)\) tensor that depends on a regular point \( p \) in the neighbourhood of a branch point and on \( q \) and its conjugated point \( \bar{q} \) for which \( x(q) = x(\bar{q}) \) and \( y(q) = y(\bar{q}) \). In fact it can be computed from \( W_{1}^{0}(p) \) and \( W_{2}^{2}(p, p) \) which are symmetric differentials of order one and two respectively. Actually, all \( W_{k}^{f} \) are meromorphic symmetric differentials but we will continue to refer to them as correlation functions. Since our approach will be purely algebraic, and in order to soften the notation, we will not explicitly mention the sum of the residues over the branch points when referring to this formula.

3. The Loday–Ronco Hopf algebra of planar binary trees

We collect here some important facts of the Loday–Ronco Hopf algebra. Details and proofs can be found in [1, 17]. Let \( S_n \) be the symmetric group of order \( n \) with the usual product \( \rho \cdot \sigma \) given by the composition of permutations. When necessary we denote a permutation \( \rho \) by its image \( (\rho(1)\rho(2)\ldots\rho(n)) \). Recall that a shuffle \( \rho(p, q) \) of type \( (p, q) \) in \( S_n \) is a permutation such that \( \rho(1) < \rho(2) < \ldots < \rho(p) \) and \( \rho(p + 1) < \rho(p + 2) < \ldots < \rho(p + q) \). For instance the shuffles of type \((1, 2)\) in \( S_3 \) are \((123), (213) \) and \((312)\). We denote the set of \((p, q)\) shuffles by \( S(p, q) \). Take

\[ k \left[ S^\infty \right] = \bigoplus_{n=0}^{\infty} k \left[ S_n \right] \]

with \( S_0 \) identified with the empty permutation. \( k \left[ S^\infty \right] \) is a vector space over a field \( k \) of characteristic 0 generated by linear combinations of permutations. It is graded by the order of permutations and \( k \left[ S_n \right] \), which contains the empty permutation, is identified with the field \( k \). For two permutations \( \rho \in S_n \) and \( \sigma \in S_q \) there is a natural product on \( S^\infty \) denoted by \( \rho \times \sigma \), which is a permutation on \( S_{p+q} \) given by letting \( \rho \) act on the first \( p \) variables and \( \sigma \) acting on the last \( q \) variables.
There is a unique decomposition of any permutation $\sigma \in S_n$ in two permutations $\sigma_i \in S_i$ and $\sigma_{n-i} \in S_{n-i}$ for each $i$ such that

$$\sigma = (\sigma_i \times \sigma_{n-i}) \cdot w^{-1}$$

(4)

where $w$ is a shuffle of type $(i, n-i)$. With the $*$ product

$$\rho \ast \sigma = \sum_{\alpha_{n,m} \in S_{n,m}} \alpha_{n,m} \cdot (\rho \times \sigma)$$

(5)

and the coproduct

$$\Delta \sigma = \sum \sigma_i \otimes \sigma_{n-i}$$

(6)

$k[S^\infty]$ becomes a bi-algebra and since it is graded and connected it is automatically a Hopf algebra.

A planar binary tree is a graph with no loops embedded in the plane with only trivalent vertices. In every planar binary tree there are paths that start on a special edge called the root and end on the terminal edges called leaves. The leaves can be left or right oriented. The order $|t|$ of a planar binary tree $t$ is the number of its vertices and on each planar binary tree of order $n$ there are $n + 1$ leaves that are usually numbered from 0 to $n$ from left to right. It is frequent to visualize planar binary trees from the bottom to the top, with the root as its lowest vertical edge and the leaves as the highest edges, oriented SW–NE or SE–NW. We will denote the set of planar binary trees of order $n$ by $Y^n$ and by $k[Y^n]$ the vector space over $k$ generated by planar binary trees of all orders. Additionally, a planar binary tree with levels is a planar binary tree such that on each horizontal line there is at most one vertex. It is clear that, reading the vertices from left to right and from top to bottom, it is possible to assign a permutation of order $n$ to a planar binary tree with levels and that this assignment is unique. For example, in figure 3 the tree corresponds to the permutation (132). In this way it is completely equivalent to consider the Hopf algebra $k[S^\infty]$ or the Hopf algebra of planar binary trees with levels because they are isomorphic. However, Loday and Ronco show in [17] that the $*$ product and the coproduct are internal on the algebra of planar binary trees, which is then isomorphic to a sub-Hopf algebra of $k[S^\infty]$ with the same product and coproduct. The identity of the Hopf algebra $k[Y^n]$ is the tree with a single edge and no vertices, following the convention of considering only internal vertices, which represents the empty permutation, and the trivial permutation of $S_1$ is represented by the tree with one vertex and two leaves, see figure 4. In fact, this element is the generator of the augmented algebra by the $*$ product. See figure 5 for an example of an order three product.

Figure 3. Planar binary tree with levels, the image of (132).
The grafting \( t_1 \lor t_2 \) of two trees \( t_1 \) and \( t_2 \) is the operation of producing a new tree \( t \) by inserting \( t_1 \) on the left and \( t_2 \) on the right leaves of \( t_1 \). It is clear that any tree of order \( n \) can be written as \( t \lor t_2 \) with \( t_1 \) of order \( p \), \( t_2 \) of order \( q \) and \( n = p + q + 1 \). If a tree has only leaves on the right branch besides the first leaf then it can be written as \( \_ \lor t_2 \) and, reciprocally, if it has only leaves on the left branch besides the last leaf. Note that in particular \( (1) = \_ \lor \_ \). In [17] Loday and Ronco show that the \( * \) product restricted to planar binary trees satisfies the identity

\[
 t * t' = t_1 \lor (t_2 * t') + (t * t'_1) \lor t'_2
\]

and

\[
 t * \_ = \_ * t = t
\]

with \( t = t_1 \lor t_2 \) and \( t' = t'_1 \lor t'_2 \).

If \( t_1 \) is a tree of order \( p \) and a representative element of \( W^0_{p+2}(p, L) \) and \( t_2 \) is a tree of order \( q \) and a representative element of \( W^0_{q+2}(p, M) \) then \( t = t_1 \lor t_2 \) is a tree of order \( n = p + q + 1 \) and a representative element of...
\[ W_{n+2}^0(p, K) = K_p(q, \bar{q}) W_{p+2}^0(q, L) W_{q+2}^0(\bar{q}, M) \]

with \( K = \{ p_1, \ldots, p_{n+1} \} = L \cup M \). We will clarify this in what follows.

**4. The solution of topological recursion**

**4.1. Genus 0**

A representation map \( \psi \) from the vector space of correlation functions of genus \( g \) to the vector space of graphs with loops should be defined such that, in particular, to a correlation function \( W_{n+2}^0(p, p_1, \ldots, p_{n+1}) \) of Euler characteristic \( \chi = -n \) would correspond the trees of order \( n \). In fact, we will state below that the representation of \( W_{n+2}^0(p, p_1, \ldots, p_{n+1}) \) is the sum of all trees of order \( n \). It is not clear that this map gives a true representation in the strict mathematical sense. It is linear by definition and it is obvious that it is surjective, as we can associate some instance of a correlation function \( W_{n+2}^0(p, p_1, \ldots, p_{n+1}) \) to any tree of order \( n \). If it is injective and a homomorphism is a more delicate issue because even if one considers \( W_{n+2}^0 \) as being a sum of all instances of correlation functions of Euler characteristic \( -n \), each represented by a tree \( t \in Y^n \) in the same way as in particle physics, where different Green functions contribute to the same scattering amplitude, it is not evident that the space of correlation functions has a product with an identity that would correspond to the trivial tree. Note that this would give at least a ring structure and, in the case of topological quantum field theory where correlations functions are identified with topological surfaces with punctures, cobordism is a good candidate for such a product. In fact it is a consequence of the axioms of topological quantum field theory, as stated by Atiyah for example in [2], that the cylinder \( \Sigma \times I \), where \( \Sigma \) is a topological surface without border and \( I \) is an interval of real numbers, is identified with the identity map between two vector spaces. In any case, we will not elaborate more on this here and use the word representation in a somewhat rough sense. In particular, when referring to the inverse image of a tree or a sum of trees we will refrain of using the inverse \( \psi^{-1} \) but will use instead \( \psi^* \), as for the pullback.

**Definition 1.** Consider the planar binary tree with one vertex (1). The 3-point correlation function \( W^0_3(p, p_1, p_2) \) is represented by the sum of two planar binary trees with one vertex, obtained by the permutation of the leaf labels \( p_1 \) and \( p_2 \):

\[
\psi(\psi^* W^0_3(p, p_1, p_2)) = \sum_{\text{perm. of leaf labels } (p_1, p_2)} (1)
\]

The trees that represent \( W^0_3(p, p_1, p_2) \) are given by the permutations of the leaf labels of \( \{ 1 \} \). Then it is natural to represent the operation of grafting two trees by the insertion of the recursion kernel \( K_p(q, \bar{q}) \) on its roots. Therefore the symbol \( \gamma \) has two meanings. When isolated it represents \( W^0_3(p, p_1, p_2) \) because the two cylinders \( W^0_3(q, p_1) \) and \( W^0_3(\bar{q}, p_2) \) are implicitly identified by its leaves. When it is an internal vertex of a more complex tree it is the recursion kernel \( K_p(q, \bar{q}) \) with suitable labels of its variables.

**Definition 2.** def:W3. The propagator or cylinder (also named the Bergman kernel in the literature) \( W^0_3(q, \bar{q}) \) is represented through \( \psi \) by the empty permutation \( 1 \) and the recursion kernel is represented through \( \psi \circ \gamma \) when in an internal vertex of some tree. Then each planar binary tree of order \( n \) is a representation of an instance of some correlation function in genus 0
with each vertex identified with a recursion kernel and each left leaf identified with the cylinder \( W^0_p(q_i, p_i) \) or each right leaf identified with the cylinder \( W^0_q(q_i, p_i) \). Finally the image under \( \psi \) of a correlation function \( W^0_{n+2}(p, p_1, \ldots, p_{n+1}) \) with \( \chi = -n \) is the sum of all planar binary trees of order \( n \) considering all permutations of their leaf labels and with the identifications mentioned above,

\[
\psi\left(W^0_{n+2}(p, p_1, \ldots, p_{n+1})\right) = \sum_{\text{perm. of leaf labels } \{p_0, \ldots, p_{n+1}\}} t_i.
\] (10)

Hence definition 1 becomes the following example.

**Example 1.** Consider the planar binary tree with one vertex. The 3-point correlation function \( W^0_3(p, p_1, p_2) \) is represented by the sum of two planar binary trees with one vertex, obtained by the permutation of the leaf labels \( p_1 \) and \( p_2 \).

\[
\psi\left(W^0_3(p, p_1, p_2)\right) = \psi\left(K_p(q, \bar{q})W^0_1(q, p_1)W^0_2(\bar{q}, p_2)\right) + \text{perm. of } \{p_1, p_2\}
= \sum_{\text{perm. of } \{p_1, p_2\}} t_1 \lor t_2
= \sum_{\text{perm. of } \{p_1, p_2\}} (1).
\] (11)

**Proposition 1.** If \( W^0_{n+2}(p, p_1, \ldots, p_{n+1}) \) is a correlation function with Euler characteristic \( \chi = -n \) that is a solution of (2) then we have

\[
\psi\left(W^0_{n+2}(p, p_1, \ldots, p_{n+1})\right) = \sum_{\text{|t|=p, |t|=q}} t_1 \lor t_2
+ \text{perm. of leaf labels } \{p_1, \ldots, p_{n+1}\}.
\] (12)

**Proof.** This is the topological recursion in genus 0 written with planar binary trees. For \( n = 1 \), this is example 1. For arbitrary \( n \) by definition 1

\[
\psi\left(W^0_{n+2}(p, p_1, \ldots, p_{n+1})\right) = \sum_{\text{perm. of leaf labels } \{p_1, \ldots, p_{n+1}\}} t_i.
\] (13)

Decompose uniquely any \( t \) of order \( n \) into \( t = t_1 \lor t_2 \) of orders \( |t_1| = p \) and \( |t_2| = q \) with \( p + q + 1 = n \) to get

\[
\psi\left(W^0_{n+2}(p, p_1, \ldots, p_{n+1})\right) = \sum_{\text{perm. of leaf labels } \{p_1, \ldots, p_{n+1}\}} t_1 \lor t_2.
\] (14)
Then $t_1$ and $t_2$ are on the image by $\psi$ of $W_{p+2}$ and $W_{q+2}$ for $p$ and $q$ varying from 0 to $n - 1$ and constrained by $p + q + 1 = n$. Since the operation of grafting two trees is represented by attaching the recursion kernel to its roots then, summing for all $t_1 \in Y^p$, $t_2 \in Y^q$ and for $p + q + 1 = n$, we get the topological recursion formula for $g = 0$ after taking the pre-image of (14) by $\psi$:

$$W_{n+2}^0(p, p_1, \ldots, p_{n+1}) = K_p(q, \bar{q}) \sum_{L|M|+1} W_{|L|+1}^0(q, L) W_{|M|+1}^0(\bar{q}, M).$$

(15)

Remark 1. Note that by $W_{n+2}^0(p, K)$ with $\left|K\right| = n + 1$ we understand all instances of the correlation function with $g = 0$ and $n + 2$ labels. This is similar to the situation in high energy physics, where for the same physical process described by a scattering amplitude there are several Feynman diagrams that contribute.

It is well known that the dimension of the vector space generated by planar binary trees of order $n$ is given by the Catalan number (see for instance [14]),

$$c_n = \frac{2n!}{n!(n + 1)n!}.$$

It is also known that correlation functions in matrix models have a large $N$ or planar expansion that is given in terms of Catalan numbers. Therefore it is of no surprise that there exists a correspondence between planar binary trees and correlation functions of genus 0.

**Theorem 1.** The $n$-order solution $W_{n+2}^0(p_1, \ldots, p_{n+1})$ of the topological recursion in genus 0 is represented by the linear combination

$$\sum t = (1)^* (1)^* \ldots (1)^*$$

with $n$ factors of (1) followed by the sum over all permutations of its labels.

In this way $W_{n+2}^0(p, p_1, \ldots, p_{n+1})$ is represented by $\sum$ followed by the identification of cylinders $W_2^0(q, p_1)$ with the left leaves or $W_2^0(\bar{q}, p_1)$ with the right leaves and finally by summing over all permutations of the labels $p_1, p_2, \ldots, p_{n+1}$. In other words, the $*$ product $(1) * (1) * \ldots * (1)$ gives all possible insertions of recursion kernels of $W_{n+2}^0$.

**Proof.** By induction on the Euler characteristic or equivalently on the order $n$. For $n = 1$ we saw that $W_2^0(p, p_1, p_2)$ is just a sum of two planar binary trees with one vertex, with the leaves in correspondence with $W_2^0(q, p_1)$ and $W_2^0(\bar{q}, p_2)$ or its permutations and the root labeled by $p$. So we start the induction at $n = 2$: we want to show that $(1) * (1)$ represents $W_4^0(p, p_1, p_2, p_3)$. In the Loday–Ronco Hopf algebra we have that $(1) * (1) = (12) + (21)$. On the other hand, the topological recursion formula gives for $W_4^0$
\[ W_2^0(p, p_1, p_2, p_3) = K_p(q, \bar{q}) \left( W_2^0(q, q_1) W_2^0(q_2, p_2, p_3) + \text{perm. of } \{p_1, p_2, p_3\} \right) + W_2^0(q, p_1) W_2^0(q_2, p_2, p_3) + \text{perm. of } \{p_1, p_2, p_3\} \]

which gives the two terms from the * product of the Loday–Ronco Hopf algebra, identifying the vertices with the recursion kernel and the leaves with the cylinders (see figure 6).

Next assume the induction hypothesis for \( n - 1 \) and note that if \( W_{n+1}^0 \) is represented by the linear combination \( \sum_{t \in \mathcal{Y}^{n-1}} (\mathbf{1}) \ast t_1 \ast t_2 \) with \( |t_1| = a, |t_2| = b \) and \( a + b + 1 = n - 1 \). Using (7) and (8) we get

\[
\sum_{t \in \mathcal{Y}^{n-1}} (\mathbf{1}) \ast t = \sum_{t \in \mathcal{Y}^{n-1}} \mathcal{Y}^{n-1} \ast t + \sum_{a+b+c=n-1} (\mathbf{1}) \ast t_1 \ast t_2.
\]

Because each \( t \) is of the order \( n - 1 \) but otherwise arbitrary, each \( t_1 \) is at most of the order \( n - 2 \) and then by the induction hypothesis \( (\mathbf{1}) \ast t_1 \) is in the image by \( \psi \) of a solution that is at most \( W_{n+1}^0 \) and at least \( W_3^0 \). Summing also over all permutations of \( K = \{p_1, \ldots, p_{n+1}\} \) we get

\[
\psi^* \left( \sum_{t \in \mathcal{Y}^{n-1}} (\mathbf{1}) \ast t \right)
\]

\[
= \sum_{L, \mathcal{M} = K, |L| = 1} K_p(q, \bar{q}) W_2^0(q, L) W_{n+1}^0(\bar{q}, M) + \sum_{L, \mathcal{M} = K, |L| > 1} K_p(q, \bar{q}) W_{n-1}^0(q, L) W_{n+1}^0(\bar{q}, M)
\]

\[
= \sum_{L, \mathcal{M} = K} K_p(q, \bar{q}) W_{n+1}^0(q, L) W_{n+1}^0(\bar{q}, M)
\]

\[
= W_{n+2}^0(p, p_1, \ldots, p_{n+1}).
\]
4.2. Genus higher than 0

The procedure of attaching an edge to two consecutive leaves and producing a graph with loops allows us to represent correlation functions with genus \( g > 0 \). This is equivalent to extracting the outermost cylinders \( W^0_2(x, p_j), W^0_2(y, p_{j+1}) \), \( x = q_i \) or \( q_j \), \( y = q_i \) or \( q_j \), and to couple a cylinder \( W^2_2(x, y) \) to two recursion kernels \( K_{q_i}(q_j, \tilde{q}_i) \) and \( K_{q_j}(q_i, \tilde{q}_j) \), for some convenient choice of indices, that are identified with two internal vertices. This procedure does not change the Euler characteristic of the associated correlation functions because the number of pairs of leaf labels is reduced exactly as the genus is increased. For instance, with this procedure we can make the sequence

\[
W^0_2(p, p_1, \ldots, p_k) \rightarrow W^1_2(p, p_1, p_2) \rightarrow W^2_2(p)
\]

and remain in the same graded vector space that contains \( k[Y^3] \). How this changes the Hopf algebra structure is not yet clear. For now, we define the operation \( i \leftrightarrow i+1 \) on a planar binary tree.

**Definition 3.** Starting with a planar binary tree of order \( n \) and \( n+2 \) labels (including the root label \( p \)) the operation \( i \leftrightarrow i+1 \) consists of erasing the labels of the leaves \( i \) and \( i+1 \) then connecting them by an edge and finally relabeling the remaining leaves, now numbered \( j \) with \( j = 0, \ldots, n-2 \), with the \( p_{j+1} \) labels, thus producing a graph with one loop.

Therefore we represent a correlation function \( W^g_k(p, p_1, \ldots, p_{k-1}) \) of genus \( g \) by graphs with loops \( t^g \) that are obtained by successive applications of the \( i \leftrightarrow i+1 \) operation.

**Definition 4.** A correlation function \( W^g_k(p, p_1, \ldots, p_{k-1}) \) of genus \( g \) and Euler characteristic \( \chi = 2 - 2g - k \) is represented by a sum of all different graphs with loops \( t^g \in (Y^n)^g \) for \( n = -\chi \):

\[
\psi(W^g_k(p, p_1, \ldots, p_{k-1})) = \sum_{t^g \in (Y^n)^g} t^g.
\]

**Remark 2.** Two graphs \( (t)^g, (t')^g \in (Y^n)^g \) are considered different in the obvious way. Either the underlying binary trees \( t, t' \in Y^n \) are distinct as base elements of \( k[Y^n] \) or the tree \( t \) has a pair of leaves, say \( (i, i+1) \) that are identified with an edge in \((t)^g\), producing a loop, and are free in \((t')^g \) (and reciprocally, because the two graphs have the same genus).

In particular \( W^1_1(p) \) is represented by a single graph with one loop denoted by \((1)^1\) whose underlying planar binary tree is \((1)\). More generally we have:

**Proposition 2.** The second summand of the topological recursion formula for the correlation function \( W^g_k(p, p_1, \ldots, p_{k-1}) \) with \( \chi = -n \) is represented by the sum

\[
\sum_{t^1 \in (Y^n)^g} (t^1)^1 \vee t_2 + \sum_{t^2 \in (Y^n)^g} t_1 \vee (t_2)^2
\]
where each underlying planar binary tree of order $n$ is decomposed as $t = t_1 \lor t_2$, with $|t_1| = p$, $|t_2| = q$, $p + q + 1 = n$.

**Proof.** For $n = 2$ the underlying trees of $W^2(p, p)$ are the same as of $W^2(p, p_1, p_2, p_3)$, namely $t = (12) + (21)$. Remembering that $(12) = (1) \lor |$ and $(21) = | \lor (1)$ and applying $\tau_{\gamma + 1}$ to (1) for $i = 0$ (27) is the same as

$$(1)^i \lor | + | \lor (1)^i$$

which is the image by $\psi$ of the second term of $W^2(p, p)$ in the topological recursion formula.

For arbitrary $n$ consider the trees $t' = t'_1 \lor t'_2$, $t' \in Y^{n-1}$ with $|t'_1| = a$, $|t'_2| = b$, $a + b + 1 = n - 1$ whose sum represents solutions $W^n_{n+1}(p, p_1, \ldots, p_n)$ by the induction hypothesis. By (7) and (8) we have

$$(1) \star t' = | \lor t' + (1 \star t'_1) \lor t'_2.$$  

(22)

Noting that $t'' = (1) \star t'_1$ are at most of order $n - 1$ and at least of order 1, and identifying pairs of leaves on the same branches by applying $\tau_{\gamma + 1}$ to each component of the grafting operation, the last formula gives

$$(\tau_{\gamma + 1})_{\text{same branches}} (1) \star t' = | \lor (t''_1 \lor t'_2) \lor t''_2 \lor (t''_2)$$

(23)

for each pair of leaves $(i, i+1)$ on the left or right branches. To collect all different terms produced in this way is equivalent to summing over $t'_1 \in Y^a$, also over $t''_1 \in Y^b$ and $(t''_2)_i \in Y^b$ for $0 \leq a \leq n - 2$, and finally over $t'_2 \in Y^b$ for $0 \leq b < n - 2$. Then the sum of all different terms given by (23) is

$$\sum_{i=0}^{n-1} (\tau_{\gamma + 1})_{\text{same branches}} (1) \star t' = \sum_{t''_1 \in Y^{n-1}} | \lor t''_1$$

$$+ \sum_{t''_1} (t''_1 \lor t'_2) \lor t''_2 + \sum_{t''_1} t''_2 \lor (t''_2)$$

$$= \sum_{t''_1} t''_1 \lor t'_2 + \sum_{t''_1} t''_2 \lor t'_2$$

(24)

where now $t'_1$ and $t'_2$ are the left and right branches of a tree $t = t'_1 \lor t'_2$ of order $n$. Note that on the right the first sum starts at $p = 1$ which is the lowest possibility for a $g = 1$ graph. Then the highest value of $q$ is $n - 2$. For the same reason, the second sum on the right starts at $q = 1$ which implies that $p \leq n - 2$. Translating (24) to the topological recursion we have

\[\star \quad (25)\]
which is the second term of the topological recursion formula for $W_n^j(p, p, \ldots, p_{n-1})$.

Now we consider the first term in the topological recursion which for $W_n^j(p, p, \ldots, p_{n-1})$ is

$$K_p(q, \bar{q}) W_{n+1}^0(q, \bar{q}, p_1, \ldots, p_{n-1}).$$

We start with a definition.

**Definition 5.** The ungrafting operation $\lor$ is defined by removing from $t$ the tree (1) that contains the root producing a forest with two trees $t_1$ and $t_2$. When $t$ represents an instance of a correlation function then the roots of $t_1$ and $t_2$ are labeled by $q$ and $\bar{q}$ and, as before, the tree (1) represents $K_p(q, \bar{q})$.

**Remark 3.** The operations $\lor$ and $\lor$ are similar to the operations $B^+$ and $B^-$ of the Connes–Kreimer Hopf algebra described, for instance, in [6].

If we start with the planar binary tree $t \in Y^n$ with leaves labels $p_1, \ldots, p_{n+1}$ and root label $p$ and identify two nearest neighbor leaves in opposite branches then we get a 1-loop graph $t^i \in (Y^n)^+$ with a relabeling $p_1, \ldots, p_{n-1}$. Then, by applying $\lor$, we get another tree $t'$ with two more edges with labels $q$ and $\bar{q}$ besides the leaves labelled by $p_1, \ldots, p_{n-1}$. This tree is isomorphic as a graph to a planar binary tree in $Y^{n-1}$ that we also denote by $t'$ by promoting the edge with the label $q$ to the root and the other edge to the rightmost leaf. See figure 7 for
an example with $W^1_n(p, p_1, p_2)$. In this way we get a representation of (26) by summing over all planar binary trees $t' \in Y^{n-1}$. It is clear that what was left to be done in (23) was the identification of two consecutive leaves in opposite branches. In principle there are several graphs $t^i \in (Y^n)^i$ of this type. The first one corresponds to the first term in (22) which gives a 1-loop graph with no leaves on the left branch. All other $t'$ come from the second term of that formula and depend on the type of the left branch $t'_i$ of the decomposition of the tree $t' \in Y^{n-1}, t' = t'_i \vee t'_2$. If $t'_i = t'$ then $(1) \ast t'_i = (1)$ and the identification of leaves gives a $t^i$ with a single leaf on the left branch. The next case is $(1) \ast (1) = (12) + (21)$ and this gives a sum of two one-loop graphs with two leaves on the left branch. The procedure continues until $t'_2 = t'$ which then gives a 1-loop graph with no leaves on the right branch. Thus we have proved the following proposition

**Proposition 3.** The representation of the first term of the topological recursion formula for $W^1_n(p, p_1, \ldots, p_{n-1})$ is given by the identification of leaves on opposite branches of the decomposition $t = t_1 \vee t_2$ with $t \in Y^n$, $t_1 \in Y^p$, $t_2 \in Y^q$, $p + q + 1 = n$:

$$
\sum_{\tau \in \text{permutations}} t_1 \vee t_2.
$$

The obvious notation $\vee$ means that two consecutive leaves in opposite branches are identified.

Therefore we have exhausted all possibilities of obtaining 1-loop graphs from planar binary trees of order $n$ and the two previous propositions imply the following theorem.

**Theorem 2.** The $n$ order solution $W^1_n(p, p_1, \ldots, p_{n-1})$ of the topological recursion in genus 1 is given by $(1) \ast (1) \ast \ldots \ast (1)$, with $n$ factors, followed by the identification of pairs of nearest neighbor leaves producing 1-loop graphs and finally by summing over all permutations of leaves labels $p_1, p_2, \ldots, p_{n-1}$:

$$
\psi \left( W^1_n(p, p_1, \ldots, p_{n-1}) \right)
= \sum_{\text{perm. of } \{p_1, \ldots, p_{n-1}\}} \sum_{i=0}^{n-1} \sum_{\gamma_i+1} ((1) \ast (1) \ast \ldots \ast (1)).
$$

Figure 8. A symmetric $2g + 1$-loop graph with $t^g$ a $g$-loop graph. The graph $t^g$ is the reflection of $t^g$ on the vertical axis that passes through the root. After the ungrafting operation the first graph is obtained by exchanging $q$ and $\bar{q}$ on the original graph.
Next we prove a simple lemma regarding symmetric graphs as in figure 8. Note that the resulting ungrafted graphs have the left–right order of the right branch of the original graph exchanged:

**Lemma 1.** If a graph that enters in the representation of the correlation function $W_{k+1}^2(p, K)$ has nearest neighbor leaves identified in different branches and is symmetric with respect to the vertical axis that passes through the root then it has a weight factor of 2, that is, it appears twice in the complete graph representation of $W_{k+1}^2(p, K)$.

**Proof.** First note that such a graph has an even number of leaves, say $2a$, possibly 0. After being ungrafted the resulting graph represents the following term in the topological recursion formula:

$$K_p(q, \tilde{q})W_{2a+2}^g(q, \tilde{q}, K) = K_p(q, \tilde{q})K_q(q_1, \tilde{q}_1)\left(W_{a+2}^g(q_1, q, L)W_{a+1}^g(\tilde{q}_1, M) + W_{a+1}^g(\tilde{q}_1, L)W_{a+2}^g(q_1, \tilde{q}, M) + \ldots\right)$$

for $|L| = |M| = a$, where the dots represent other terms that are not symmetric and that are represented by other graphs. This shows that the same graph after being ungrafted originates two symmetrical graphs and, as such, it has a weight factor of 2 when counting the number of graphs in the complete representation.

**Example 2.** See figure 9 for the graph representation of $W_3^2(p, p_1, p_2)$ with a graph of weight 2.

A simple but important fact is that $r^{-\gamma_{l+1}}$ acts as a derivation when applied independently to the left and right branches of a tree $t = t_1 \cup t_2$. This is apparent in proposition 2. However, when summing over all graphs we must take care with overcounting. If we start with $t^l = (t_1)^l \cup t_2 + t_1 \cup (t_2)^l$ with $|t| = n$, $|t_1| = p$, $|t_2| = q$, $p + q + 1 = n$ and apply $r^{-\gamma_{l+1}}$ to the two branches independently and sum over all different graphs we get

$$\sum_{i=0}^{n-1} (r^{-\gamma_{l+1}})_{\text{same branches}} \left\{ \sum_{\gamma t_1 \subset (r^p, t_2) \subset t \gamma} (t_1)^l \cup t_2 + \sum_{\gamma t_1 \subset (r^p, t_2) \subset t \gamma} 2(t_1)^l \cup (t_2)^l \right. $$

$$+ 2t_1 \cup (t_2)^l \right\}$$

whenever the operation is well defined. The reasoning for the two factors is the double counting of identical graphs. For instance, if $(t_1)^l$ has a loop starting at leaf 2 and $(t_1)^2$ was

3 The operation is not well defined if there is only one leaf available before and/or after a certain loop or if there are no more leaves to contract. In this case we set to 0 the result of acting with $r^{-\gamma_{l+1}}$.
obtained by producing a second loop starting at leaf 0, then this $(t_i)^2$ is identical to the 2-loop graph that was obtained by producing a second loop at leaf 2 in $(t_i)$ that already had a loop starting at leaf 0.

Figure 9. $W^i_j(p, p_1, p_2)$. The root label $p$ and the leaf labels $\{p_1, p_2\}$ are omitted.
Therefore the sum of different 2-loop graphs obtained from all planar binary trees \( t = t_1 \lor t_2 \) is

\[
\sum_{(t)^2 \in (Y)^2} (t)^2 = \sum_{(q)^2 \in \frac{I^2}{\frac{I^2}{\frac{I}{I}}}} (q)^2 \lor t_2 \\
+ \sum_{(p)^2 \in \frac{I^2}{\frac{I^2}{\frac{I}{I}}}} (t_1)^3 \lor (t_2)^3 \\
+ \sum_{(n)^2 \in \frac{I^2}{\frac{I^2}{\frac{I}{I}}}} t_1 \lor (t_2)^2.
\]

(31)

As for the identification of two consecutive leaves in separate branches, we start again from \( t^1 = (t_1)^1 \lor t_2 + t_1 \lor (t_2)^1 \) to get

\[
t^2 = (t_1)^1 \lor t_2 + t_1 \lor (t_2)^1.
\]

(32)

The fact that it may not always be possible to contract two leaves in opposite branches is important for counting the dimensions of the vector spaces \( k \left[ (Y^n)^g \right] \) generated by graphs with \( g \) loops but we will not consider this.

It is clear that we can continue this procedure and generate graphs \( t^g \in (Y^n)^g \) with an increasing number of \( g \) loops and \( k \) labels (including the root) up to the consistence of the relation \(-n = 2 - 2g - k\). Hence we have the following proposition.

**Proposition 4.** The graphs with loops \( t^g \in (Y^n)^g \) and \( k \) labels, including the root, that are compatible with \(-n = 2 - 2g - k\) are obtained from planar binary trees \( t = t_1 \lor t_2 \) by successive applications of \( g \) loops on the same branches and \( g \) loops on the opposite branches to all \( t \in Y^n \):

\[
\sum_{(t)^g \in (Y)^g} (t)^g = \sum_{k=0}^{g} \sum_{p+q+1=n} (t_1)^k \lor (t_2)^{g-k} \\
+ \sum_{k=0}^{g} \sum_{p+q+1=n} (t_1)^{g-1-k} \lor (t_2)^k.
\]

(33)

**Proof.** It follows from the discussion above and a simple inductive argument. The first sum on the right is a consequence of the fact that \((e^{-g+1})_{\text{br.}}\) acts as a derivation, after taking into account double counting and summing over all different graphs, and the second sum results from applying \((e^{-g+1})_{\text{opp.br.}}\), to the \((g-1)\)-loop graph \( t^{g-1} \) without leaves from opposite branches identified. To see that any graph \( t^g \) in the sum on the left can be obtained in this way just take the same graph but with some pair of leaves free, say \((i, i + 1)\). This is a \( t^{g-1} \) graph that admits a decomposition \((t_1)^k \lor (t_2)^{g-1-k}\) or \((t_1)^k \lor (t_2)^{g-2-k} \). In the first case, if \((i, i + 1)\) belong to the left or right branches, apply \((e^{-g+1})_{\text{br.}}\) to get an element in the first sum on the right of (33). If \((i, i + 1)\) belong to opposite branches apply \((e^{-g+1})_{\text{opp.br.}}\) to get an element of the second sum. The second case works in a similar way except that \((i, i + 1)\) must belong to the same branches. \(\square\)
Theorem 3. The order solution $W^g_{2g-2+n}$ of the topological recursion in genus $g > 0$ and $k = 2 - 2g + n > 0$ variables is given by $(1) * (1) * \ldots * (1)$, with $n$ factors, followed the identification of pairs of nearest neighbor leaves producing graphs with loops as in proposition 4 and finally by summing over all permutations of $p_1, p_2, \ldots, p_{1-2(g+n)}$.

Proof. The proof now follows easily from proposition 4. Remember that the sum of all planar binary trees of order $n$ is obtained by $n$ factors of $(1)$ with the $*$ product. By ungrafting the graphs given by (33) and using the representation map $\psi$ we get

$$W^g(p, p_1, \ldots, p_{k-1}) = \psi_* \left( \sum_{\text{perm. of leaf labels } K = \{ p_1, \ldots, p_{k-1} \}} (t)^g \right)$$

$$= \sum_{h=0}^{g} \sum_{\text{perm. of } K} K_p(q, \bar{q}) \left( W_{g+1}^{g-1}(q, \bar{q}, K) + W_{g+1}^h(q, \bar{q}, L) W_{g+1}^{g-h}(\bar{q}, M) \right)$$

which is the topological recursion formula for arbitrary genus.

Remark 4. After being ungrafted the second term in (33) can still have leaves in opposite branches identified. In general this happens if $(t)^g = (t_1)^{g_1} \lor (t_2)^{g_2}$ with $g_1 + g_2 = g - 1$ and $(t_2)^{g_2}$ has a decomposition $(t_2)^{g_2'} \lor (t_2)^{g_2''}$ with $g_2' + g_2'' = g - 1$.

Example 3. See figure 10 for the graph representation of $W^2_{f}(p)$.

5. The antipode

In a graded connected Hopf algebra there is a canonical antipode $S$ whose expression is given by the convolution inverse of the identity:
with $m$ the product, $\Delta$ the co-product, $\eta$ the unit and $\epsilon$ the co-unit. Explicitly, in the Loday–Ronco Hopf algebra, we have

$$S(t) = -t - S(t_1) \ast t_2$$

where in Sweedler notation

$$\Delta t = \sum t_1 \otimes t_2$$

is the co-product in $k[Y^\infty]$ induced by (6). For instance, $S(1) = -1$ because $1$ is primitive and $S(12) = (21)$ and also $S(21) = (12)$. This suggests that a map $\psi^S$ induced by the antipode on the vector space of correlation functions should give

$$(\psi^S)(W_0^0) = W_2^0.$$ 

More generally, from $S((1)) = -(1)$ we see that

$$S((1) \ast (1) \ast \ldots \ast (1)) = (-1)^n (1) \ast (1) \ast \ldots \ast (1)$$

with $n$ factors in the $\ast$ product and then

$$(\psi^S)(W_0^n) = (-1)^n W_0^n.$$ 

Since $n$ is identified with the Euler characteristic we see that the induced map respects the grading of $k[Y^\infty]$ for $g = 0$. For the moment it is not clear how to extend this simple computation to the case of graphs with loops. The tree $|$ being the identity in $k[Y^\infty]$ and representing $W_2^0$ has trivially $S(|) = |$, but it is not clear if the change of topology from a tree to a graph with loops shouldn’t change dramatically the Hopf algebra structure, or if even the full algebra of these class of graphs with loops of arbitrary order is yet a Hopf algebra. There are examples of Hopf algebra of general graphs that are well documented in the literature (see for instance [16, 18]) but we do not know at the moment if they can be adapted to this framework. In particular, the Euler characteristic of a graph has a different meaning to the one used here.

**6. Discussion**

We have seen that an extension of the Hopf algebra planar binary trees described by Loday and Ronco provides a representation of a vector space, whose nature is still to be clarified, generated by the set of correlation functions typical of matrix models and that satisfy the recursion formula discovered by Eynard and Orantin. This extension, obtained by identifying nearest neighbor leaves through a single edge, is only necessary for $g > 0$. This procedure moves from planar binary trees to planar graphs with loops, binary in the internal vertices in the sense that each internal vertex has two children. In the process of showing that this class of graphs satisfies the Eynard–Orantin full recursion formula, we have provided an explicit formula for the solutions that are obtained first by obtaining the sum of all planar binary trees of order $n$ through the $\ast$ product of $n$ factors of $(1)$ computed in $k[Y^\infty]$, and then by connecting two consecutive leaves in all possible ways up to the genus $g$.

**References**

[1] Aguiar M and Sottile F 2006 Structure of the Loday-Ronco Hopf algebra of trees *J. Algebra* **295** 473–511
[2] Atiyah M F 1988 Topological quantum field theory *Publications Mathématiques de l'IHÉS* **68** 175–86

[3] Brouder C 2000 Runge–Kutta methods and renormalization *Eur. Phys. J. C* **12** 521–34

[4] Brouder C 2000 On the trees of quantum fields *Eur. Phys. J. C* **12** 535–49

[5] Collins J C 1984 *Renormalization: An Introduction to Renormalization, the Renormalization group and the Operator-Product Expansion* (University Press: Cambridge)

[6] Connes A. and Kreimer D., *Hopf algebras, renormalization and noncommutative geometry,* Quantum field theory: perspective and prospective (Les Houches, 1998) NATO Sci. Ser. C Math. Phys. Sci., vol 530 (Dordrecht: Kluwer) pp 59–108

[7] Connes A and Kreimer D 2000 Renormalization in quantum field theory and the riemann-hilbert problem: I. The Hopf algebra structure of graphs and the main theorem *Comm. Math. Phys.* **210** 249–73

[8] Connes A and Kreimer D 2001 Renormalization in quantum field theory and the Riemann-Hilbert problem: II. The β-function, diffeomorphisms and the renormalization group *Comm. Math. Phys.* **216** 215–41

[9] di Francesco P, Ginsparg P H and Zinn-Justin J 1995 2-D gravity and random matrices *Phys. Rep* **254** 1–133

[10] Dumitrescu O, Mulase M, Safruk B and Sorkin A 2013 The spectral curve of the eynard-orantin recursion via the laplace transform, algebraic and geometric aspects of integrable systems and random matrices, contemp. math *Amer. Math. Soc.* **593** 263–315

[11] Eynard B and Orantin N 2007 Invariants of algebraic curves and topological expansion *Commun. Number Theory Phys.* **1** 347–452

[12] Foissy L 2002 Les algèbres de Hopf des arbres enracinés décorés I *Bull. Sci. Math.* **126** 193–239

[13] Foissy L 2002 Les algèbres de Hopf des arbres enracinés décorés II *Bull. Sci. Math.* **126** 249–88

[14] Frabetti A 2001 Simplicial properties of the set of planar binary trees *J. Algebraic Combin.* **13** 41–65

[15] Gelfand I M, Krob D, Lascoux A, Leclerc B, Retakh V S and Thibon J-Y 1995 Noncommutative symmetric functions *Adv. Math.* **112** 218–348

[16] Humpert B and Martin J L 2012 The incidence Hopf algebra of graphs *SIAM J. Discrete Math.* **26** 555–70

[17] Loday J-L and Ronco M O 1998 Hopf algebra of the planar binary trees *Adv. Math.* **139** 293–309

[18] Schmitt W R 1994 Incidence Hopf algebras *J. Pure Appl. Algebra* **96** 299–330