# A SIMPLE MODEL OF TREES FOR UNICELLULAR MAPS

GUILLAUME CHAPUY, VALENTIN FERAY, AND ÉRIC FUSY

Abstract. We consider unicellular maps, or polygon gluings, of fixed genus. In FPSAC '09 the first author gave a recursive bijection transforming unicellular maps into trees, explaining the presence of Catalan numbers in counting formulas for these objects. In this paper, we give another bijection that explicitly describes the "recursive part" of the first bijection. As a result we obtain a very simple description of unicellular maps as pairs made by a plane tree and a permutation-like structure. All the previously known formulas follow as an immediate corollary or easy exercise, thus giving a bijective proof for each of them, in a unified way. For some of these formulas, this is the first bijective proof, e.g. the Harer-Zagier recurrence formula, or the Lehman-Walsh/Goupil-Schaeffer formulas. Thanks to previous work of the second author this also leads us to a new expression for Stanley character polynomials, which evaluate irreducible characters of the symmetric group.

## 1. Introduction

A unicellular map is a connected graph embedded in a surface in such a way that the complement of the graph is a topological disk. These objects have appeared frequently in combinatorics in the last forty years, in relation with the general theory of map enumeration, but also with the representation theory of the symmetric group, the study of permutation factorizations, or the study of moduli spaces of curves. All these connections have turned the enumeration of unicellular maps into an important research field (for the many connections with other areas, see [9] and references therein; for an overview of the results see the introductions of the papers [3, 1]). The main results in the domain can be roughly separated in two families.

The first family deals with *colored* maps (maps endowed with an application from its vertex set to a set of $q$ colors). This implies "summation" enumeration formulas (see [8, 14, 11] or paragraph 3.4 below). These formulas are often elegant, and different combinatorial proofs for them have been given in the past few years [10, 6, 14, 11, 1]. The issue is that some important topological information, such as the genus of the surface, is not apparent in these constructions.

Formulas of the second family keep track explicitly of the genus of the surface. This includes inductive relations (like the Harer-Zagier recurrence formula [8]) or explict (but quite involved) closed forms (Lehman-Walsh [15] and Goupil-Schaeffer [7] formulas). From a combinatorial point of view, these formulas are harder to understand. A step in this direction was done by the first author in [3] (this construction is explained in subsection 2.2), which led to new inductions relations and to new formulas. However the link with other formulas in the second family remained

The first and third authors are partially supported by ERC grant StG 208471 – ExploreMaps. The second author is partially supported by ANR project PSYCO.

mysterious, and [3] left open the problem of finding combinatorial proofs of these formulas.

The goal of this paper is to present a new bijection between unicellular maps and surprisingly simple objects which we call *C-decorated* trees (these are merely plane trees equipped with a certain kind of permutation on their vertices). This bijection is based on the previous work of the first author [3]: we explicitly describe the "recursive part" appearing in this work. As a consequence, not only can we reprove all the aforementioned formulas in a bijective way, thus giving the first bijective proof for several of them, but we do that in unified way. Indeed, C-decorated trees are so simple combinatorial objects that all formulas follow from our bijection as an immediate corollary or easy exercise.

Another interesting application of this bijection is a new explicit way of computing the so-called Stanley character polynomials, which are nothing but the evaluation of irreducible characters of the symmetric groups, properly normalized and parametrized. Indeed, in a previous work [4], the second author expressed these polynomials as a generating function of (properly weighted) unicellular maps. Although we do not obtain a "closed form" expression (there is no reason to believe that such a form exists!), we express Stanley character polynomials as the result of a term-substitution in free cumulants, which are another meaningful quantity in representation theory of symmetric groups.

## 2. The main bijection

**2.1. Unicellular maps and C-decorated trees.** A *map* $M$ of genus $g \geq 0$ is a connected graph $G$ embedded on a closed compact oriented surface $S$ of genus $g$, such that $S \backslash G$ is a collection of topological disks, which are called the *faces* of $M$. Loops and multiple edges are allowed. The graph $G$ is called the *underlying graph* of $M$ and $S$ its *underlying surface*. Two maps that differ only by an oriented homeomorphism between the underlying surfaces are considered the same. A corner of $M$ is the angular sector between two consecutive edges around a vertex. A *rooted map* is a map with a marked corner, called the *root*; the vertex incident to the root is called the *root-vertex*. From now on, all maps are assumed to be rooted (note that the underlying graph of a rooted map is naturally vertex-rooted). A *unicellular map* is a map with a unique face. The classical Euler relation $|V|-|E|+|F| = 2-2g$ ensures that a unicellular map with $n$ edges has $n+1-2g$ vertices. A *plane tree* is a unicellular map of genus 0.

A *rotation system* on a connected graph $G$ consists in a cyclic ordering of the half-edges of $G$ around each vertex. Given a map $M$, its underlying graph $G$ is naturally equipped with a rotation system given by the *clockwise ordering* of half-edges on the surface in a vicinity of each vertex. It is well-known that this correspondence is 1-to-1, i.e. a map can be considered as a connected graph equipped with a rotation system (thus, as a purely combinatorial object). We will take this viewpoint from now on.

A *cycle-signed* permutation is a permutation where each cycle carries a sign, either $+$ or $-$. A *C-permutation* is a cycle-signed permutation where all cycles have odd length, see Figure 1(a). For each C-permutation $\sigma$ on $n$ elements, the *rank* of $\sigma$ is defined as $r(\sigma) = n - \ell(\sigma)$, where $\ell(\sigma)$ is the number of cycles of $\sigma$. Note that $r(\sigma)$ is even since all cycles have odd length. The *genus* of $\sigma$ is defined as $r(\sigma)/2$. A *C-decorated tree* on $n$ edges is a pair $\gamma = (T, \sigma)$ where $T$ is a plane tree

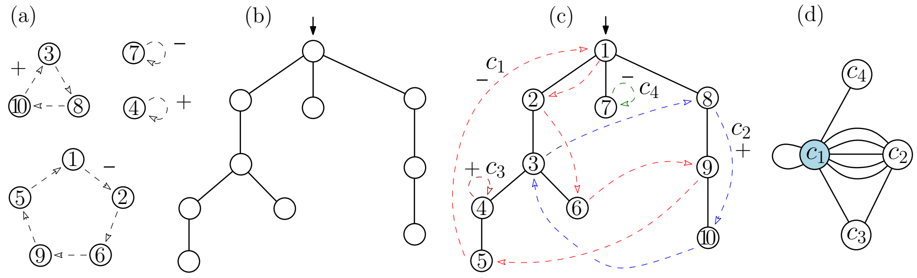


FIGURE 1. (a) A C-permutation $\sigma$. (b) A plane tree $T$. (c) The C-decorated tree $(T, \sigma)$. (d) The underlying graph of $(T, \sigma)$.

with $n$ edges and $\sigma$ is a C-permutation of $n+1$ elements. The *genus* of $\gamma$ is defined to be the genus of $\sigma$. Note that the $n+1$ vertices of $T$ can be canonically numbered from 1 to $n+1$ (e.g., following a left-to-right depth-first traversal), hence $\sigma$ can be seen as a permutation of the vertices of $T$, see Figure 1(c). The *underlying graph* of $\gamma$ is the (vertex-rooted) graph $G$ obtained from $T$ by merging into a single vertex the vertices in each cycle of $\sigma$ (so that the vertices of $G$ correspond to the cycles of $\sigma$), see Figure 1(d).

**Definition 1.** *For $n, g$ nonnegative integers, denote by $\mathcal{E}_g(n)$ the set of unicellular maps of genus $g$ with $n$ edges; and denote by $\mathcal{T}_g(n)$ the set of C-decorated trees of genus $g$ with $n$ edges.*

For $\mathcal{A}$ a finite set, $k\mathcal{A}$ denotes the set made of $k$ disjoint copies of $\mathcal{A}$. For two finite sets $\mathcal{A}$ and $\mathcal{B}$, we write $\mathcal{A} \simeq \mathcal{B}$ if there is a bijection between $\mathcal{A}$ and $\mathcal{B}$. Our main result will be to show that $2^{n+1}\mathcal{E}_g(n) \simeq \mathcal{T}_g(n)$, with a bijection which preserves the underlying graphs of the objects.

2.2. **Recursive decomposition of unicellular maps.** In this section, we briefly recall a combinatorial method developed in [3] to decompose unicellular maps.

**Proposition 1** (Chapuy [3])**.** *For $k \geq 1$, denote by $\mathcal{E}_g^{(2k+1)}(n)$ the set of maps from $\mathcal{E}_g(n)$ in which a set of $2k+1$ vertices is distinguished. Then for $g > 0$ and $n \geq 0$,*

$$2g\,\mathcal{E}_g(n) \simeq \mathcal{E}_{g-1}^{(3)}(n) + \mathcal{E}_{g-2}^{(5)}(n) + \mathcal{E}_{g-3}^{(7)}(n) + \cdots + \mathcal{E}_0^{(2g+1)}(n). \tag{1}$$

*In addition, if $M$ and $(M', S')$ are in correspondence, then the underlying graph of $M$ is obtained from the underlying graph of $\gamma'$ by merging the vertices in $S'$ into a single vertex.*

We now sketch briefly the construction of [3]. Although this is not really needed for the sequel, we believe that it gives a good insight into the objects we are dealing with (readers in a hurry may take Proposition 1 for granted and jump directly to subsection 2.3). We refer to [3] for proofs and details.

We first explain where the factor $2g$ comes from in (1). Let $M$ be a rooted unicellular map of genus $g$ with $n$ edges. Then $M$ has $2n$ corners, and we label them from 1 to $2n$ incrementally, starting from the root, and going clockwise around the (unique) face of $M$ (Figure 2). Let $v$ be a vertex of $M$, let $k$ be its degree, and let $(a_1, a_2, \ldots, a_k)$ be the sequence of the labels of corners incident to it, read in

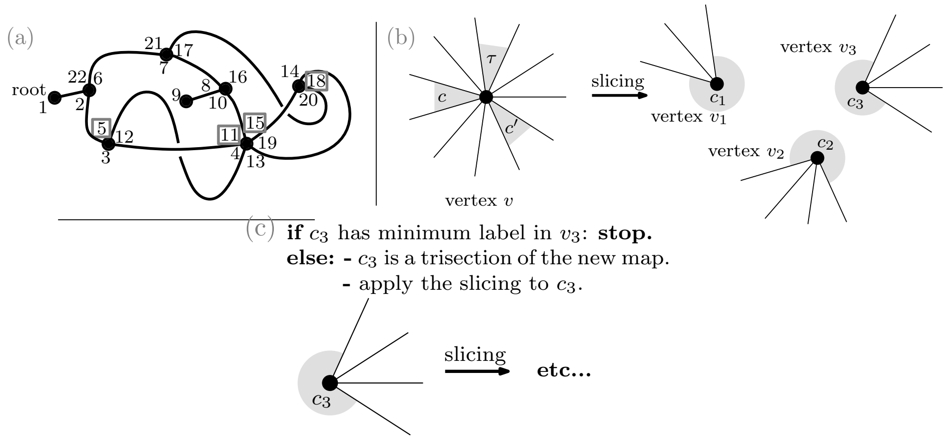


FIGURE 2. (a) A unicellular map of genus 2 equipped with its corner labelling. Labels corresponding to trisections are boxed. (b) Given a trisection $\tau$, two other corners of interest $c$ and $c'$ are canonically defined (see text). “Slicing the trisection” then gives rise to three new vertices $v_1, v_2, v_3$, with distinguished corners $c_1, c_2, c_3$. (c) The recursive procedure of [3]: if $c_3$ is the minimum corner of $v_3$, then stop; else, as shown in [3], $c_3$ is a trisection of the new map $M'$: in this case, iterate the slicing operation on $(M', c_3)$.

*counterclockwise direction around* $v$ starting from the minimal label $a_1 = \min a_i$. If for some $j \in [\![1, k-1]\!]$, we have $a_{j+1} < a_j$, we say that the corner of $v$ labelled by $a_{j+1}$ is a *trisection of* $M$. Figure 2(a) shows a map of genus two having four trisections. More generally we have:

**Lemma 2** ([3])**.** *A unicellular map of genus $g$ contains exactly $2g$ trisections. In other words, the set of unicellular maps of genus $g$ with $n$ edges and a marked trisection is isomorphic to $2g\,\mathcal{E}_g(n)$.*

Now, let $\tau$ be a trisection of $M$ of label $l(\tau)$, and let $v$ the vertex its belongs to. We denote $c$ the corner of $v$ with minimum label and $c'$ the corner with minimum label among those which appear between $c$ and $\tau$ clockwise around $v$ and whose label is greater than $l(\tau)$. By definition of a trisection, $c'$ is well defined. We then construct a new map $M'$, by *slicing* the vertex $v$ into three new vertices using the three corners $c, c', \tau$ as on Figure 2(b). We say that the map $M'$ is obtained from $M$ by *slicing the trisection* $\tau$. As shown in [3], the new map $M'$ is a unicellular map of genus $g-1$. We can thus relabel the $2n$ corners of $M'$ from 1 to $2n$, according to the procedure we already used for $M$. Among these corners, three of them, say $c_1, c_2, c_3$ are naturally inherited from the slicing of $v$, as on Figure 2(b). Let $v_1, v_2, v_3$ be the vertices they belong to, respectively. Then the following is true [3]: *In the map $M'$, the corner $c_i$ has the smallest label around the vertex $v_i$, for $i \in \{1, 2\}$. For $i = 3$, either the same is true, or $c_3$ is a trisection of the map $M'$.*

We now finally describe the bijection promised in Proposition 1. It is defined recursively on the genus, as follows. Given a map $M \in \mathcal{E}_g(n)$ with a marked trisection $\tau$, let $M'$ be obtained from $M$ by the slicing of $\tau$, and let $c_i, v_i$ be defined as above for $i \in \{1,2,3\}$. If $c_3$ has the minimum label in $v_3$, set $\Psi(M,\tau) := (M', \{v_1, v_2, v_3\})$, which is an element of $\mathcal{E}_{g-1}^{(3)}(n)$. Else, let $(M'', S) = \Psi(M', c_3)$, and set $\Psi(M,\tau) := (M'', S \cup \{v_1, v_2\})$. Note that this recursive algorithm necessarily stops, since the genus of the map decreases and since there are no trisections in unicellular maps of genus 0 (plane trees). Thus this procedure yields recursively a mapping that associates to a map $M$ with a marked trisection $\tau$ another map $M''$ of a smaller genus, with a set $S''$ of marked vertices (namely the set of vertices which have been involved in a slicing at some point of the procedure). The set $S''$ of marked vertices necessarily has odd cardinality, as easily seen by induction. Moreover, it is clear that the underlying graph of $M$ coincides with the underlying graph of $M''$ in which the vertices of $S''$ have been identified together into a single vertex. One can show that $\Psi$ is a bijection by constructing explicitly the inverse mapping [3].

2.3. **Recursive decomposition of C-decorated trees.** We now propose a recursive method to decompose C-decorated trees, which can be seen as parallel to the decomposition of unicellular maps given in the previous section. Denote by $\mathcal{C}(n)$ (resp. $\mathcal{C}_g(n)$) the set of C-permutations on $n$ elements (resp. on $n$ elements and of genus $g$). A *signed sequence* of integers is a pair $(\epsilon, S)$ where $S$ is an integer sequence and $\epsilon$ is a sign, either $+$ or $-$.

**Lemma 3.** *Let $X$ be a finite non-empty set of positive integers. Then there is a bijection between signed sequences of distinct integers from $X$ —all elements of $X$ being present in the sequence— and C-permutations on the set $X$. In addition the C-permutation has one cycle if and only if the signed sequence has odd length and starts with its minimal element.*

*Proof.* Let $\gamma$ be a signed sequence, e.g. $\gamma = {}^{+}(4731562)$. If $\gamma$ has odd length and starts with its minimal element, return $\gamma$ seen as a unicyclic C-permutation (where the unique cycle is written sequentially). Otherwise cut $\gamma$ as $\gamma = \gamma_1\gamma_2$, where $\gamma_2$ starts with the minimal element in $\gamma$ (in our example, $\gamma_1 = {}^{+}(473)$ and $\gamma_2 = (1562)$). If $\gamma_2$ has odd length, then "produce" the signed cycle ${}^{+}\gamma_2$. If $\gamma_2$ has even length, move the second element of $\gamma_2$ to the end of $\gamma_1$, and "produce" the signed cycle ${}^{-}\gamma_2$. Then (in both cases), restart the same process on $\gamma = \gamma_1$, producing one (signed) cycle at each step, until $\gamma$ is odd and starts with its minimal element, in which case one produces $\gamma$ as the last signed cycle. (In our example, the signed cycles successively produced are ${}^{-}(162)$, ${}^{-}(3)$, and ${}^{+}(475)$.) The process clearly yields a collection of signed cycles of odd lengths, i.e., yields a C-permutation. The mapping is easy to invert (we omit details in this extended abstract), so it gives a bijection. □

An element of a C-permutation is called *non-minimal* if it is not the minimum in its cycle. Non-minimal elements play the same role for C-permutations (and C-decorated trees) as trisections for unicellular maps. Indeed, a C-permutation of genus $g$ has $2g$ non-minimal elements (compare with Lemma 2), and moreover we have the following analogue of Proposition 1:

**Proposition 4.** *For $k \geq 1$, denote by $\mathcal{T}_g^{(2k+1)}(n)$ the set of C-decorated trees from $\mathcal{T}_g(n)$ in which a set of $2k+1$ cycles is distinguished. Then for $g > 0$ and $n \geq 0$,*

$$2g\,\mathcal{T}_g(n) \simeq \mathcal{T}_{g-1}^{(3)} + \mathcal{T}_{g-2}^{(5)} + \mathcal{T}_{g-3}^{(7)} + \cdots .$$

*In addition, if $\gamma$ and $(\gamma', S')$ are in correspondence, then the underlying graph of $\gamma$ is obtained from the underlying graph of $\gamma'$ by merging the vertices corresponding to cycles from $S'$ into a single vertex.*

*Proof.* For $k \geq 1$ let $\mathcal{C}_g^{(2k+1)}(n)$ be the set of C-permutations from $\mathcal{C}_g(n)$ where a subset of $2k+1$ cycles are marked. Let $\mathcal{C}_g^{\circ}(n)$ be the set of C-permutations from $\mathcal{C}_g(n)$ where a non-minimal element is marked. Note that $\mathcal{C}_g^{\circ}(n) \simeq 2g\, \mathcal{C}_g(n)$ since a C-permutation in $\mathcal{C}_g(n)$ has $2g$ non-minimal elements. Moreover $\mathcal{C}_g^{\circ}(n) \simeq \sum_{k\geq 1} \mathcal{C}_{g-k}^{(2k+1)}(n)$ (apply Lemma 3 to the cycle —represented as a signed sequence— containing the marked non-minimal element; this produces a collection of $(2k+1) \geq 3$ signed cycles of odd length, which we take as the marked cycles). Hence $2g\, \mathcal{C}_g(n) \simeq \sum_{k\geq 1} \mathcal{C}_{g-k}^{(2k+1)}(n)$. Since $\mathcal{T}_g(n) = \mathcal{E}_0(n) \times \mathcal{C}_g(n+1)$, we conclude that $2g\, \mathcal{T}_g(n) \simeq \sum_{k\geq 1} \mathcal{T}_{g-k}^{(2k+1)}(n)$. The statement on the underlying graph just follows from the fact that the procedure in Lemma 3 merges the marked cycles into a unique cycle. □

## 2.4. **The main result.**

**Theorem 5.** *For each non-negative integers $n$ and $g$ we have*

$$2^{n+1}\mathcal{E}_g(n) \simeq \mathcal{T}_g(n).$$

*In addition the cycles of a C-decorated tree naturally correspond to the vertices of the associated unicellular map, in such a way that the respective underlying graphs are the same.*

*Proof.* The proof is a simple induction on $g$, whereas $n$ is fixed. The case $g = 0$ is obvious. Let $g > 0$. The induction hypothesis ensures that for each $g' < g$, $2^{n+1}\mathcal{E}_{g'}^{(2k+1)}(n) \simeq \mathcal{T}_{g'}^{(2k+1)}(n)$, where the underlying graphs (taking marked vertices vertices into account) of corresponding objects are the same. Hence, by Propositions 1 and 4, we have $2g\, 2^{n+1}\mathcal{E}_g(n) \simeq 2g\, \mathcal{T}_g(n)$, where the underlying graphs of corresponding objects are the same. Finally, one can extract from this $2g$-to-$2g$ correspondence a 1-to-1 correspondence, which still preserves underlying graphs: think of extracting a perfect matching from a $2g$-regular bipartite graph, which is possible according to Hall's marriage theorem. Hence $2^{n+1}\mathcal{E}_g(n) \simeq \mathcal{T}_g(n)$. □

2.5. **A fractional, or stochastic, formulation.** Even if this does not hinder enumerative applications to be detailed in the next section, we do not know of an effective (polynomial-time) way to implement the bijection of Theorem 5; indeed the last step of the proof is to extract a perfect matching from a $2g$-regular bipartite graph whose size is exponential in $n$.

What can be done effectively is a *fractional* formulation of the bijection. For a finite set $X$, let $\mathbb{C}\langle X\rangle$ be the set of linear combinations of the form $\sum_{x\in X} u_x \cdot x$, where the $x \in X$ are seen as independent formal vectors, and the coefficients $u_x$ are in $\mathbb{C}$. Let $\mathbb{R}_1^+\langle X\rangle \subset \mathbb{C}\langle X\rangle$ be the subset of linear combinations where the coefficients are nonnegative and add up to 1. Denote by $\mathbf{1}_X$ the vector $\sum_{x\in X} x$. For two finite sets $X$ and $Y$, a *fractional mapping* from $X$ to $Y$ is a linear mapping $\varphi$ from $\mathbb{C}\langle X\rangle$ to $\mathbb{C}\langle Y\rangle$ such that the image of each $x \in X$ is in $\mathbb{R}_1^+\langle Y\rangle$; the subset of elements of $Y$ with strictly positive coefficients in $\varphi(x)$ is called the *image-support* of $x$. Note that $\varphi(x)$ identifies to a probability distribution on $Y$; a "call to $\varphi(x)$" is meant as picking up $y \in Y$ under this distribution. A fractional mapping is *bijective* if $\mathbf{1}_X$ is mapped to $\mathbf{1}_Y$, and is *deterministic* if each $x \in X$ is mapped to some $y \in Y$. Note that, if there is a fractional bijection from $X$ to $Y$, then $|X| = |Y|$ (indeed in that case the matrix of $\varphi$ is bistochastic).

One can now formulate by induction on the genus an effective (the cost of a call is $O(gn)$) fractional bijection from $2^{n+1}\mathcal{E}_g[n]$ to $\mathcal{T}_g(n)$, and similarly from $\mathcal{T}_g[n]$ to $2^{n+1}\mathcal{E}_g(n)$. The crucial property is that, for $k \geq 1$ and $E$, $F$ finite sets, if there is a fractional bijection $\Phi$ from $kE$ to $kF$ then one can *effectively* derive from it a fractional bijection from $E$ to $F$: map each $x \in E$ to $\frac{1}{k}(\iota(\Phi(x_1)) + \cdots + \iota(\Phi(x_k)))$, where $x_1, \ldots, x_k$ are the representatives of $x$ in $kE$, and where $\iota$ is the projection from $kF$ to $F$. Hence by induction on $g$, Propositions 1 and 4 (where the stated combinatorial isomorphisms are effective) ensure that there is an effective fractional bijection from $2^{n+1}\mathcal{E}_g(n)$ to $\mathcal{T}_g[n]$ and similarly from $\mathcal{T}_g[n]$ to $2^{n+1}\mathcal{E}_g[n]$, such that if $\gamma'$ is in the image-support of $\gamma$ then the underlying graphs of $\gamma$ and $\gamma'$ are the same.

Note that, given an effective fractional bijection between two sets $X$ and $Y$, and a uniform random sampling algorithm on the set $X$, one obtains immediately a uniform random sampling algorithm for the set $Y$. In the next section, we will use our bijection to prove several enumerative formulas for unicellular maps, starting from elementary results on the enumeration of trees or permutations. In all cases, we will also be granted with a uniform random sampling algorithm for the corresponding unicellular maps, though we will not emphasize this point in the rest of the paper.

## 3. Counting formulas for unicellular maps

It is quite clear that C-decorated trees are much simpler combinatorial objects than unicellular maps. In this section, we use them to give bijective proofs of several known formulas concerning unicellular maps. We focus on the Lehman-Walsh and Goupil-Schaeffer formulas, and the Harer-Zagier recurrence, of which bijective proofs were long-awaited. We also sketch a bijective proof of the Harer-Zagier summation formula (prototype for a family of formulas for which bijective proofs were already known). We insist on the fact that all these proofs are elementary consequences of our main bijection (Theorem 5).

### 3.1. Two immediate corollaries.

The set $\mathcal{T}_g(n) = \mathcal{E}_0(n) \times \mathcal{C}_g(n+1)$ is the product of two sets that are easy to count. Precisely, let $\epsilon_g(n) = |\mathcal{E}_g(n)|$ and $c_g(n) = |\mathcal{C}_g(n)|$. Recall that $\epsilon_0(n) = \mathrm{Cat}(n)$, where $\mathrm{Cat}(n) := \frac{(2n)!}{n!(n+1)!}$ is the $n$th Catalan number. Therefore Theorem 5 gives $\epsilon_g(n) = 2^{-n-1}\mathrm{Cat}(n)c_g(n+1)$.

It is immediate to give for $c_g(n+1)$ a closed form (by summing over all possible cycle types) or an explicit generating series. This yields two classical results for the enumeration of unicellular maps.

For $\gamma = (\gamma_1, \ldots, \gamma_\ell) = 1^{m_1} \ldots k^{m_k}$ a partition of $g$, the number $a_\gamma(n+1)$ of permutations of $n+1$ elements with cycle-type equal to $1^{n+1-2g-\ell}3^{m_1} \ldots (2k+1)^{m_k}$ is classically given by the quotient $a_\gamma(n+1) = (n+1)!/((n+1-2g-\ell)!\prod_i m_i!(2i+1)^{m_i})$, and the number of C-permutations with this cycle-type is just $a_\gamma(n+1)2^{n+1-2g}$ (since each cycle has 2 possible signs). Hence, we get the equality $c_g(n+1) = 2^{n+1-2g}\sum_{\gamma \vdash g} a_\gamma(n+1)$. We thus recover:

**Proposition 6** (Walsh and Lehman [15]). *The number $\epsilon_g(n)$ is given by*

$$\epsilon_g(n) = \frac{(2n)!}{n!(n+1-2g)!2^{2g}} \sum_{\gamma \vdash g} \frac{(n+1-2g)_\ell}{\prod_i m_i!(2i+1)^{m_i}},$$

*where $(x)_k = \prod_{j=0}^{k-1}(x-j)$, $\ell$ is the number of parts of $\gamma$, and $m_i$ is the number of parts of length $i$ in $\gamma$.*

The exponential generating function $C(x,y) := \sum_{n,g} \frac{1}{(n+1)!} c_g(n+1) y^{n+1} x^{n+1-2g}$ of signed cycles of odd length ($y$ marks the number of elements, which are labelled, and $x$ marks the number of cycles) is

$$C(x,y) = \exp\Big(2x \sum_{k\geq 1} \frac{y^{2k+1}}{2k+1}\Big) - 1 = \exp\Big(x \log\Big(\frac{1+y}{1-y}\Big)\Big) - 1 = \Big(\frac{1+y}{1-y}\Big)^x - 1.$$

Since $c_0(1) = 2$ and $\frac{1}{(n+1)!} c_g(n+1) = \frac{2^{n+1} n!}{(2n)!} \epsilon_g(n) = \frac{2}{(2n-1)!!} \epsilon_g(n)$ for $n \geq 1$, we recover:

**Proposition 7** (Harer-Zagier series formula [8, 9]). *The generating function* $E(x,y) := 1 + 2xy + 2 \sum_{g\geq 0, n>0} \frac{\epsilon_g(n)}{(2n-1)!!} y^{n+1} x^{n+1-2g}$ *is given by*

$$E(x,y) = \Big(\frac{1+y}{1-y}\Big)^x.$$

3.2. **Harer-Zagier recurrence formula.** Elementary algebraic manipulations on the expression of $E(x,y)$ yield a very simple recurrence satisfied by $\epsilon_g(n)$, known as the Harer-Zagier recurrence formula (stated in Proposition 10 hereafter). We now show that the model of C-decorated trees makes it possible to derive this recurrence directly from a combinatorial isomorphism, that generalizes Rémy's beautiful bijection for plane trees [13].

It is convenient here to consider C-decorated trees as *unlabelled structures*: precisely we see a C-decorated tree as a plane tree where the vertices are partitioned into parts of odd size, where each part carries a sign + or −, and such that the vertices in each part are cyclically ordered (the C-permutation can be recovered by numbering the vertices of the tree according to a left-to-right depth-first traversal), think of Figure 1(c) where the labels have been taken out. We denote by $\mathcal{P}(n) = \mathcal{E}_0(n)$ the set of plane trees with $n$ edges, and by $\mathcal{P}^{\mathrm{v}}(n)$ (resp. $\mathcal{P}^{\mathrm{c}}(n)$) the set of plane trees with $n$ edges where a vertex (resp. a corner) is marked. Rémy's procedure, shown in Figure 3, realizes the isomorphism $\mathcal{P}^{\mathrm{v}}(n) \simeq 2\,\mathcal{P}^{\mathrm{c}}(n-1)$, or equivalently

$$(n+1)\mathcal{P}(n) \simeq 2(2n-1)\mathcal{P}(n-1). \tag{2}$$

Let $\mathcal{T}_g^{\mathrm{v}}(n)$ be the set of C-decorated trees from $\mathcal{T}_g(n)$ where a vertex is marked. Let $\mathcal{A}$ (resp. $\mathcal{B}$) be the subset of objects in $\mathcal{T}_g^{\mathrm{v}}(n)$ where the signed cycle containing the marked vertex has length 1 (resp. length greater than 1). Let $\gamma \in \mathcal{T}_g^{\mathrm{v}}(n)$, with $n \geq 1$. If $\gamma \in \mathcal{A}$, record the sign of the 1-cycle containing $v$ and then apply the Rémy's procedure to the plane tree with respect to $v$ (so as to delete $v$). This reduction, which does not change the genus, yields $\mathcal{A} \simeq 2 \cdot 2(2n-1)\mathcal{T}_g(n-1)$. If $\gamma \in \mathcal{B}$, let $c$ be the cycle containing the marked vertex $v$; $c$ is of the form $(v, v_1, v_2, \ldots, v_{2k})$ for some $k \geq 1$. Move $v_1$ and $v_2$ out of $c$ (the successor of $v$ becomes the former successor of $v_2$). Then apply the Rémy's procedure twice, firstly

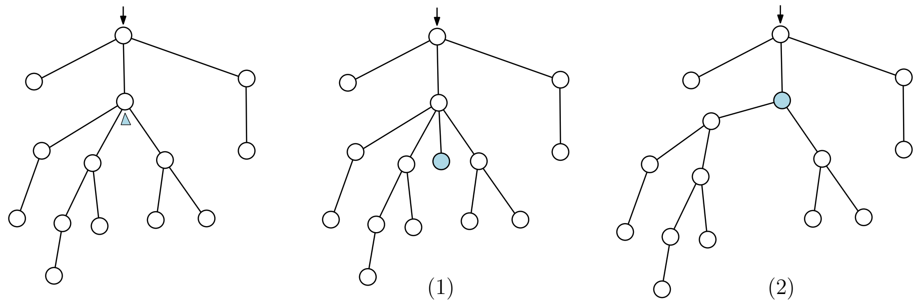


FIGURE 3. Rémy's procedure gives two ways to obtain a plane tree with $n$ edges and a marked vertex $v$ from a plane tree with $n-1$ edges and a marked corner: (1) in the first way $v$ is a leaf, (2) in the second way $v$ is a non-leaf.

with respect to $v_1$ (on a plane tree with $n$ edges), secondly with respect to $v_2$ (on a plane tree with $n-1$ edges). This reduction, which decreases the genus by 1, yields $\mathcal{B} \simeq 2(2n-1)2(2n-3)\mathcal{T}^{\mathrm{v}}_{g-1}(n-2)$, hence $\mathcal{B} \simeq 4(n-1)(2n-1)(2n-3)\mathcal{T}_{g-1}(n-2)$. Since $\mathcal{T}^{\mathrm{v}}_g(n) = \mathcal{A} + \mathcal{B}$ and $\mathcal{T}^{\mathrm{v}}_g(n) \simeq (n+1)\mathcal{T}_g(n)$, we finally obtain the isomorphism

$$(n+1)\mathcal{T}_g(n) \simeq 4(2n-1)\mathcal{T}_g(n-1) + 4(n-1)(2n-1)(2n-3)\mathcal{T}_{g-1}(n-2), \tag{3}$$

which holds for any $n \geq 1$ and $g \geq 0$ (with the convention $\mathcal{T}_g(n) = \emptyset$ if $g$ or $n$ is negative). Since $2^{n+1}\mathcal{E}_g(n) \simeq \mathcal{T}_g(n)$, we recover:

**Proposition 8** (Harer-Zagier recurrence formula [8, 9])**.** *The coefficients $\epsilon_g(n)$ satisfy the following recurrence relation valid for any $g \geq 0$ and $n \geq 1$ (with $\epsilon_0(0) = 1$ and $\epsilon_g(n) = 0$ if $g < 0$ or $n < 0$):*

$$(n+1)\epsilon_g(n) = 2(2n-1)\epsilon_g(n-1) + (n-1)(2n-1)(2n-3)\epsilon_{g-1}(n-2).$$

To the best of our knowledge this is the first proof of the Harer-Zagier recurrence formula that directly follows from a combinatorial isomorphism. The isomorphism (3) also provides a natural extension to arbitrary genus of Rémy's isomorphism (2).

3.3. **Refined enumeration of bipartite unicellular maps.** In this paragraph, we explain how to recover a formula due to Goupil and Schaeffer [7, Theorem 2.1] from our bijection. Let us first give a few definitions. A graph is *bipartite* if its vertices can be colored in black and white such that each edge connects a black and a white vertices. If the graph has a root-vertex $v$, then $v$ is required to be black; and if the graph is also connected, then such a bicoloration of the vertices is unique. From now on, a connected bipartite graph with a root-vertex is assumed to be endowed with this canonical bicoloration.

The degree distribution of a map/graph is the sequence of the degrees of its vertices taken in decreasing order (it is a partition of $2n$, where $n$ is the number of edges). If we consider a bipartite map/graph, we can consider separately the *white vertex degree distribution* and the *black vertex degree distribution*, which are two partitions of $n$.

Let $\ell, m, n$ be positive integers such that $n+1-\ell-m$ is even. Fix two partitions $\lambda$,$\mu$ of $n$ of respective lengths $\ell$ and $m$. We call $\mathrm{Bi}(\lambda,\mu)$ the number of bipartite unicellular maps, with white (resp. black) vertex degree distribution $\lambda$ (resp. $\mu$). The corresponding genus is $g=(n+1-\ell-m)/2$. It will be convenient to change a little bit the formulation of the problem and to consider *labelled maps* instead of the usual non-labelled maps: we call a *labelled map* a map whose vertices are labelled with integers $1, 2, \cdots$. If the map is bipartite, we require instead that the white and black vertices are labelled separately (with respective labels $w_1, w_2, \cdots$ and $b_1, b_2, \cdots$). The degree distribution(s) of a (bipartite) labelled map with $n$ edges can be seen as a composition of $2n$ (resp. two compositions of $n$). For $\boldsymbol{I} = (i_1, \cdots, i_\ell)$ and $\boldsymbol{J} = (j_1, \cdots, j_m)$ two compositions of $n$, we denote by $\mathrm{BiC}(\boldsymbol{I}, \boldsymbol{J})$ the number of labelled bipartite unicellular maps with white (resp. black) vertex degree distribution $\boldsymbol{I}$ (resp. $\boldsymbol{J}$). The link between $\mathrm{Bi}(\lambda,\mu)$ and $\mathrm{BiC}(\boldsymbol{I}, \boldsymbol{J})$ is straightforward: $\mathrm{BiC}(\boldsymbol{I}, \boldsymbol{J}) = m_1(\lambda)!m_2(\lambda)!\cdots m_1(\mu)!m_2(\mu)!\cdots \mathrm{Bi}(\lambda,\mu)$, where $\lambda$ and $\mu$ are the sorted versions of $\boldsymbol{I}$ and $\boldsymbol{J}$. We now recover the following formula:

**Proposition 9** (Goupil and Schaeffer [7, Theorem 2.1])**.** *:*

$$\mathrm{BiC}(\boldsymbol{I}, \boldsymbol{J}) = 2^{-2g} \cdot n \cdot (\ell+2g_1-1)!(m+2g_2-1)! \cdot \sum_{g_1+g_2=g} \sum_{\substack{p_1+\cdots+p_\ell=g_1 \\ q_1+\cdots+q_m=g_2}} \prod_{r=1}^{\ell} \frac{1}{2p_r+1}\binom{i_r-1}{2p_r} \prod_{r=1}^{m} \frac{1}{2q_r+1}\binom{j_r-1}{2q_r}. \tag{4}$$

*Proof.* For $g=0$ the formula is simply

$$\mathrm{BiC}(\boldsymbol{I}, \boldsymbol{J}) = n(\ell-1)!(m-1)!, \tag{5}$$

which can easily be established by a bivariate version of the cyclic lemma, see also [5, Theorem 2.2]. (Note, that in that case, the cardinality only depend on the lengths of $\boldsymbol{I}$ and $\boldsymbol{J}$.)

We now prove the formula for arbitrary $g$. Consider some lists $\boldsymbol{p} = (p_1, \cdots, p_\ell)$ and $\boldsymbol{q} = (q_1, \cdots, q_m)$ of nonnegative integers with total sum $g$: let $g_1 = \sum p_i$ and $g_2 = \sum q_i$. We say that a composition $\boldsymbol{H}$ refines $\boldsymbol{I}$ along $\boldsymbol{p}$ if $\boldsymbol{H}$ is of the form $(h_1^1, \cdots, h_1^{2p_1+1}, \cdots, h_\ell^1, \cdots, h_\ell^{2p_\ell+1})$, with $\sum_{t=1}^{2p_r+1} h_r^t = i_r$ for all $r$ between 1 and $\ell$. Clearly, there are $\prod_{r=1}^{\ell}\binom{i_r-1}{2p_r}$ such compositions $\boldsymbol{H}$. One defines similarly a composition $\boldsymbol{K}$ refining $\boldsymbol{J}$ along $\boldsymbol{q}$.

Consider now the set of labelled bipartite plane trees of vertex degree distributions $\boldsymbol{H}$ and $\boldsymbol{K}$, where $\boldsymbol{H}$ (resp. $\boldsymbol{K}$) refines $\boldsymbol{I}$ (resp. $\boldsymbol{J}$) along $\boldsymbol{p}$ (resp. $\boldsymbol{q}$). By (5), there are $n\cdot(\ell+2g_1-1)!(m+2g_2-1)!$ trees for each pair $(\boldsymbol{H}, \boldsymbol{K})$, so in total, with $\boldsymbol{I}$, $\boldsymbol{J}$, $\boldsymbol{p}$ and $\boldsymbol{q}$ fixed, the number of such trees is:

$$n \cdot (\ell+2g_1-1)!(m+2g_2-1)! \prod_{r=1}^{\ell}\binom{i_r-1}{2p_r} \prod_{r=1}^{m}\binom{j_r-1}{2q_r}. \tag{6}$$

As the parts of $\boldsymbol{H}$ (resp. $\boldsymbol{K}$) are naturally indexed by pairs of integers, we can see these trees as labelled by the set $\{w_r^t; 1 \le r \le \ell, 1 \le t \le 2p_r+1\} \sqcup \{b_r^t; 1 \le r \le m, 1 \le t \le 2q_r+1\}$. There is a canonical permutation of the vertices of the trees with cycles of odd sizes and which preserves the bicoloration: just send $w_r^t$ to $w_r^{t+1}$ (resp. $b_r^t$ to $b_r^{t+1}$), where $t+1$ is meant modulo $2p_r+1$ (resp. $2q_r+1$). If we additionally put a sign on each cycle, we get a C-decorated tree (with labelled cycles) that corresponds to a labelled bipartite map with white (resp. black) vertex degree distribution $\boldsymbol{I}$ (resp. $\boldsymbol{J}$). Conversely, to recover a labelled bipartite plane tree from such a C-decorated tree, one has to choose in each cycle which vertex gets the label $w_r^1$ or $b_r^1$, and one has to forget the signs of the $(n+1-g)$ cycles.

## 4. Computing Stanley character polynomials

We now consider the following enumerative problem. For $n$ a fixed integer, we would like to compute the generating series $F_n(p_1, p_2, \cdots; q_1, q_2, \cdots) = \sum \mathrm{wt}(M, \varphi)$ of pairs $(M, \varphi)$ where $M$ is a rooted bipartite unicellular map with $n$ edges, and $\varphi$ is a mapping from the vertex set $V_M$ of $M$ to positive integers, satisfying the following order condition: *for each edge $e$ of $M$, one has $\varphi(b_e) \geq \varphi(w_e)$, where $b_e$ and $w_e$ are respectively the black and white extremities of $e$.* The weight of such a pair is $\mathrm{wt}(M, \varphi) := \prod_{v \in V_M^\circ} p_{\varphi(i)} \prod_{v \in V_M^\bullet} q_{\varphi(i)}$, where $V_M^\bullet$ and $V_M^\circ$ are respectively the sets of black (resp. white) vertices of $M$.

Our motivation comes from representation theory of the symmetric group. This topic is linked to map enumeration by the following formula [4]. Let $\boldsymbol{p} = p_1, \cdots, p_r$ and $\boldsymbol{q} = q_1, \cdots, q_r$ be two finite lists of positive integers of the same length. Then the evaluation $F_n(p_1, \cdots, p_r, 0, \cdots; q_1, \cdots, q_r, 0, \cdots)$ of the generating series considered above is equal to $L(L-1)\cdots(L-n+1)\hat{\chi}^\lambda((1\ 2\ \cdots\ n))$, where:

- $\lambda$ is the partition with $p_1$ parts equal to $q_1 + \cdots + q_r$, $p_2$ parts equal to $q_2 + \cdots + q_r$, and so on...
- $L = \sum_{1 \leq i \leq j \leq r} p_i q_j$ is its number of boxes ;
- $\hat{\chi}^\lambda$ is the normalized character of the irreducible representation of $S_L$ associated to $\lambda$;
- $(1\ 2\ \cdots\ n)$ is an $n$-th cycle seen as a permutation of $S_L$ (if $n > L$, it is not defined but, as the numerical factor is 0, it is not a problem).

Our main bijection allows us to express the generating series $F_n$ in terms of the corresponding generating series for plane trees. A. Rattan has proved [12] that this generating series is the $n+1$-th free cumulant $R_{n+1}$ of the transition measure of $\lambda$ (as $\lambda$ depends on $\boldsymbol{p}$ and $\boldsymbol{q}$, $R_{n+1}$ can be seen as a series in $\boldsymbol{p}$ and $\boldsymbol{q}$). Free cumulants have become in the last few years an important tool in (asymptotic) representation theory of the symmetric groups, see for example the work of P. Biane [2].

Let us define an operator $D$ by $D(x^k) := \sum_{g \geq 0} c_g(k) x^{k-2g} = k! \sum_{r=1}^k 2^r \binom{k-1}{r-1}\binom{x}{r}$, $D$ being extended multiplicatively to monomials in distinct variables, and then extended linearly to multivariate polynomials and series (in particular, series in the variables $\boldsymbol{p}$ and $\boldsymbol{q}$).

**Theorem 11.** *For any $n \geq 1$, one has $2^{n+1} F_n = D(R_{n+1})$.*

*Proof.* A pair $(M, \varphi)$ as above corresponds by the bijection of Section 2 to a bipartite $C$-decorated tree $T$, together with a function $\varphi : V_T \to \mathbb{N}$ which fulfills the order condition and such that all vertices in a given cycle have the same image by $\varphi$. The result follows directly. □

The free cumulant $R_{n+1}$ is the compositional inverse of an explicit series [12]. Hence Theorem 11 gives an efficient, easily implemented way of computing Stanley character polynomials $F_n$.

## References


[1] O. Bernardi. An analogue of the Harer-Zagier formula for unicellular maps on general surfaces. arXiv:1011.2311, 2010.

[2] P. Biane. Representations of symmetric groups and free probability. *Adv. Math.*, 138(1):126–181, 1998.

[3] G. Chapuy. A new combinatorial identity for unicellular maps, via a direct bijective approach. *Advances in Applied Mathematics*, 47(4):874 – 893, 2011.

[4] V. Féray. Stanley's formula for characters of the symmetric group. *Ann. Comb.*, 13(4):453–461, 2010.

[5] I. P. Goulden and D. M. Jackson. The combinatorial relationship between trees, cacti and certain connection coefficients for the symmetric group. *European J. Combin.*, 13(5):357–365, 1992.

[6] I. P. Goulden and A. Nica. A direct bijection for the Harer-Zagier formula. *J. Comb. Theory, Ser. A*, 111(2):224–238, 2005.

[7] A. Goupil and G. Schaeffer. Factoring $n$-cycles and counting maps of given genus. *European J. Combin.*, 19(7):819–834, 1998.

[8] J. Harer and D. Zagier. The Euler characteristic of the moduli space of curves. *Invent. Math.*, 85:457–486, 1986.

[9] S. K. Lando and A. K. Zvonkin. *Graphs on Surfaces and Their Applications.* Springer, 2004.

[10] B. Lass. Démonstration combinatoire de la formule de Harer-Zagier. *C. R. Acad. Sci. Paris*, 333, Série I:155–160, 2001.

[11] A. Morales and E. Vassilieva. Bijective enumeration of bicolored maps of given vertex degree distribution. *DMTCS Proceedings*, AK:661–672, 2009.

[12] A. Rattan. Stanley's character polynomials and coloured factorizations in the symmetric group. *J. Combin. Theory Ser. A*, 114(4):535–546, 2008.

[13] J.-L. Rémy. Un procédé itératif de dénombrement d'arbres binaires et son application à leur génération aléatoire. *RAIRO Inform. Théor.*, 19(2):179–195, 1985.

[14] G. Schaeffer and E. A. Vassilieva. A bijective proof of Jackson's formula for the number of factorizations of a cycle. *J. Comb. Theory, Ser. A*, 115(6):903–924, 2008.

[15] T. R. S. Walsh and A. B. Lehman. Counting rooted maps by genus. I. *J. Combin. Theory Ser. B*, 13:192–218, 1972.

G. Chapuy: CNRS, LIAFA - UMR 7089, Université Paris 7, 75205 Paris Cedex, France
*E-mail address*: `chapuy@liafa.univ-paris-diderot.fr`

V. Feray: CNRS, LaBRI - UMR 5800, Université de Bordeaux, 33405 Talence Cedex, France
*E-mail address*: `feray@labri.fr`

É. Fusy: CNRS, LIX - UMR 7161, École Polytechnique, 91128 Palaiseau Cedex, France
*E-mail address*: `fusy@lix.polytechnique.fr`